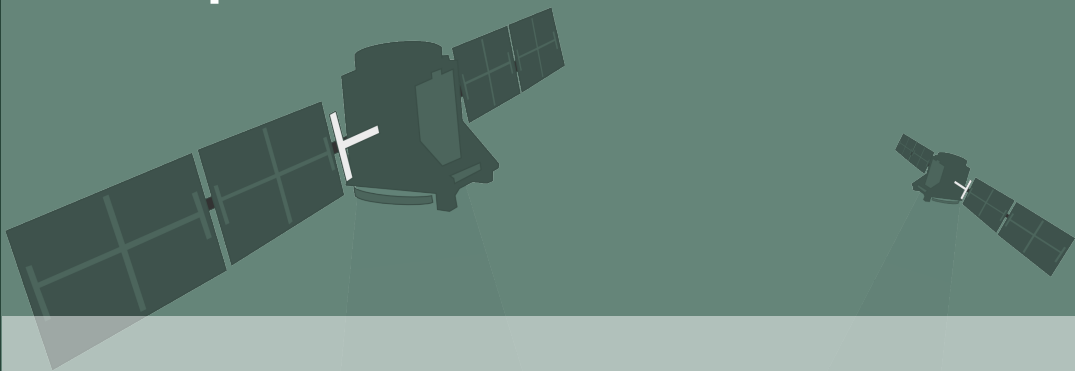




Geografische Empfehlungssysteme

Christian Matyas



University
of Bamberg
Press

15 Schriften aus der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich- Universität Bamberg

Schriften aus der Fakultät Wirtschaftsinformatik
und Angewandte Informatik der Otto-Friedrich-
Universität Bamberg

Band 15

Geografische Empfehlungssysteme

von Christian Matyas

Bibliographische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Informationen sind im Internet über <http://dnb.ddb.de/> abrufbar

Diese Arbeit hat der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität als Dissertation vorgelegen.

1. Gutachter: Prof. Dr. Christoph Schlieder

2. Gutachter: Prof. Dr. Ute Schmid

Tag der mündlichen Prüfung: 12.11.2013

Dieses Werk ist als freie Onlineversion über den Hochschulschriften-Server (OPUS; <http://www.opus-bayern.de/uni-bamberg/>) der Universitätsbibliothek Bamberg erreichbar. Kopien und Ausdrücke dürfen nur zum privaten und sonstigen eigenen Gebrauch angefertigt werden.

Herstellung und Druck: docupoint, Magdeburg

Umschlaggestaltung: University of Bamberg Press, Andra Brandhofer

© University of Bamberg Press Bamberg 2014
<http://www.uni-bamberg.de/ubp/>

ISSN: 1867-7401

ISBN: 978-3-86309-204-7 (Druckausgabe)

eISBN: 978-3-86309-205-4 (Online-Ausgabe)

URN: urn:nbn:de:bvb:473-opus4-60852

Zusammenfassung

Mobile Geräte werden immer häufiger mit Sensoren zur Bestimmung der eigenen Position ausgestattet, zum Beispiel mit GPS. Mit Hilfe der Ortsinformationen dieser Sensoren können beispielsweise moderne Bildmanagementanwendungen digitale Fotos automatisch nach geografischen Regionen gruppieren oder passende Schlagworte generieren. Dies führt unter anderem zu einer besseren Suchbarkeit dieser digitalen Daten.

Grundsätzlich geben Ortsinformationen in digitalen Fotos nicht nur Hinweise auf das Foto selbst, sondern machen auch sichtbar, welche geografischen Entscheidungen der Fotograf bei deren Erstellung getroffen hat. Diese Arbeit nutzt diese Entscheidungen für die Berechnung von weiteren Empfehlungen für den Nutzer, beispielsweise einer Bildmanagementanwendung. Ein konkreter Anwendungsfall lautet folgendermaßen: Einem Nutzer sollen für eine frei wählbare geographische Region (z.B. einer Stadt), mehrere Bilder empfohlen werden, die zum einen typisch für diese Region sind, zum anderen aber auch für ihn persönlich interessant sein könnten. Um diese geografischen Mehr-Objekt-Empfehlungen zu berechnen, wurde ein neuartiger Algorithmus entwickelt, der zunächst die Ortsinformationen aller Nutzer zu einem geografischen Modell bündelt. Auf Grundlage dieser prototypischen Konzeptualisierung von einzelnen Regionen, können dann typische Bilder empfohlen werden. Weiterhin werden diese geografischen Modelle in einem zweiten Schritt für die zusätzliche Gewichtung der einzelnen geografischen Entscheidungen der Nutzer verwendet, um über den Ansatz

eines kollaborativen Filters zu einer persönlichen Empfehlung zu gelangen. Dazu wurden mehrere Verfahren entwickelt und miteinander verglichen.

Diese Arbeit ist im Rahmen des europäischen Projektes Tripod entstanden, für das der entwickelte geografische Empfehlungsalgorithmus als Softwaremodul prototypisch implementiert wurde. Damit wurden die Empfehlungen mit Hilfe von georeferenzierten Bildern evaluiert, die auf den Online-Galerien Panoramio.com und Flickr.de veröffentlicht wurden. Durch die Auswertung der geografischen Informationen und der daraus berechneten Ortsmodelle, ließen sich deutlich präzisere Empfehlungen vorschlagen, als mit anderen bekannten Empfehlungsverfahren.

Inhaltsverzeichnis

1	Einleitung	8
1.1	Beiträge	12
1.2	Gliederung	13
2	150 Jahre Fotografie	17
2.1	You press the button, we do the rest	17
2.2	Neue Möglichkeiten für Empfehlungssysteme	28
2.3	Ausgangsfragen	32
3	Geografische Bildsammlungen	34
3.1	Übersicht über aktuelle Bildsammlungen	35
3.1.1	www.geograph.org	36
3.1.2	www.panoramio.com	40
3.1.3	www.flickr.com	41
3.2	Fokus der einzelnen Bildsammlungen	42
3.3	Arbeiten zu räumlichen Benutzerdaten	45
3.4	Arbeiten mit geographischen Bildsammlungen	46
3.5	Inhaltsbasierte Verfahren	50
4	Empfehlungssysteme	52
4.1	Informationsbedarf und Informationsfilterung	52
4.2	Weitere Überlegungen zu Typikalität und Variabilität	63
4.3	Systematisierung von Empfehlungssystemen	64
4.4	Überblick über die einzelnen Ansätze	67
4.4.1	Kollaboration	67

4.4.2	Gegenstands-basierte Filterung	73
4.4.3	Demographische Filterung	74
4.4.4	Hybride Empfehlungssysteme	76
4.5	Bestehende Empfehlungssysteme für Bilder	78
4.6	Ortsbasierte Empfehlungssysteme	79
4.7	Geographische Empfehlungssysteme	80
4.8	Zusammenfassung	81
5	Typische Bildgruppen	82
5.1	Anwendungsfall: typische Bildgruppen einer bestimmten Region	83
5.2	Bildgruppen in der Wirklichkeit	89
5.3	Kollaborative Semantik in unstrukturierten Bildsammlungen	94
5.4	Räumliche Entscheidungen bei georeferenzierten Bildern	98
5.4.1	Gruppierung von räumlichen Entscheidung . . .	102
5.4.2	Popularität eines CPV	109
5.4.3	Die Heatmap als geografisches Modell einer Region	113
5.5	Kollaborative Semantik in Texten	121
5.5.1	Untersuchung von Konzeptualisierung einer Region in Texten	124
5.5.2	Untersuchung einer Folksonomy	130
5.5.3	Kookkurrenzanalyse eines Textkorpus	134
5.6	Zusammenfassung	137
6	Persönliche Bildempfehlungen	139
6.1	Anwendungsfall: Personalisierte Bildgruppen	139
6.2	Generischer kollaborativer Empfehlungsansatz	145
6.3	Geografischer Empfehlungsansatz	149
6.3.1	Schritt 1: Benutzer-Feedback	149

6.3.2	Schritt 2: Ähnlichkeitsberechnung mit Hilfe von räumlichen Entscheidungen	156
6.3.3	Schritt 3: Nachbarschaftswahl mit Hilfe von räumlichen Entscheidungen	174
6.3.4	Schritt 4: Aggregation der räumlichen Entscheid- ungen	175
7	Tripod GeoRecommender	179
7.1	Projekt Tripod	179
7.2	Heatmapper Prototyp	189
7.2.1	Webservice	192
7.3	GeoRecommender Prototyp	194
7.3.1	Einbettung in eine Bildergalerie am Beispiel des Tripod-Prototypen	200
7.3.2	Einbettung in eine Anwendung des Social Web am Beispiel von Facebook	204
7.3.3	Auswertungstool GeoRec in Java	211
7.4	Erweiterungspunkte des Empfehlungssystems	215
8	Auswertung	216
8.1	Vergleich typischer Bildgruppen zu Patchwork-Postkarten	217
8.2	Tripot	220
8.2.1	Aufgabenstellung	220
8.2.2	Ausgangsfragen	225
8.2.3	Methoden	226
8.2.4	Baseline	227
8.2.5	Ergebnisse der Analyse	229
8.2.6	Ergebnis	238
8.3	Offline-Evaluierung	238
8.3.1	Vergleich und mit Hilfe eines passenden Evalua- tionsmaßes	241

8.3.2	Evaluationsmaße	241
8.4	Offline-Evaluierung - Panoramio/Flickr	248
8.4.1	Evaluationsmethode	248
8.4.2	Evaluationsdaten	252
8.4.3	Evaluationsergebnisse	254
9	Ergebnisse und Ausblick	260
9.1	Ergebnisse	260
9.2	Ausblick	263
9.2.1	Ausblick auf weitere Analyse von raum-zeitlichen Entscheidungen	263
9.2.2	Ausblick auf Sensorik in Kameras	264
9.2.3	Ausblick auf Evaluierungsframeworks	267

1 Einleitung

Den Menschen stärker in die inhaltliche Gestaltung des Internets einzubinden, ist immer noch eines der Ziele, die man sich für das Web 2.0 gesetzt hat. Dabei geht es in den meisten Fällen nicht darum, komplexes Wissen abzufragen, wie es bei Wikipedia¹ versucht wird, sondern Menschen nach ihrer Meinung zu fragen. Dies macht die Nutzung des Internets zu einer sehr viel persönlicheren Erfahrung und bindet die Nutzer auch an die entsprechenden Dienste. Beteiligt sich ein Nutzer durch ständiges Feedback an einem System, kann dieses im Umkehrschluss speziell zugeschnittene Inhalte für diesen Nutzer bereitstellen bzw. weitere Empfehlungen aussprechen. Ein Paradebeispiel für ein derartiges System ist nicht zuletzt die E-Commerce Plattform Amazon². Alleine die Startseite bietet eine Vielzahl von personalisierten Inhalten an.

Auch durch Konzepte, wie beispielsweise Cloud Computing, werden immer mehr Anwendungen nicht mehr zentral beim Endnutzer ausgeführt. Das sorgt dafür, dass die Anwendungsdaten nicht mehr beim Endnutzer liegen. Für den Anwender selbst verschwimmen die Grenzen zwischen dem eigenen Arbeitsplatz und dem eigentlichen Ausführungsort des Dienstes. Durch diesen Wandel haben die Betreiber Zugriff auf die Nutzerdaten, können diese analysieren und personalisierte Inhalte an den Nutzer kommunizieren. So schaltet Google in seinem Email-

¹ <http://www.wikipedia.org>

² <http://www.amazon.com>

Service (Gmail³) passende Werbung entsprechend der Schlagwörter, die sich in der E-Mail befindet, die der Nutzer gerade geöffnet hat.

Die technische Entwicklung einer komplett vernetzten Welt machen es möglich, dass der Mensch immer gläserner und transparenter wird. Bewusst und/oder unbewusst gibt er immer mehr über sich preis. In beiden Fällen kann das Verhalten des Nutzers analysiert werden, um ihm dadurch ein persönlicheres Anwendungsereignis anzubieten. Abgesehen von rechtlichen Bedenken über die Speicherung dieser personenbezogenen Daten, eröffnen sie dennoch auch neue Anwendungsfälle.

Ein Aspekt des Nutzers steht dabei bisher noch nicht im primären Fokus der Forschung: sein räumliches Verhalten. In dieser Arbeit wird das räumliche Verhalten über diskrete Punkte im Raum bestimmt. Eine Intention, den eigenen Standpunkt im Raum zu veröffentlichen, kann zum Beispiel die persönliche Vorliebe für diesen Ort sein. Eine Quelle, aus der man räumliches Verhalten rekonstruieren kann, sind georeferenzierte Bildsammlungen auf die viele Firmen bereits Zugriff haben. Georeferenzierte Bilder haben dabei einen besonderen Reiz auf Firmen, die bereits eigene Kartenprodukte oder Bildmanagementsysteme auf dem Markt etabliert haben. Die prominentesten Vertreter dabei sind:

1. Google: bietet selbst die Kartensoftware GoogleEarth⁴ und die Webanwendung GoogleMaps⁵ an und kaufte 2007 die spanische Firma Panoramio⁶ mit mehreren Millionen georeferenzierten Bilder auf. Zusätzlich bietet Google über Picasa⁷ eine eigene Bild-

³ <http://mail.google.com>

⁴ <http://www.google.de/earth>

⁵ <http://maps.google.de>

⁶ <http://www.panoramio.com/>

⁷ <http://www.google.de/picasa/>

managementsoftware an, die es ebenfalls ab Version 3.5 erlaubt Bilder zu verorten.

2. Yahoo: eigenes Kartenprodukt mit Yahoo!Maps⁸. Die Firma kauft im Jahr 2005 die Foto-Plattform Flickr⁹ von einem kleinen kanadischen Startup, die zu einer der bekanntesten Foto-Sharing-Seiten erwachsen ist. Yahoo!Maps wird kurz darauf in die Flickr-Plattform integriert und bietet momentan über 178 Millionen georeferenzierte Bilder an (Stand: März 2012).
3. Microsoft: startet mit Virtual Earth in den webbasierten GIS Bereich, wechselte zuerst den Namen in Live Maps und später in Bing Maps¹⁰. 2008 stellt Microsoft das System Photosynth¹¹ vor, das es ermöglicht, Bilder im dreidimensionalen Raum anzuzeigen und miteinander zu verbinden. Ebenfalls 2008 gibt Microsoft das Photosync-System frei und sammelt damit georeferenzierte Bilder. Momentan können diese Bilder über Bing Maps gefunden werden. Zusätzlich integriert Microsoft Photosynth in die eigene Windows Live Plattform, was ermöglicht lokale Bilder sehr unkompliziert in den Datenbestand von Photosynth hinzuzufügen. Weltweit bietet Microsoft momentan (März 2012) Zugriff auf immer mehr als 300.000 sogenannten PhotosynthsTM mit über 12 Millionen Bildern. Im September 2010 waren vorerst 40.000 Synth zugänglich, die aus ungefähr zwei Millionen Bildern bestanden.

Diese Entwicklungen deuten an, dass sehr viel Geld in die Beschaffung von georeferenzierten Bildsammlungen investiert wird. Hinzu kommt, dass die Vertriebswege, gerade für digitale Bilder immer vielfältiger

⁸ <http://maps.yahoo.com>

⁹ <http://www.flickr.com>

¹⁰ <http://www.bing.com/maps/>

¹¹ <http://photosynth.net/>

werden. Der Nutzer kann durch Online-Dienste nicht nur seine Datensammlungen archivieren, sondern diese auch mit anderen teilen. Reine Bildanzeigergeräte, wie digitale Bilderrahmen oder auch vernetzte Hintergrundbilder, bieten dem Nutzer die Möglichkeit Bilder auf diese Geräte in seinem eigenen Wohnzimmer auszuliefern — Abbildung 1.1 illustriert dies. Die momentanen Anwendungsfälle bzw. Marketingstrategien sehen allerdings vor, dass die nötige Bildauswahl vom Nutzer selbst, seinen Freunden oder Familienangehörigen bewältigt wird. Für diese Geräte bieten Empfehlungsansätze eine neue Möglichkeiten, immer wieder neue Bilder anzuzeigen, die automatisch nach dem Geschmack des Besitzers ausgewählt werden. Zum Beispiel wäre ein Abonnement von Bildern aus interessanten Urlaubsorten über einen solchen Dienst denkbar.

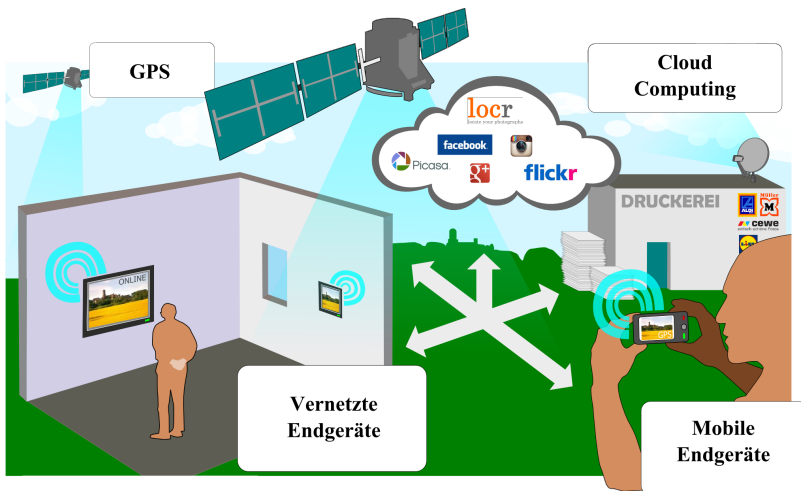


Abbildung 1.1: Vernetzte Endgeräte und neue Dienstleistungen

1.1 Beiträge

Das Ziel dieser Arbeit ist es, zu erforschen, wie mit Hilfe von georeferenzierten Bildsammlungen eine Empfehlung für einen Nutzer berechnet werden kann, vergleichbar den Empfehlungen, die zum Beispiel Amazon und Google auf Grundlage von Bewertungen und Textanalysen errechnen. Als zentraler technischer Beitrag wird ein Empfehlungssystem auf Basis von georeferenzierten Bildern vorgestellt, das insbesondere den aktuellen Standort eines Nutzers beziehungsweise die Standorte, an denen er Bilder erstellt hat, bei der Erstellung der Empfehlungen berücksichtigt.

Ausgangspunkt dieser Arbeit ist, die Definition von zwei neuen Anwendungsfällen für dieses geografische Bild-Empfehlungssystem. Zunächst sollen Gruppen von Bildern empfohlen werden, die einen Ort illustrieren, indem die typischsten Ansichten berechnet werden. Im zweiten Schritt sollen solche Bilder empfohlen werden, die auch für einen bestimmten Nutzer von Interesse sind. Diese Anwendungsfälle bringen neue Anforderungen an die Zusammenstellung der Empfehlungsmenge mit sich. Diese werden in den entsprechenden Kapiteln herausgearbeitet. Die grundlegende Idee dieser Arbeit ist, dass der Nutzer durch den Umgang mit georeferenzierten Informationen (in diesem Fall durch das Erstellen eines Bildes) bestimmte Standorte positiv bewertet.

Diese räumliche Entscheidungen aus einem geographischen Bezugsrahmen werden dann dazu genutzt, Empfehlungen für einen neuen geographischen Kontext zu erzeugen. Mit räumlichem Kontext ist zum Beispiel eine bestimmte Stadt gemeint. Um dies zu erreichen

- werde ich die Gesamtsumme aller räumlichen Entscheidungen analysieren, damit ein geografisches Modell des Raumes erstellt werden kann. Dieses Modell gibt an, welche Orte typisch oder untypisch für einen festgelegten räumlichen Kontext sind.

- werde ich auf Grundlage dieses Modells verschiedene Maße erarbeiten, mit deren Hilfe man die Ähnlichkeit von zwei Nutzern durch dieses geografische Zusatzwissen weiter differenzieren kann. Die berechnete geografische Ähnlichkeit wird genutzt, um die Orte des Modells kollaborativ neu zu bewerten.
- werde ich die Ergebnisse durch die verschiedenen Ähnlichkeitsberechnungen gegeneinander evaluieren.

Die erarbeiteten räumlichen Ähnlichkeitsmaße sind auch für andere Anwendungsfelder interessant, in denen ebenfalls räumliche Positionen von Nutzern protokolliert werden. Nach der großen Flut an sozialen Netzwerken entstehen immer mehr ortsbasierte soziale Netzwerke, die den Nutzer in besonderen Bezug zu seiner aktuellen Position stellen. Jede nutzergenerierte Information beinhaltet daher auch eine Positionsangabe. Die hier vorgestellten Nutzerähnlichkeiten können auch auf diese räumlichen Informationsströme angewendet werden, um dadurch neue Beziehungen zwischen Personen im sozialen Netzwerk zu initiieren oder auch um relevante Informationen mit Hilfe des räumlichen Verhaltens zu filtern.

1.2 Gliederung

Die Arbeit gliedert sich wie folgt:

Kapitel 2 gibt eine Übersicht über die Entwicklung der Fotografie, die sich in zwei Bereichen besonders rasant entwickelt hat. Zum einen ist es mit jeder Evolution der Fotoapparate immer einfacher geworden, Bilder zu erstellen. Dies liegt vor allem an den Sensoren in den Kameras, die immer genauer werden und so die Kameras immer intelligenter werden. Zum anderen ermöglicht es die Entwicklung der Speicherkonzepte, angefangen mit einzelnen Bildplatten, über Filmstreifen bis hin zur digitalen Speicherung, immer mehr Bilder zu produzieren. Der Nachteil

dieser neuen Möglichkeiten lässt sich mit einer einfachen Gleichung ausdrücken: einfacher fotografieren = mehr fotografieren = mehr verwalten. Einen möglichen Weg, mit diesem erhöhten Verwaltungsaufwand umzugehen, bieten Empfehlungssysteme.

Kapitel 3 gibt eine Übersicht über georeferenzierte Datenquellen. Dabei ist beeindruckend, dass hinter den meisten Sammlungen mittlerweile die großen Suchmaschinenbetreiber zu finden sind (Google, Yahoo! und Microsoft). Diese Datensammlungen sind Beispiele für Volunteered Geographic Information (VGI) (Goodchild, 2007). In diesem Kapitel werden mögliche Analyseansätze dieser Daten abgehandelt.

Kapitel 4 gibt einen Überblick über die Forschung im Bereich Empfehlungssysteme. Neben der ausführlichen Beschreibung der verschiedenen Empfehlungsansätze in der üblichen Kategorisierung wird auch ein Problem aufgezeigt, das in vielen Ansätzen zu finden ist: die fehlende Variabilität in den Empfehlungen.

Kapitel 5 beschäftigt sich mit der Frage, welche Orte im geografischen Raum besonders populär sind. Dabei wird ein Verfahren vorgestellt, mit dessen Hilfe georeferenzierte Bildsammlungen auf diese Fragestellung hin analysiert werden können. Berechnet wird ein Modell des Raumes, das eine konkrete Definition des Ortes aufweist und diese Orte nach Popularität sortiert. Es werden die Ergebnisse, die durch die festgehaltenen visuellen Eindrücke der Nutzer entstanden sind, mit verschiedenen Möglichkeiten der Textanalyse verglichen. Beide Analysen kommen dabei aufgrund der Ausdrucksmächtigkeit der unterschiedlichen Kommunikationsformen (Text und Bild) und der Art der Datenerhebung zu anderen Ergebnissen. Aufgezeigt werden in diesem Kapitel auch die Einschränkungen und die Eigenarten des geografischen Verfahrens.

Kapitel 6 beschäftigt sich damit, die Ergebnisse aus Kapitel 5 für ein Verfahren zu verwenden, das die Ergebnisse weiter personalisiert. Dieses Verfahren bedient sich dabei dem Grundgerüst eines nutzerbasierten

kollaborativen Filters, wie es in Kapitel 4 definiert wird. Ziel ist es, das geografische Modell aus Kapitel 5 zu verwenden, um die einzelnen Orte durch die Vorlieben von ähnlichen Nutzern neu zu gewichten. Dadurch lässt sich eine personalisierte Empfehlung berechnen.

Kapitel 7 befasst sich mit dem Project Tripod¹² (Fan u. a., 2010), in welchem die beschriebenen Ansätze aus Kapitel 5 und Kapitel 6 für ein vitales System implementiert wurden. Dieses Kapitel erörtert zudem die weiteren Anwendungsmöglichkeiten, die geografische Metadaten in Bildern ermöglichen. Bilder mit Texten, wie Titel oder Schlagwörter, zu beschreiben, war dabei das Hauptaugenmerk des Projekts Tripod. Zusätzlich sollten neue Filteransätze erforscht werden, mit denen man auf diese georeferenzierten Datenquellen zugreifen kann. Es wurde im Rahmen des Projekts Tripod nach Möglichkeiten geforscht, wie man zum einen gezielt nach Bildern über ihre geografischen Zusatzinformationen suchen kann, zum anderen, welche neuen Empfehlungsansätze mit diesen Informationen möglich sind.

In Kapitel 8 wird mit Hilfe von verschiedenen Evaluationsverfahren die Qualität der Empfehlungen untersucht, was insbesondere bei nutzerbasierten Empfehlungsverfahren nicht immer einfach ist. Evaluiert wurden die Ergebnisse mit einer Kreuzevaluierung, was auch die gängigste Methode bei dieser Art von Empfehlung darstellt. Durch das gezielte Ausblenden von Bildern aus einer bestimmten Region war es möglich, Empfehlungen für diese Region auf Offlinedaten zu überprüfen. Der Versuch wurde auf zwei Bildsammlungen mit unterschiedlichen thematischen Schwerpunkten durchgeführt, hierbei wird auch auf die Schwierigkeiten bei der Anwendung des geografischen Empfehlungssystems eingegangen.

¹² <http://www.projecttripod.org>

Kapitel 9 schließt die Arbeit mit einer Zusammenfassung und einem Ausblick auf Forschungsfragen und Anknüpfungspunkte für weitere Arbeiten ab.

2 150 Jahre Fotografie

2.1 You press the button, we do the rest

Mit dem Fotoapparat Kodak Nr. 1 legte George Eastmann im Jahre 1888 den Grundstein für die moderne analoge Fotografie. Es war die erste Kamera, die mit einem Rollfilm auf Papierbasis ausgestattet war und somit hundert Bilder auf einem Medium speichern konnte. Dadurch war es möglich um einiges schneller, einfacher und mobiler Bilder aufzunehmen, im Gegensatz zu den bis dahin üblichen Apparaten, die Glasplatten als Aufzeichnungsmedium verwendeten. Mit seiner Firma Kodak machte er das Geschäft mit Fotografien zum Massenmarkt. Er baute dabei vor allem auf geschicktes Marketing, indem er auf erschwingliche Kameras für jedermann und einen besonderen Service bei der Filmentwicklung baute. Dieser Service lässt sich am besten mit dem Motto von Kodak selbst beschreiben „Sie drücken den Knopf, wir machen den Rest“ (im Original: „You press the button, we do the rest“).

1969 wurde das lichtempfindliche elektronische Bauteil mit der Bezeichnung Charge-couple Device (CCD) zur Datenspeicherung erfunden. Als Nebenerscheinung konnte ein CCD zudem auch als Sensor genutzt werden, um ein zweidimensionales Bild zu erfassen. Dadurch war die technische Grundlage für die digitale Fotografie geschaffen. Die beiden Erfinder Willard Boyle und George E. Smith wurden dafür 2009 mit dem Nobelpreis für Physik ausgezeichnet. Zwischen 1969 und 2009 hat der CCD-Sensor eine neue Evolution der Fotografie ermöglicht. 1981 wur-

de von Sony das erste Bildaufnahmegerät mit digitaler Speicherung für den Massenmarkt mit dem Namen Mavica veröffentlicht. Auch andere namhafte Hersteller folgten mit eigenen digitalen Geräten (Canon ION – 1984, Apple Quicktake – 1994, Nikon Coolpix – 1997). Die direkte Speicherung in einem digitalen Format hatte für den Kunden vor allem einen großen Mehrwert: Die Bilder waren sofort sichtbar und konnten auf jedem Computer angezeigt oder verarbeitet werden. Eine weitere positive Entwicklung für die digitale Fotografie waren die fallenden Kosten für Speicher bei gleichzeitiger Zunahme der Speichergröße. Allerdings war dadurch vieles, was Kodak als Service-Infrastruktur aufgebaut hatte, überflüssig und musste durch neue Konzepte ersetzt werden.

Alle Entwicklungen der letzten 135 Jahre sorgen dafür, dass immer mehr Bildaufnahmen erzeugt werden. Gleichzeitig stieg aber auch der Aufwand, Bildsammlungen zu organisieren und relevante Bilder aus der Masse auffindbar zu machen. Im Privatbereich lässt sich eine gewisse Ordnung und Übersicht noch mit geschickten Ordnerstrukturen aufrechterhalten. Allerdings ist diese Lösung für umfangreichere Bildsammlungen, zum Beispiel von professionellen Bildagenturen, nicht mehr denkbar. Die beiden führenden Bildarchive Getty Image¹³ und Corbis¹⁴ verwalten zusammen um die 170 Millionen Bildaufnahmen. Um diese Bildmassen zugänglich zu machen, bedient man sich klassischen bibliothekarischen Methoden, wie der Kategorisierung und Sortierung nach zusätzlichen Informationen über die Bilder. Sinnvolle Kategorien sind etwa: der Name des Fotografen bzw. des Rechteinhabers, der Typ der Aufnahme (Landschaftsaufnahme, Detailaufnahme, Portrait, ...) oder auch das Entstehungsjahr und der Ort. Diese Daten über Daten nennt man auch Metadaten oder Metainformationen.

¹³ <http://www.gettyimages.de/>

¹⁴ <http://www.corbisimages.com/>

Definition 2.1 *Metadaten sind Daten, die Information über andere Daten geben. Man nennt diese Daten auch Sekundärdaten im Gegensatz zu den Primärdaten. Bei einem Bild bestehen die Primärdaten aus der digitalen Repräsentation des Bildes (Pixel, Vektoren, ...). Sekundärdaten sind zum Beispiel die Beschreibung des Inhalts oder der Name des Urhebers.*

Ein dauerhaftes Problem ist nach wie vor die automatische Erzeugung dieser Zusatzinformationen. Gerade für Bilder sind Metainformationen schlecht vom Computer berechenbar, im Gegensatz zu anderen digitalen Kommunikationsmitteln. Das Kommunikationsmittel Text kann zum Beispiel einfacher vom Computer analysiert werden (Markov, 2007). Das liegt vor allem an der beschreibbaren Syntax von Text, die es erlaubt, aus Texten sinntragende Einheiten herauszulesen. Bei einem Text besteht die kleinste sinntragende Einheit aus einem Buchstaben. Die Buchstaben wiederum setzen sich durch festgeschriebene Regeln zu Wörtern und Wörter zu Sätzen zusammen und somit zu einer funktionellen Einheit, die ein Computer weiter analysieren kann. In Form von formalen Sprachen (zum Beispiel Programmiersprachen) kann er sie sogar verstehen.

In Analogie dazu kann man für Rasterbilder sagen, dass ein einzelnes Pixel (= picture elements) die kleinste festgelegte Einheit bildet. Als nächst höhere Abstraktionsebene folgen einzelne Regionen oder Gruppierungen von Pixeln, in denen der Mensch Objekte im Bild erkennen kann, man spricht dabei von Segmentierung. Diese Regionen mit einem entsprechenden Algorithmus aufzufinden, ist sehr berechnungsintensiv. Bei Text markiert eine Leerstelle oder ein Satzzeichen das Ende eines Wortes, die Grenzen bei Bildern sind eher fließend. Hinzu kommt, dass Objekte im Bild sich meist aus mehreren Regionen zusammensetzen, was die Erkennung eines Objektes nochmals schwieriger macht. Es wird zwar an Lösungen für die Objekterkennung in Bildern geforscht,

zum Beispiel über Kantendetektion (Belongie u. a., 2001) oder über den Wechsel der Textur (Lavrenko u. a., 2003), allerdings sind diese inhaltsbasierten Verfahren für Bilder noch recht unzuverlässig. Im praktischen Umgang mit Bildern werden diese Verfahren kaum angewendet. Für die Verschlagwortung von Bildern ist daher der Mensch immer noch unersetzlich. Ein Ansatz, dieses Problem zu bearbeiten, stützt auf der Fragestellung, wie Menschen sich motivieren lassen, solche Arbeiten freiwillig zu tätigen; das ESP Spiel ist dabei eines der bekanntesten Beispiele (Robertson u. a., 2009),(Seneviratne und Izquierdo, 2010).

Ein inhaltsbasiertes Verfahren, das hinreichend stabil arbeitet, ist die Gesichtserkennung. Hsu u. a. (2002) und Yang u. a. (2002) geben eine gute Übersicht über bestehende Lösungsansätze. Gesichtserkennung ist heutzutage in immer mehr Bildverwaltungssoftware zu finden. Auch in Kameras wird die Gesichtserkennung als Unterstützung für die passende Einstellung der Kameraoptik verwendet. Somit ist die Erkennung von Regionen im Bild, in denen wahrscheinlich ein Gesicht zu erkennen ist, schon heute möglich.

An die Segmentierung und Kategorisierung der Gesichter in den Bildern schließt sich allerdings gleich das nächste Problem an: die Benennung der Regionen. Im Fall von Gesichtern gruppiert man Bilder mit ähnlichen Gesichtern und vergleicht sie mit bereits bekannten Gesichtern. In den meisten Fällen wird man dennoch den Nutzer entscheiden lassen, ob die berechnete Ähnlichkeit der Gesichter auch der Wirklichkeit entspricht. Ähnliche Ansätze findet man auch für unbestimmte Regionen, wobei es im Allgemeinen zu vielen Doppeldeutigkeiten kommt, zum Beispiel bei der Erkennung, ob es sich um einen Vogel, ein Flugzeug oder doch ein anderes fliegendes Objekt handelt.

Weitere Verfahren verwenden Kontextinformation, wie die Zeit und den Ort an dem ein Bild aufgenommen wurde, um beschreibende Informationen über ein Bild zu erhalten. Verfahren, die speziell für georefe-

renzierte Bildsammlungen entwickelt wurden, werden in Kapitel 3 kurz vorgestellt. Auch im Projekt Tripod, auf das Kapitel 7 genauer eingeht, versuchte man über die Position des Fotografen und einfache inhaltliche Verfahren, wie zum Beispiel die Erkennung von Gebäuden in den Bildern, Metainformationen über Bilder zu bestimmen.

Ausgehend von den einzelnen Objekten, sind allerdings auch beschreibende Metadaten über die Gesamtkomposition wichtig; so ist es für ein Bildarchiv entscheidend zu wissen, ob auf einem Bild ein Wald in der Natur oder eine Allee in einer Stadt zu erkennen ist, wenn auf dem Bild mehrere Regionen mit Bäumen erkannt wurden.

All diese Probleme bei der automatischen Objekterkennung haben dazu geführt, dass in den letzten Jahren die meisten Archive ihre Bilder ausschließlich durch den Menschen verschlagwortet haben. Die vorhandenen Techniken wurden zunächst nur unterstützend eingesetzt, mit dem Ziel, die Qualität der textlichen Metadaten so hoch wie möglich zu halten.

Die Qualität von textlichen Metadaten bestimmt die Sichtbarkeit eines Bildes in einem Archivierungssystem. Die gängigste Suchmethode ist die Eingabe eines Suchbegriffs, der mit den textlichen Informationen in den Metadaten verglichen wird. Wichtigste Information für diese Art der Suche sind dabei einzelne Schlüsselwörter, die die Szene, wie auch die einzelnen Objekte des Bildes, beschreiben sollen. Für die Vergabe von Schlüsselwörtern findet man zwei unterschiedliche Ansätze:

- Ein Schlüsselwort kann aus einem zuvor festgelegten Schlagwortkatalog oder Thesaurus stammen. Alternativ werden auch komplexere Ansätze wie Ontologien verwendet. In der Praxis hat es sich allerdings erwiesen, dass die Suche nach einem geeigneten Deskriptor aus einem Thesaurus recht zeitaufwendig ist. Vor allem für Nutzer, die mit dem Vokabular nicht vertraut sind, ist eine lange Einarbeitungszeit notwendig. Getty Image bietet beispiels-

weise einen Thesaurus aus dem Bereich Kunst und Architektur an, den Art & Architecture Thesaurus (AAT). Das Vokabular beinhaltet über 34.000 Konzepte, sowie über 131.000 Begriffe, Beschreibungen und historische Belege aus dem Bereich Kunst und Architektur. Der AAT ist ein Beispiel für einen Thesaurus nach den beiden offiziellen Standards für Thesauri, dem ISO Standard ISO 2788 und dem NISO Standard ANSI/NISO Z39.19.

- Die zweite Möglichkeit ist die Verwendung von freien Texteingaben durch den Nutzer (meistens nur einzelne Wörter), die Tags genannt werden. Diese Vergabe eines Schlagworts wird daher Tagging genannt. Im Zuge der Web 2.0 Bewegung findet man die Möglichkeit des Taggings vor allem auf sozialen Plattformen, in denen der Endnutzer die Inhalte der Plattform selbst mitgestaltet. Ein Tag kann als eine zweiteilige Relation aufgefasst werden, $tagging(Objekt, Schlagwort)$, wobei das Schlagwort ein beliebiger Text sein kann. Meistens wird dieser Text allerdings auf ein Wort beschränkt.

In einer erweiterten Form eines Tagging-Systems ist es möglich, dass jeder Nutzer ein beliebiges Objekt mit Tags versehen kann, nicht nur seine eigenen Inhalte. Man spricht dabei auch von kollaborativem Tagging. In dieser Form wird für das Tagging eine dreifache Relation verwendet, die auch den Urheber dieser Zusatzinformation kennt: $tagging(Objekt, Schlagwort, Benutzer)$ (Gruber, 2005).

Definition 2.2 *Ein Tag ist eine zweiteilige Relation $tagging(Objekt, Schlagwort)$, die ein Schlagwort an ein Objekt bindet. Zusätzlich kann beim Tagging auch der Urheber an das Objekt gebunden werden, dann verwendet man die dreiteilige Relation $tagging(Objekt, Schlagwort, Benutzer)$.*

Besonders die Verwendung von Tags stellte sich als sehr effizient und benutzerfreundlich heraus, brachte aber seine eigenen Probleme mit sich (Golder und Huberman, 2006). Jeder Benutzer verwendet, durch die Freiheit bei der Texteingabe verschuldet, gerne sein eigenes Vokabular. So finden sich in den Tags die verschiedensten Sprachen, Kunstwörter, Umgangssprachen oder auch Wörter mit Schreibfehlern. Dadurch bringt diese Art der Organisation nicht nur einen gewissen Aufwand mit sich, sondern bietet meistens nur für den jeweiligen Autor einen direkten Nutzen. Besonderes Interesse liegt daher bei kollaborativen Tagging-Systemen. Hier wird jedes Objekt von jedem Nutzer mitgetagt. Das hat den Vorteil, dass sich bei reger Beteiligung der Nutzer ein gemeinschaftliches Vokabular bildet und dabei zwar die persönlichen Tags noch vorhanden sind, diese aber durch Tags mit einem breiten Zuspruch in der Gemeinschaft verdrängt werden (Golder und Huberman, 2006).

Eine neue Art, Bilder unabhängig von zusätzlichen Texten zu organisieren, bietet die Verortung auf einer Karte. Durch die Information über den Aufnahmeort des Bildes lassen sich auch Fragen, wie „Welche Fotos habe ich in der Nähe von der Stadt Berlin gemacht?“ beantworten. Die geografische Verortung, oder auch Georeferenzierung, kann als spezielle Ausprägung der freien Verschlagwortung (Tagging) aufgefasst werden. Eine Georeferenzierung wird als Tripple definiert, das anstatt eines freien Schlagwortes die gewünschten GPS-Koordinaten enthält, *geoTagging(Objekt, GPS-Position, Benutzer)*. Die GPS-Position ist ein geografisches Datum, das meistens aus einem Tupel von Längen- und Breitengrad besteht. Diese Art der Relation wird in seiner Kurzform GeoTag genannt.

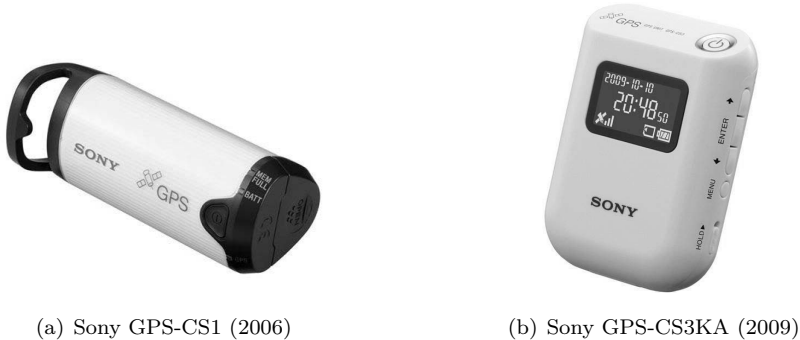
Definition 2.3 *Ein GeoTag ist eine dreiteilige Relation aus Objektkennung, der Position an der diese Information erstellt worden ist und*

der Kennung des Benutzers: *geoTagging(Objekt, GPS-Position, Benutzer)*.

Generell gibt es zwei gängige Möglichkeiten für den Endanwender, die eigenen Bilder zu verorten. Die erste Möglichkeit nutzt das Global Positioning System, kurz GPS. Zusammen mit dem Zeitstempel der Bilder lassen sich diese mit den GPS Informationen verbinden. Zunächst war diese Art der Verschlagwortung mit speziellen GeoTags eine Freizeitbeschäftigung von GPS-Begeisterten, also sogenannten „Early Adoptern“¹⁵ (Schenk, 2007). Im Jahr 2006 brachte Sony das erste GPS-Gerät auf den Markt, das speziell für den Bereich der privaten Fotografie konzipiert wurde (siehe Abbildung 2.1). Der Sony GPS-CS1 sollte Geotagging auch Menschen nahebringen, die nicht das ansonsten nötige Wissen über die GPS-Technologie besitzen. Daher wurde das Handling des Gerätes selbst auf das Nötigste reduziert: Das Gerät lässt sich durch einen einzigen Knopf ein- und wieder ausschalten. Die Qualität des GPS-Signals wird durch ein einziges LED ausgegeben (kein Signal, schwaches Signal, optimales Signal). Heutzutage kennt man solche Geräte unter dem Namen GPS-Tracker oder auch Data-Logger. Der Vorteil eines GPS-Trackers liegt darin, dass sie die Positionen über einen langen Zeitraum speichern können. Da die Position des Gerätes nur in regelmäßigen Abständen aufgezeichnet werden (zum Beispiel alle 15 Sekunden), können auf einem relativ kleinen internen Speicher mehrere Wochen lang die Positionen des Gerätes gespeichert werden.

Ende 2009 haben die ersten Hersteller auch damit begonnen, GPS-Sensoren direkt in ihre Kameras zu integrieren, sodass eine umständliche Nachbereitung der Bilder entfällt. Es ist dennoch interessant zu sehen, dass, trotz immer mehr technischer Unterstützung, der Zuwachs

¹⁵ Der Begriff Early Adopter bezeichnet Menschen, die die neuesten technischen Errungenschaften oder die neuesten Varianten von Produkten oder modischen Accessoires nutzen



(a) Sony GPS-CS1 (2006)

(b) Sony GPS-CS3KA (2009)

Abbildung 2.1: Spezielle GPS-Geräte für den GeoTagging Einsatz

von verorteten Bildern auf der Webseite Flickr sehr konstant bleibt und der erwartete Anstieg bislang ausbleibt. Nach eigenen Stichproben über einen Zeitraum von ungefähr 3 Jahren, ließ sich dieser konstant lineare Zuwachs an verorteten Bildern feststellen (siehe Abbildung 2.3). Ein Erklärungsgrund ist der hohe Preiskampf bei Kleinbildkameras, bei dem schon 50 Euro zusätzlich für eine Funktion wie GPS viel Aufklärungsarbeit über den Nutzen dieser neuen Technik benötigt, bevor sie von Käufern akzeptiert werden.

Durchgesetzt hat sich GPS-Technik und andere räumliche Lagesensoren im Bereich der Mobiltelefone mit eingebauter Kamera, was vor allem auf den großen Erfolg des iPhones zurückzuführen ist. Das iPhone ist seit der Version 3GS mit einem GPS Empfänger ausgestattet. Im Durchschnitt werden von Nutzern dieses Mobiltelefons täglich mehrere tausend Bilder bei Flickr eingestellt, die zumindest die Möglichkeit haben, verortet zu sein. Momentan (Februar 2012) ist das GPS-Fotohandy damit die beliebteste „Kamera“ in der gesamten Flickr-Gemeinschaft. Abbildung 2.2 (Quelle: www.flickr.com/cameras/) zeigt die Anzahl der Nutzer, die im Laufe des letzten Jahres mit dem jeweiligen Kameratyp ein Bild in Flickr eingestellt haben. Alle bisherigen

iPhone-Versionen belegten jeweils Platz eins. Die Verwendung des mobilen Internets macht diese Geräte zusätzlich interessant für die Verwendung von Diensten wie einem geografischen Empfehlungssystem, da die erstellten Bilder schon bei der Entstehung dem Service zur Verfügung gestellt werden können.

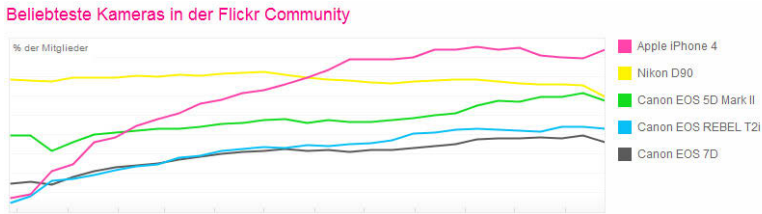


Abbildung 2.2: Die fünf beliebtesten Kameras auf Flickr (November 2011)

Die zweite Möglichkeit ein Bild zu verorten, ist die Position der Bilder über eine Karte selbst zu bestimmen. Während ein GPS-Signal die Position der Bilder in einer Genauigkeit von bis zu einer Mikrosekunde Grad bestimmt, liefert die Kartenmethode teilweise sehr ungenaue Informationen. Die meisten Karteninterfaces bieten die Möglichkeit beliebig zu zoomen und dadurch kommt es stark auf den Nutzer an, wie genau er seine Bilder positioniert. Zudem gibt es zwei grundsätzlich verschiedene Konzepte, wie man die Position eines Bildes angeben kann. Dadurch, dass ein GPS-Tracker mit der Kamera mitgeführt wird, wird automatisch die Position der Kamera festgehalten. Im Gegensatz dazu legt ein Nutzer, der sein Bild frei auf einer Karte positionieren kann, oft die Position des fotografierten Objektes als Position des Bildes fest. Daher kann es zu einem deutlichen Unterschied kommen. Zum Beispiel sind die meisten guten Standpunkte, um den Eiffelturm in Paris zu fotografieren, einige hundert Meter vom eigentlichen Objekt entfernt. Um eine Datenkonsistenz zu erreichen, legen die meisten Bibliotheken auf eine

der beiden Positionen fest, was in den meisten Fällen der Standpunkt des Fotografen sein wird, um die automatische Positionierungsmethode nicht auszuschließen.

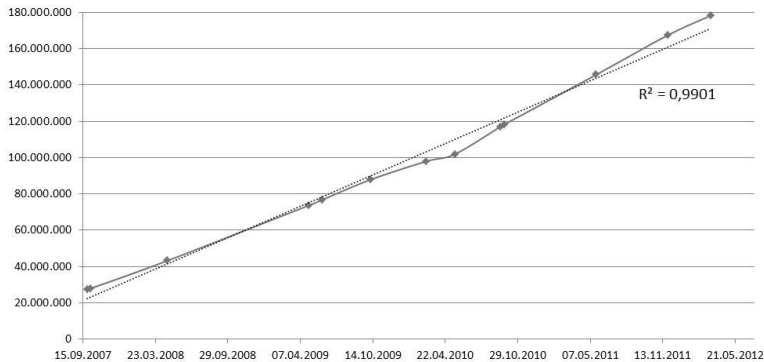


Abbildung 2.3: Anzahl von georeferenzierten Bilder auf Flickr.

Wissenschaftlich interessant sind vor allem Bibliotheken, in denen Privatleute ihre georeferenzierten Bilder online stellen, da hier angenommen werden kann, dass diese Bilder aus Eigeninteresse aufgenommen wurden. Im Gegensatz dazu findet man bei anderen Bibliotheken viele Auftragsarbeiten. Aber auch in diesen Bibliotheken lassen sich unterschiedliche Schwerpunkte bei der gewünschten Bildauswahl feststellen. Es gibt Webseiten, wie Flickr, die grundsätzlich keinen Themenschwerpunkt für die verwalteten Bilder vorschreiben. Der Nutzer kann Bilder zu jedem beliebigen Thema einstellen. Auch die Quote an georeferenzierten Bildern ist unterschiedlich. Flickr bietet zum Beispiel seit 2006 die Möglichkeit zur Verortung der Bilder an. Im Durchschnitt werden um die 3.000 Bilder pro Minute eingestellt, allerdings nur ein Bruchteil davon wird mit einer Ortsangabe versehen. Insgesamt wurden um die 180 Millionen verortete Bilder weltweit eingestellt (Flickr verfügt über ungefähr 5 Billionen Bilddateien insgesamt). Viel interessanter sind daher reine GeoTagging Communities wie Panoramio, die sich

komplett auf die GPS Fotografie spezialisiert haben und daher gezielt diese Nutzergruppen ansprechen. Schon auf Startseite befindet sich der Slogan „show your favorite places“. Seiten wie Panoramio oder auch Locr schränken sich bei Auswahl der Bildmotive auf Bilder von Gebäuden und Landschaften ein und wollen Menschen als Bildmotiv eher ausschließen.

2.2 Neue Möglichkeiten für Empfehlungssysteme

Wie vorher bereits erwähnt, ermöglicht die Verortung von Bildern, diese auf einer Karte darzustellen oder nach geografischen Bedingungen zu filtern. Die simple Anzeige aller Bilder aus einer privaten Sammlung auf einer Karte scheint noch recht übersichtlich und hilfreich zu sein. Nimmt man allerdings alle Bilder von verschiedenen Benutzern für eine bekannte Stadt wie Amsterdam zusammen, findet man zum Beispiel auf Panoramio über 5.000 Bilder, die über den Stadtkern verteilt sind. Dies macht es notwendig, eine Auswahl an Bilder zu treffen, die eine sinnvolle Ansicht ermöglicht. Alleine für die Darstellung der Bilder kann es sinnvoll sein, die Bilder so auszuwählen, dass sich die stellvertretenden Abbildungen auf der Karte nicht zu sehr überlappen. Dies kann mit recht simplem Verfahren, wie einer gleichmäßigen Raumunterteilung oder räumlichen Clusterverfahren, erreicht werden. Abbildung 2.4 zeigt die Darstellung von Bildern im Stadtbereich von Barcelona auf GoogleMaps ähnlich einem Mosaik. Bei dieser Art von Auswahl spricht man von einer Darstellung mit einer hohen räumlichen Diversität (Tang und Sanderson, 2010). Der Begriff der räumlichen Diversität wird in 5 näher diskutiert. Abbildung 2.4 verdeutlicht allerdings auch recht anschaulich, dass eine rein räumliche Verteilung der Ergebnisse nicht unbedingt eine gute Auswahl ausmacht, wenn man auf der Suche nach typischen Ansichten einer Stadt ist. Keine der gezeigten

Bilder zeigt berühmte Baudenkmäler oder besonders charakteristische Gegenden der Stadt. Die Bilder wirken austauschbar und gleichwertig. In den meisten Fällen bleiben typische Ansichten in der Ergebnismenge verborgen, daher muss für diese Fragestellung ein anderer Ansatz verwendet werden.

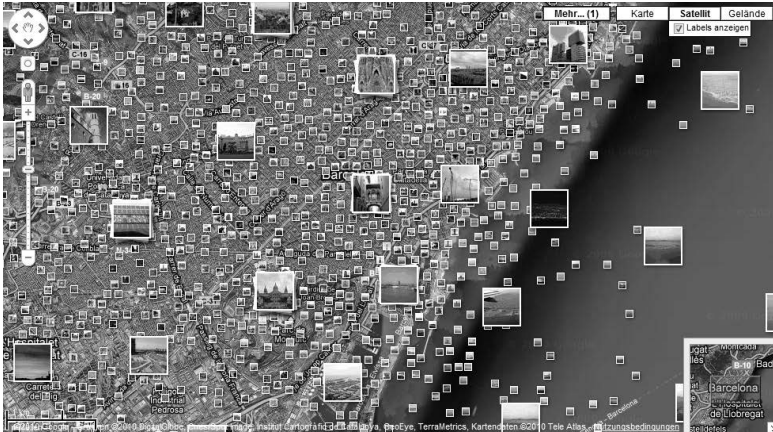


Abbildung 2.4: Geographische Diversität

Eine ganz eigene Art von georeferenzierten Informationen werden über sogenanntes Social Media erzeugt. Social Media enthält Informationen, die in sozialen Plattformen entstehen und über die diese Informationen an bestimmte Nutzer weitergegeben werden. Die neuste Entwicklung dieser Plattformen geht dazu über, jede Information an die momentane Position des Nutzers zu binden; dadurch generiert der Nutzer nicht nur eine georeferenzierte Information, sondern hinterlässt auch einen ständigen Datenstrom im geografischen Raum, an dem man bestimmte Aktivitäten des Nutzer ablesen kann. Die Webseite bdnooz.com führt beispielsweise eine Liste mit über 60 verschiedenen ortsbezogenen sozialen Netzwerken. Wenn ein Bild entsteht, weiß das System automatisch, an welchem Ort es entstanden ist. Gerade weil diese Information

auch an den Benutzer gebunden ist, macht es diese Netzwerke zu einer perfekten Plattform für die vorgestellten geografischen Empfehlungssysteme.

Mit jedem Bild, das ein Benutzer erzeugt, und durch die Möglichkeiten des sozialen Internets, sich als Urheber dieses Bildes zu identifizieren, wird neben dem Aufenthaltsort des Bildes auch der Standpunkt des Nutzers festgehalten. Dies ist eine neue Art des Nutzerfeedbacks, welches bisher noch nicht für Empfehlungssysteme verwendet wurde. Kollaborative Filter machen sich das persönliche Feedback zunutze, um Gemeinsamkeiten zwischen den einzelnen Benutzern zu finden.

Die Chancen der verstärkten Personalisierung, besonders auf e-Commerce Plattformen, beschreibt Anderson (2006) sehr treffend in seinem Buch „The Long Tail - Der lange Schwanz: Nischenprodukte statt Massenmarkt - Das Geschäft der Zukunft“. Anderson beschreibt hier den Wandel der Handelsgeschäfte, der sich immer mehr durch verteilte Handelsnetze anstatt zentralen Großhandel auszeichnet. Waren werden somit nicht mehr an einem zentralen Ort ausgestellt, verkauft und gelagert, wie es noch bei den großen Kaufhäusern der Fall war, sondern auf kleinere Lager verteilt und auch von dort versendet. Verkauft wird dabei auf virtuellen Plattformen, wie Amazon. Dies macht es möglich, auch Waren vorrätig zu halten, die ganz selten gekauft werden, sogenannte Nischenprodukte. Abbildung 2.5 zeigt eine typische Verteilung der verkauften Stückzahlen auf einer e-Commerce Plattform. Ungefähr 20% der verkauften Waren befinden sich am Kopf der Verteilung. In der Regel finden sich hier die momentanen Bestseller. So ziemlich jedes Produkt aus dem Kopf der Verteilung kann bedingungslos an jeden Benutzer empfohlen werden, da sie die gegenwärtigen Trends in der Gesellschaft widerspiegeln.

Interessant ist nun allerdings der lange Schwanz der Verkäufe, in denen sich hauptsächlich Nischenprodukte finden. Auch wenn diese - immer-

hin 80% der verkauften Produkte - für bestimmte Personen interessant sind, ist es eine Herausforderung, die passenden Produkte hieraus für einen bestimmten Benutzer zu empfehlen.

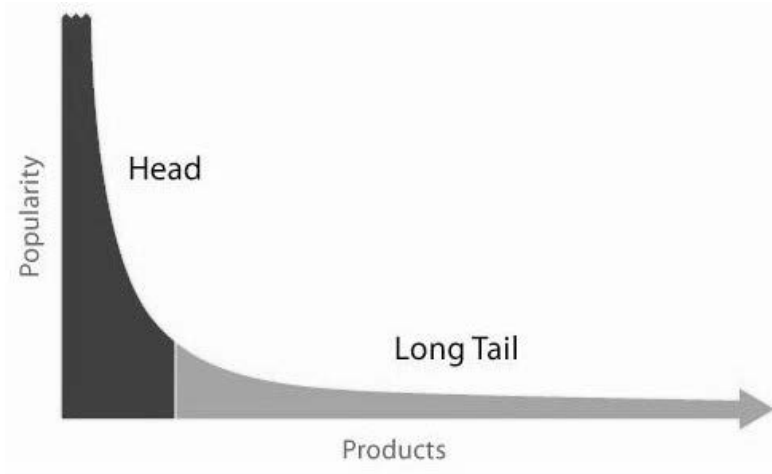


Abbildung 2.5: Verteilung von Produkten nach Beliebtheit (Grafik in Anlehnung an Anderson (2006))

Meine Arbeit beschäftigt sich in Kapitel 5 mit der Frage, ob man so eine Long-Tail-Verteilung auch in Datensätzen von verorteten Bildern finden kann. Folglich müsste es Orte geben, die besonders häufig aufgesucht werden und daher für eine breite Masse interessant sind. Im Gegensatz dazu wird es bestimmte Orte geben, die nicht für jeden interessant sind und daher einen Anwendungsfall für Empfehlungssysteme darstellt. Dabei werden räumliche Entscheidungen, die in diesen Bildern kodiert sind, stellvertretend für die Produkte des e-Commerce Handels betrachtet. Zusätzlich muss ein geeignetes Maß für die Popularität dieser Entscheidungen gefunden werden. Letztendlich muss eine Methode entwickelt werden, die die relevanten Bilder für eine bestimmte Person im langen Schwanz der Verteilung identifizieren kann.

2.3 Ausgangsfragen

Die Entwicklungen in den letzten Jahren im Bereich der Fotografie, explosionsartiger Anstieg durch den Umstieg auf digitale Speicherung und die neuen Möglichkeiten durch mehr Metadaten, sind der Ausgangspunkt für diese Arbeit. Für die Umsetzung eines geografischen Empfehlungssystems steht dabei eine grundsätzliche Frage im Mittelpunkt der Überlegungen:

- Wie kann man die Bewertung von geografischen Orten, die sich aus räumlichen Verhalten der Nutzer ergibt, für geografische Empfehlungssysteme nutzen?

Weiterhin wurde nach einem konkreten Anwendungsfall gesucht, für den ein geografisches Empfehlungssystem sinnvoll ist. Für diesen Anwendungsfall sind Antworten auf die folgenden Fragen ausschlaggebend:

- Was sind typische Bildgruppen mit geografischem Bezug?
- Was sind Anwendungsfälle, bei denen Nutzen aus einem geografischen Empfehlungssystem gezogen werden kann?

Für die Umsetzung des Verfahrens müssen bestimmte Teilaspekte eines nutzerbasierten kollaborativen Filters neu definiert und zusätzlich Metriken erarbeitet werden, die auch die vorhandenen geografischen Metadaten mit in Betracht ziehen. Um die georeferenzierten Daten als Eingangsdaten für ein Empfehlungssystem nutzen zu können, muss vor allem der Begriff eines Ortes als kleinste Größe in diesem System definiert werden. Daraus entsteht im Laufe der Arbeit der Begriff der Heatmap, welche ein Model des geografischen Bildraumes beschreibt. Für jeden dieser festgelegten Orte kann man entscheiden, ob ein bestimmter Nutzer an diesem Ort aktiv teilnimmt, indem eines seiner Bilder dort verortet wurde, oder nicht. Wir werden im späteren den Begriff

einer räumlichen Entscheidung für diese Orte einführen. Insgesamt sind folgende Überlegungen nötig:

- Wie können benutzerspezifische räumliche Entscheidungen im geografischen Raum begrifflich geklärt werden?
- Was sind Anforderungen an Popularitätsmaße für die Bewertung von räumlichen Entscheidungen? Welche Metriken genügen diesen Ansprüchen?
- Was sind mögliche Berechnungen von Nutzerähnlichkeiten, die auf den Metriken für räumliche Entscheidungen aufbauen, um einen benutzer-basierten kollaborativen Empfehlungsansatz zu ermöglichen?

Die Umsetzung der eigenen Ansätze sowie deren Evaluation warf abschließend folgende Fragen auf:

- Wie kann dieser theoretische Ansatz in Software umgesetzt und dabei dem ursprünglichen Anwendungsfall gerecht werden?
- Welche Möglichkeiten bestehen, die Ergebnisse gegen die eigenen Anforderungen und andere Systeme evaluieren zu können?

Die Beantwortung dieser Fragen wird der Bestandteil der weiterführenden Arbeit sein.

3 Geografische Bildsammlungen

In den letzten Jahren sind eine Vielzahl von Web 2.0 Anwendungen entstanden, die sich gezielt auf georeferenzierte Medien und Objekte konzentriert haben. Dabei sind mittlerweile nicht nur Bilder mit Geo-Tags von großem Interesse auch Informationen wie Text (Geonotes (Espinoza u. a., 2001)) oder momentane persönliche Standort (Google Latitude (Page und Kobsa, 2009)). Diese Information ermöglichen neue interessante Anwendungsfälle, wie soziale ortsbasierte Netzwerke oder ortsbasierte Filterung. Der Schwerpunkt dieser Arbeit liegt zwar auf georeferenzierten Bildsammlungen, Teile davon lassen sich aber auch für andere georeferenzierte Daten übertragen. Voraussetzung ist, dass der georeferenzierte Ort einer persönlichen Entscheidung gewählt wurde. Alle diese georeferenzierten Daten lassen sich zu dem Oberbegriff „Volunteered Geographic Information“ (VGI) zusammenfassen, da sie von den Nutzern freiwillig erhoben werden. Dabei ist zu beachten, dass die Ersteller der geographischen Informationen keineswegs Experten sind und somit mit unterschiedlichstem intellektuellem Hintergrund Informationen bereitstellen. Coleman u. a. (2009) prägten für die Nutzer dieser Anwendungen den Begriff „Produser“, ein Mischwort aus Producer und User, und leitet aus der vorherrschenden Diskussion über das Thema fünf verschiedene heterogene Nutzergruppen für Produser geografischer Informationen ab: Neophyte, Interested Amateur, Expert Amateur, Expert Professional und Expert Authority. Zusätzlich bestimmten sie vier verschiedene Kontexte in den sich eine mögliche VGI Anwendung abspielen kann: Mapping and Navigation, Social Networks, Civic/Governmental, Emergency Reporting. Daraus ergeben sich

20 verschiedene Rollen, in denen sich ein Nutzer befinden kann. Die Zielgruppe der hier vorgestellten Bildsammlungen sind dabei vor allem interessierte Amateure. Die Anwendungen zur Bildverwaltung selbst bieten meistens Möglichkeiten an, wie man auch aus sozialen Netzwerken kennt.

Das Forschungsfeld der VGI ist eine recht junge Richtung in der GIScience Gemeinschaft. In Goodchilds Aufsatz „Citizens as sensors“ Goodchild (2007) wird der Begriff VGI zum ersten Mal als Spezialfall eines allgemeineren Phänomens im World Wide Web dem „user-generated content“ (Inhalte, die durch Nutzer erzeugt werden) eingeführt. Kuhn (2007), Parsons (2008) und Matyas (2011) geben einen guten Einstieg für eine weitere Vertiefung in das Thema VGI.

3.1 Übersicht über aktuelle Bildsammlungen

Im Folgenden werden drei beispielhafte VGI Anwendungen vorgestellt, die sich mit georeferenzierten Bildern beschäftigen. In allen drei Fällen wird es dem Nutzer ermöglicht, Bilder zu einer bestehenden Bildsammlung hinzuzufügen. Dabei ist von auch von Interesse, welche Anforderungen die jeweiligen Systeme an die Bilder der Anwender stellen und wie diese Anforderungen von der Benutzerseite adaptiert werden. Die Anwendungsfälle reichen dabei über Spiele und geografische Projekte bis hin zu semiprofessioneller Bildverwaltung. Technisch gesehen handelt es sich bei diesen VGI Anwendungen um Mashups (Michael u. a., 2010), die mit Hilfe von webbasierten GIS Anwendungen, wie GoogleMaps oder YahooMaps, umgesetzt werden. Auf den eingebundenen Karten werden die entsprechenden Daten visualisiert und eingetragen.

3.1.1 www.geograph.org

Geograph ist eine Mischung aus Spiel, einem Fotografie- und einem Forschungsprojekt über unterschiedliche geografische Fragestellungen. Ziel dabei ist es, jede der ein Quadratkilometer großen Kacheln, die den Vermessungen des britischen Vermessungsamtes Ordnance Survey (OS) zugrunde liegen, mit Bildern zu beschreiben. Dabei sind Bilder erwünscht, die die geographischen Eigenschaften der Landschaften beschreiben, die in dieser Kachel zu finden sind. Ansichten auf Orte, die entweder durch den Menschen entstanden (Gebäude, Denkmäler, ...) oder auf natürliche Art entstanden sind (Berge, Flüsse, ...), sind bevorzugte Motive. In der Anleitung des Spiels wird diese Einschränkung folgendermaßen beschrieben:

Photos need not be master works of art, the important thing is the human and physical geographical information they contain.

Diese Einschränkung macht nicht nur für Geograph einen Sinn, sondern insbesondere auch für alle Verfahren, die mit Hilfe eines Bildes, Aussagen über den Aufnahmeort treffen wollen. Die Verfahren, die im späteren Verlauf vorgestellt werden (siehe Kapitel 5 und Kapitel 6), dienen dabei als Beispiel. In diesem Sinne sind Bilder von beweglichen Objekten, wie etwa Menschen, oder kurzweiligen Ereignissen, wie etwa Straßenfeste, nicht interessant für die Bildersammlung Geograph. An der Qualitätskontrolle können sich alle Teilnehmer des Systems beteiligen, indem sie Verbesserungsvorschläge für ein bestimmtes Bild anbringen. Diese Vorschläge werden dann wiederum von einem Moderator gegengelesen, bevor sie veröffentlicht werden.

Der spielerische Ansatz dieser Sammlung wird daran deutlich, dass man in Geograph für jedes Bild eine bestimmte Anzahl von Punkten erhält. Die Höhe der Punkte für ein hochgeladenes Bild steht allerdings in Relation zu der Anzahl der Bilder, die schon in einer Kachel zu finden sind.

Das Gebiet an dem man teilnimmt, wird somit immer weniger „wert“, je mehr Bilder dort bereits zu finden sind. Daher liegt der Fokus des Projekts nicht darauf, möglichst viele Aufnahmen für wenige Gebiete zu erreichen, sondern darauf, eine möglichst gute Abdeckung über alle Kacheln von Großbritannien zu erreichen. Im Moment hat Geograph eine Abdeckung von 79,7% (Stand: März 2012) mit 2.815.490 Bildern in 264.563 fotografierten Kacheln. Dennoch lässt sich gut erkennen, dass Nutzer vor allem in den Ballungsgebieten der englischen Großstädte deutlich mehr Bilder einstellen, als auf dem Land, selbst wenn sie dafür keine Punkte mehr erhalten (nur die ersten vier Bilder bekommen Punkte). Die Kachel in London zum Beispiel, in der der Trafalgar Square liegt, beinhaltet über 1.500 Bilder. Die Nutzer scheinen sich daher auch bei den sehr strengen Vorgaben des Spiels nicht davon beirren zu lassen, trotzdem noch Bilder einzustellen; im Durchschnitt beinhaltet jede Kachel 7,5 Bilder. Den meisten Nutzern scheint es also nicht wirklich wichtig zu sein, eine hohe Punktzahl zu erreichen, sondern für das Projekt als solches einen Beitrag zu liefern bzw. ihre eigenen Bilder online zu präsentieren.

Wie in Abbildung 3.1 auf der kleinen Übersicht zu sehen ist, bietet Geograph dem Nutzer eine farbcodierte Darstellung, die eine feste Kategorisierung nach der Anzahl der Bilder auf einer ansteigenden Skala zeigt. Man kann schnell erkennen, dass manche Gegenden recht häufig von Nutzern fotografiert wurden, während andere kaum besucht wurden. Zum Beispiel scheint der Süd-Westen von Leeds viel beliebter zu sein als der Nord-Osten. Diese Beobachtungen werden wir auch in anderen Datensätzen finden, insbesondere wenn die Abdeckung durch fotografische Aufnahmen nicht im Vordergrund steht. Im Fall von Leeds ist der Grund wahrscheinlich die höhere Bevölkerungsdichte im Süd-Westen, was den Aufwand für die Bewohner dieser Gebiete, ein Bild zu machen, deutlich reduziert. Aber auch andere Faktoren spielen eine maßgebende

Rolle. Im Osten von Manchester zum Beispiel ist die Bevölkerungsdichte fast null. Allerdings befindet sich dort der „Peak District“, ein sehr bekanntes Naturgebiet und beliebtes Ausflugsziel. So scheint sich die Popularität eines Ortes ebenfalls in der Anzahl der Bilder an diesem Ort wiederzuspiegeln.

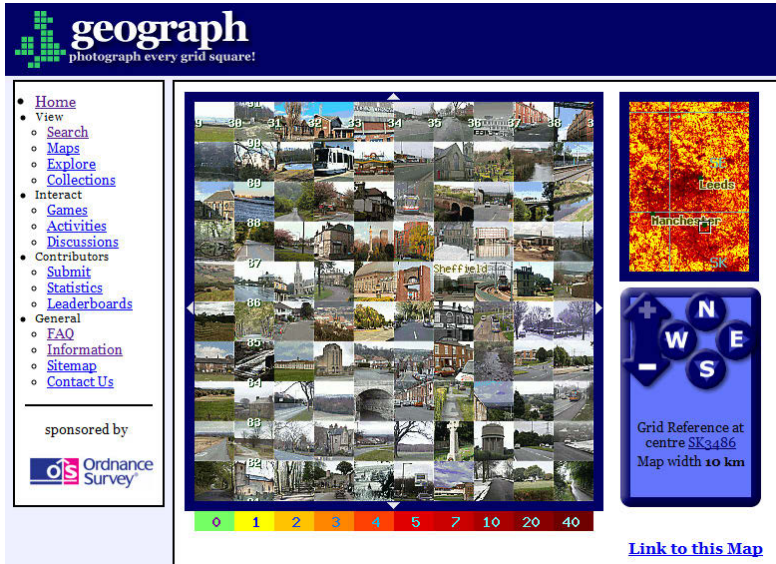


Abbildung 3.1: Kachelansicht für 100 Bilder aus Sheffield auf Geograph. Übersichtskarte rechts oben zeigt die Verteilung der Bilder in der Umgebung.

Während sich Geograph um eine Bildsammlung für ganz Großbritannien bemüht, hat sich ein ähnliches Projekt zur Aufgabe gemacht, jeden Schnittpunkt der ganzzahligen Höhen- und Breitengrade zu bebildern und das weltweit. Auf der offiziellen Webseite des Degree Confluence Projects findet man die folgende Einleitung:

The goal of the project is to visit each of the latitude and longitude integer degree intersections in the world, and to take pictures at each

location. The pictures, and stories about the visits, will then be posted here.

Abbildung 3.2 zeigt den Ausschnitt der bebilderten Weltkarte für Europa. Auffällig ist auch hier, dass an unterschiedlichen Schnittpunkten eine ganz unterschiedliche Anzahl von Bildern zu finden ist, abermals an größeren Städten und touristischen Highlights, an denen eine hohe Popularität auch erwartet werden kann.

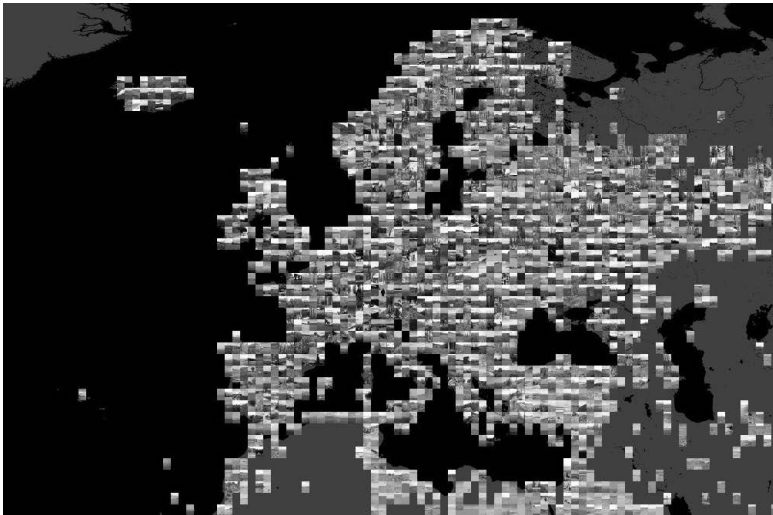


Abbildung 3.2: Screenshot des Degree Confluence Project für Europa
(<http://www.confluence.org/>)

Interessant wird vor allem der Datensatz von Geograph dadurch, dass beinahe jedes Bild mit einer umfangreichen textlichen Beschreibung online gestellt wird. Eine kurze Beschreibung der Dinge, die auf diesen Bildern zu sehen sind, macht eines der Qualitätsmerkmale für ein gutes Geograph-Bild aus.

Für die vorliegende Arbeit war vor allem die Darstellung der Häufigkeitsverteilung ein wichtiges Indiz dafür, dass durch die Georeferenzie-

rung auch immer eine starke Häufung von Bildaufnahmen an populären oder einfach zu erreichenden Orten zu finden ist. Für eine genauere Untersuchung des räumlichen Verhaltens der Nutzer sind die Ortsdaten mit dem festgelegten Meilenraster allerdings zu grob, um sie im Rahmen eines geographischen Empfehlungssystems, wie es in dieser Arbeit beschrieben wird, direkt nutzen zu können.

3.1.2 www.panoramio.com

Panoramio kann ähnlich wie Geograph als geographisches Projekt gesehen werden, da auch hier der Fokus ganz stark auf Plätzen (Places) liegt. Im Fall von Panoramio wird allerdings nicht versucht, die Ansichten auf so viele Orte, wie möglich zu streuen, sondern sich auf die favorisierten Plätzen der Nutzer zu beschränken. Das Motto der Webseite lautet dementsprechend „Show your favorite places“ oder zu Deutsch „Zeigen Sie uns Ihre Lieblingsplätze“. Panoramio startete im Sommer 2005 als unabhängige Webapplikation. Mitte 2007 war die Webanwendung so erfolgreich, dass Google auf die kleine Firma Panoramio aus Spanien aufmerksam wurde und diese komplett aufgekauft hat. Seither erforscht Google auf Grundlage dieser Bildbasis neue Einsatzmöglichkeiten von verorteten Bildern. Erste Ergebnisse dieser Arbeit kann man heute schon auf den Kartenprodukten von Google (GoogleMaps und GoogleEarth) sehen. Abbildung 3.3 zeigt zum Beispiel eine Ansicht von Parc Güell auf der Seite von Panoramio mit Hilfe von GoogleMaps. Panoramio verwendet im Gegensatz zum Gridsquare von Geograph eine genaue Georeferenzierung mit Hilfe von GPS-Koordinaten. Entweder werden diese aus den Metadaten übernommen oder der Nutzer georeferenziert die Bilder auf einer Karte. Die Datenqualität wird ebenfalls über die gegenseitige Kontrolle der Nutzer garantiert, die neue Positionen für ein Bild vorschlagen können. Durch den Fokus auf persönliche Lieblingsplätze steht diese Bilderdatenbank im Fokus meiner Arbeit,

da hier persönliche Entscheidungen getroffen werden, die sich in den eingestellten Bildern ablesen lassen.



Abbildung 3.3: Bilder von Park Güell in Barcelona auf Panoramio.

3.1.3 www.flickr.com

Flickr ist sicherlich einer der bekanntesten und beliebtesten Web 2.0 Anwendungen, die sich von Anfang an auf das Management und das Hosting von Bildern spezialisiert hat. Der Grundgedanke zu Beginn war, Privatpersonen einen Ort zu bieten, um ihre Bilder online mit anderen Teilnehmern der Webseite oder Freunden zu teilen. 2008 wurde zusätzlich die Zusammenarbeit von Flickr mit der professionellen Bildagentur Getty beschlossen. Die ermöglichte es Getty, bestimmte Bilder von privaten Nutzern in ihre Sammlung einzukaufen und im Gegenzug auch Bilder aus ihrer eigenen Sammlung über Flickr zu verkaufen.

Im Jahr 2006 wurden die ersten externen Applikationen für Geotagging in Flickr eingebaut. Kurze Zeit später war dieses Feature ein grundlegender Bestandteil in Flickr und stand als eingebaute Funktionalität zur Verfügung. Abbildung 3.4 zeigt zum Beispiel Bilder, die im Stadtgebiet von Paris aufgenommen wurden. Technisch gesehen unterscheidet

sich diese Möglichkeit allerdings nicht von Seiten, wie “lo-alize.us”, die ebenfalls die bereits vorhandenen Metadaten in Flickr-Bilder auslesen oder dem Nutzer seine Bilder auf einer Karte verorten lassen. Neuere Ansätze für eine geographische Verortung wurden während des Projekt Tripod erforscht, auf das Kapitel 7 einen Überblick bietet.



Abbildung 3.4: Bilder aus dem Stadtgebiet von Paris auf Flickr.

3.2 Fokus der einzelnen Bildsammlungen

Abbildung 3.5 zeigt 100 Bilder von Geograph aus einem zentralen Grid in Sheffield. Es ist auffällig, dass keines der Bilder einen Menschen zeigt. Das liegt an dem sehr strengen Fokus der Webseite auf Aufnahmen von nicht-beweglichen Gegenständen. In nahezu allen Grids finden sich daher kaum Bilder, in denen Personen zu erkennen sind, und so gut wie gar keine Bilder, auf denen Personen das Hauptmotiv des Bildes darstellen. Aufgrund dieses strengen Fokus kann bei fast allen Bildern davon ausgegangen werden, dass der Anblick, der abgebildet wurde, auch tatsächlich nur von dem entsprechenden Standpunkt aus zu sehen ist. Für die folgenden Überlegungen in dieser Arbeit ist dieses Szenario der Idealfall, da der Aufnahmeort auch gleichzeitig den möglichen In-

halt des Bildes bestimmt (ihn zumindest sehr vorhersehbar macht). Ein Bild von einer bestimmten Person zum Beispiel ist allerdings ganz unabhängig vom Ort reproduzierbar und für den Fotografen selbst spielt der Ort, an dem er stand keine besonders große Bedeutung. Anders ist es allerdings, wenn das Bild dieser Person ein bedeutendes Gebäude im Hintergrund aufzeigt.



Abbildung 3.5: Bilder aus dem Stadtgebiet von Sheffield auf Geograph.

Bei einem geographisch motivierten Projekt scheint dieser Fokus nicht sonderlich überraschend zu sein, daher wurden zwei weitere Bildgalerien untersucht, ob man dort Ähnliches beobachten kann. Es wurden 100 zufällige Bilder von Panoramio und Flickr analysiert, um den unterschiedlichen Fokus dieser Bildersammlungen aufzuzeigen. Beide Onlinogalerien bieten die Möglichkeit, Bilder zu georeferenzieren. Allerdings legt Panoramio seinen Fokus, wie auch schon Geograph, auf Bilder mit

Gebäuden und Landschaften. Panoramio bietet dem Nutzer sogar an, Bilder zu markieren, die nicht konform mit dieser Einschränkung sind.

Unterschieden wurden bei der Untersuchung drei Gruppen: zum einen Bilder von Gebäuden oder unbeweglichen Gegenständen; zudem wurde differenziert, ob zusätzlich noch Menschen zu sehen waren. War der Hauptgegenstand des Bildes eine Person (zum Beispiel ein reines Portrait), wurde es in die letzte Gruppe einsortiert. Untersucht wurden Bilder aus vier verschiedenen Städten, die im Verlauf dieser Arbeit noch öfter als Analysebeispiele dienen werden.

Stadt	Bilder ohne Menschen	Bilder mit Menschen	Bilder von Menschen
Amsterdam	88	8	4
Bamberg	99	1	-
Cardiff	82	18	-
Dublin	88	12	-
Durchschnitt	0,89	0,1	0,01

Tabelle 3.1: Motive von Bildern aus der Bildergalerie Panoramio.

Stadt	Bilder ohne Menschen	Bilder mit Menschen	Bilder von Menschen
Amsterdam	56	8	36
Bamberg	62	2	36
Cardiff	53	7	40
Dublin	62	17	21
Durchschnitt	0,58	0,08	0,34

Tabelle 3.2: Motive von Bildern aus der Bildergalerie Flickr

3.3 Arbeiten zu räumlichen Benutzerdaten

Es gibt verschiedene Ansätze, mit diesen Daten umzugehen, wobei sich viele Arbeiten mit der Fragestellung beschäftigen, wie sich aus diesen Informationen qualitative Aussagen für den geographischen Raum extrahieren lassen. Andere Arbeiten befassen sich mit der Fragestellung, welche Aussagen sich über die geographischen Informationen auf das Verhalten der Nutzer ableiten lassen. In einem geographischen Empfehlungssystem spielen beide Fragestellungen eine Rolle.

Viele Forschungsarbeiten, die sich mit dem räumlichen Verhalten von Menschen beschäftigen, verwenden komplette GPS-Bewegungsspuren und selten diskrete räumliche Entscheidungspunkten, wie Positionen an denen Bilder aufgenommen worden sind. Spezielle räumliche (oder raumzeitliche) Clusterverfahren werden dabei genutzt, um wichtige Entscheidungspunkte für den einzelnen Nutzer in diesen GPS-Spuren zu finden Ashbrook und Starner (2002).

Ähnliche Verfahren können dann wiederum auf die Entscheidungspunkte der gesamten Nutzergemeinschaft angewendet werden, um besondere Orte im Raum ausfindig zu machen. Mit georeferenzierten Bildern zu arbeiten hat den Vorteil, dass der Nutzer bereits willentlich einen persönlichen Entscheidungspunkt mit seinem Bild markiert. Daher können die Daten wesentlich differenzierter betrachtet werden als reine GPS-Tracks, bei denen eine Häufung von GPS-Signalen an einer Stelle aufgrund von eines erhöhten Interesse der Nutzer zustande kam oder er/sie doch nur an einer Ampel warten musste. Ashbrook und Starner (2003) verwendeten zum Beispiel GPS-Tracks von Testkandidaten, um persönlich relevante Punkte im geographischen Raum zu bestimmen. Dabei spielte die Verweildauer an einem Ort die entscheidende Rolle. GPS-Positionen sind aber immer fehlerbehaftet sind, auch wenn ein GPS-Gerät am selben Ort verweilt. Deshalb war für die Autoren die

Verwendung eines geographischen Clusterverfahren unvermeidbar. Auch in georeferenzierten Bildsammlungen trifft man auf dasselbe Problem, da auch hier die geographischen Koordinaten meistens von GPS-Geräten stammen. Ashbrook und Starner bedienen sich einer Variante des k-means Algorithmus, der in Kapitel 5 als Grundlage für einen eigenen geographischen Clusteransatz verwendet wird. Der Vorteil ihrer Methode ist, dass die Anzahl der Cluster nicht im Vorhinein festgelegt werden muss. Weitere räumliche Clusteralgorithmen werden in Zhou u. a. (2007) oder Matyas (2011) beschrieben.

3.4 Arbeiten mit geographischen Bildsammlungen

Einige Arbeiten zu “Volunteered Geographic Information” beschäftigen sich mit recht ähnlichen Fragestellungen, wie geographische Empfehlungsansätze. Ausgehend von den unstrukturierten, ungeordneten Datensammlungen wird versucht, neues Wissen oder neue Erkenntnisse zu erhalten. Dabei gibt es verschiedene Fragestellungen, die man versucht, mit Hilfe dieser Daten zu beantworten. Eine konkrete Fragestellung ist die Identifikation von interessanten Orte.

Ahern u. a. (2007) stützen sich bei ihrem World Explorer auch auf textliche Metadaten, die in Bildsammlungen enthalten sind. In ihrem Anwendungsbeispiel verwenden sie Bilder aus Flickr und haben somit Zugriff auf die Tags dieser Bilder. Zunächst verwenden sie ein k-means Clusterverfahren, das die euklidische Distanz als Abstandsmaß verwendet. Die Auswahl des passenden Zerlegungsfaktor k ist abhängig von der Bilddichte in dem entsprechenden Gebiet. Bei einer hohen Dichte werden mehr Cluster erzeugt als in Gebieten mit recht wenig Bilder (< 50 Bilder). Für jedes einzelne Gebiet soll nun ein Tag ausgewählt werden, der am repräsentativsten für das gesamte Gebiet ist. Dabei wird ein Verfahren verwendet, das insbesondere im Information Retrie-

val Bereich sehr bekannt ist. Man gewichtet ein Tag mit Hilfe seiner Häufigkeit in Relation zu seiner Häufigkeit im gesamten Suchgebiet. Man nennt dieses Maß TF-IDF für Term Frequency – Inverse Document Frequency. Weitere Methoden, um ortsrelevante Tags zu identifizieren, werden in Rattenbury und Naaman (2009) behandelt. Im späteren Verlauf meiner Arbeit werden wir zum Beispiel sehen, dass beinahe alle Bilder den Namen der Stadt aufweisen, in der sie aufgenommen wurden. Dieser Tag sollte daher nicht repräsentativ für Bilder aus einem kleinen Teilgebiet der Stadt sein.

In einer Folgearbeit zu dieser geht Kennedy u. a. (2007) einen Schritt weiter und versucht, aus allen Bildern zu diesen repräsentativen Tags eine Gruppe von repräsentativen Bildern zu finden. Dabei verwenden sie wiederum das k-means Clusterverfahren, um die beliebtesten Standpunkte zu finden, von denen die jeweilige Sehenswürdigkeit oder Landmarke aufgenommen wurde. Daraus entstehen weitere Bildgruppen, die nun nach “representativeness” sortiert werden. Dabei sind ähnliche Überlegungen zu finden, die auch in meiner Arbeit für das Maß der Popularität eines Ortes genutzt werden. Ein Faktor bei dem verwendeten Repräsentativitätsmaß ist unter anderem auch die Anzahl der Nutzer. Dabei spricht sich Kennedy ebenfalls gegen die Anzahl der Bilder als maßgebende Kenngröße aus und entscheidet sich dafür, die Anzahl der Nutzer zu verwenden. Andere Bewertungsfaktoren sind die Variabilität im Entstehungsdatum der Bilder, der Abstand der Bilder der verschiedenen Cluster untereinander und die visuelle Übereinstimmung der Bilder innerhalb des Clusters mit Hilfe der Extraktion von lokalen Bildmerkmalen. Das Ergebnis ist eine geordnete Liste von Bildgruppen ähnlich der Heatmap, die in Kapitel 5 näher beschrieben wird, allerdings nur für eine festgelegte Landmarke und einer festen Anzahl von Ansichten.

Werden zu den einzelnen Bildern zusätzlich detaillierte textuelle Beschreibungen der Ansichten, die in den Bildern festgehalten wurden, gespeichert, sind auch umfangreiche Textanalyseverfahren möglich. Insbesondere in den Datensätzen des Geograph Projekts bestehen diese aus mehreren umgangssprachlichen Sätzen. Ein beliebtes Mittel ist die Kookurenzanalyse, also die Fragestellung, ob zwei lexikalische Einheiten, in dem Fall Wörter, auffällig oft zusammen vorkommen. Dykes u. a. (2008) verwendet zum Beispiel zusätzlich sogenannte Treemaps. Diese werden verwendet, um die Zusammenhänge mit den vier häufigsten Szenentypen "Mountain", "Berg", "Siedlung" und "Strand" und ausgewählten zusätzlichen Beschreibungen, die häufig mit diesen Szenentypen verwendet werden (kategorisiert nach beschreibenden Adjektiven, auffindbare Objektbezeichnungen und Aktivitäten) darzustellen. Zusätzlich wird über eine Färbung der Knoten die vorherrschende geographische Lage der Bilder angezeigt. Analysen wie (Edwardes und Purves, 2007) versuchten auf Grundlage der Beschreibungen in Geograph das Konzept von "Place" weiter zu untersuchen. Dabei ging es ihnen hauptsächlich darum, welche Terme dazu benutzt werden, um bestimmte Orte wie Strände oder Berge zu beschreiben. Dieses Wissen wird dazu genutzt, Bilder mit besseren Schlagwörtern indexieren zu können. Die Ergebnisse solcher Studien haben zunächst einen starken visuellen Anspruch und versuchen, Beobachtungen für den Menschen sichtbar darzustellen, bevor diese in konkrete Modelle umgesetzt werden. Diese Modelle in Form von Konzepthierarchien könnten in Zukunft auch Empfehlungssysteme mit zusätzlichen Hintergrundwissen unterstützen, wenn es darum geht, ein übergeordnetes Konzept wie einen "Strand" mit einer Vielzahl von Bildern zu illustrieren. Dafür könnte eine Gewichtung der möglichen Unterkonzepte als zusätzlichen Filter bei der Bildauswahl verwendet werden. Im Gegensatz zu diesen Ansätzen wird sich meine Arbeit allerdings auf die räumliche Konzeptualisierung beschränken und sich mehr auf die Frage konzentrieren "wo" Bilder von

bestimmten Menschen erstellt wurden und weniger “wie” die Bilder betitelt wurden.

Sehr detailliert beschäftigt sich Girardin in seinen Arbeiten (Girardin u. a., 2007) (Girardin u. a., 2008) (Girardin u. a., 2009) mit der Analyse von sogenannten digitalen Fußabdrücken, die Menschen im geographischen Raum hinterlassen. Ein digitaler Fußabdruck entsteht zum Beispiel, wenn Mieträder an verschiedenen Orten abgeholt oder abgegeben werden und damit einen Ortswechsel des Fahrers anzeigen. Weitere Beispiele sind, wenn Eintrittskarten gekauft werden, wenn Menschen über ein Mobilfunknetz kommunizieren oder eben auch wenn an einem Ort ein georeferenziertes Bild erstellt wird. Ziel von Girardins Arbeiten ist es vor allem, die dynamischen Bewegungen von Touristen zu analysieren, um damit Fragestellungen zu beantworten, wie die Beliebtheit von Orten und typische Bewegungsmuster. Zur Analyse begnügt er sich allerdings mit einer Visualisierung der Daten, um es dem Nutzer zu ermöglichen, gewisse Phänomene in den Daten zu identifizieren, ohne sie begründen zu müssen. So findet er in (Girardin u. a., 2008) zum Beispiel heraus, dass sich touristische Bewegungsmuster je nach Herkunftsland grundsätzlich voneinander unterscheiden. In Rom bewegen sich italienische Touristen sehr viel freier durch die Stadt als andere Besucher der Stadt, die sich viel mehr geradlinig vom Kolosseum zur Vatikanstadt bewegen. Seine Arbeiten bieten daher den ersten Anhaltspunkt, dass sich verschiedene Nutzergemeinschaften über ihr räumliches Verhalten unterscheiden lassen. Meine Arbeit beschäftigt darüber hinaus mit der Frage, ob sich ein ähnliches räumliches Verhalten in einem Gebiet in andere Gebiete übertragen lassen.

3.5 Inhaltsbasierte Verfahren

Bei einem Verfahren, das mit Multimediadateien arbeitet, liegt es natürlich nahe direkt die Bildinhalten selbst zu analysieren. Wie in Kapitel 2 kurz erwähnt, versuchen die meisten inhaltsbasierten Verfahren, das Problem der Objekterkennung zu lösen, wobei eines der wenigen Objekte, das mit einer hohen Trefferquote und zusätzlich einer niedrigen Falscherkennung auskommt, das Gesicht ist.

Mein eigener Ansatz geht generell einen anderen Lösungsweg, dennoch könnten inhaltsbasierte Verfahren in Zukunft meinen Ansatz an bestimmten Stellen unterstützen. Grundsätzlich versuchen diese Verfahren, direkt in Bildern nach Ähnlichkeiten zu suchen [20]. Meistens werden einfache Eigenschaften, wie Helligkeit oder Struktur in den Bildern untersucht. Besonders beliebt sind Ansätze, die mit SIFT (Scale-invariant feature transform) arbeiten (Lowe, 1999). Dabei sucht man in Bildern nach Punkten, die invariant gegen Operationen, wie Skalierung, Drehung sind. Zusätzlich dazu sind diese Punkte auch invariant bei Beleuchtungsveränderung und affinen und dreidimensionalen Abbildungen. Dadurch lassen sich recht schnell Objekte in Bildern erkennen oder aus Bildern, die mehrere Ansichten eines Objektes aufzeigen, eine 3D Rekonstruktion erzeugen.

Snavely u. a. (2006) demonstrieren mit ihrem “Photo Explorer” zum ersten Mal eindrucksvoll, wie mit Hilfe von SIFT virtuelle Bildräume erschaffen werden können. Für jedes Bild wird die dreidimensionale Kameraposition zurückberechnet. Innerhalb dieser Sicht lassen sich die restlichen Bilder in dem lokalen Koordinatensystem der Kamera anzeigen, um dem Nutzer damit alternative Sichten aus der Ansicht des aktuellen Bildes heraus anzubieten. Durch diese dreidimensionale Repräsentation der Bilder lassen sich auch flüssige Übergänge von einem Bild zum nächsten berechnen und dem Nutzer das Gefühl vermitteln,

sich frei durch eine dreidimensionale Welt zu bewegen. Im Photo Explorer wird ein Bildabgleich lediglich für eine räumliche Navigation durch größere Bilddatenmengen verwendet. Die Auswahl der relevanten Bilder wird dem Benutzer selbst überlassen und das System unterstützt ihn dabei nur visuell.

Denkbar ist es, mit Hilfe von SIFT bestehende Clusterverfahren zu unterstützen, was seither aber wegen der hohen Komplexität der Berechnungen nicht umgesetzt wurde. Kennedy u. a. (2007) sprechen sich in ihrem Clusterverfahren noch gegen SIFT in ihrem Gruppierungsschritt aus, verwenden SIFT allerdings um die innere Homogenität der Bildgruppe zu berechnen. Die visuelle Übereinstimmung der einzelnen Bilder in einer Gruppe von Bildern ist für sie ein Faktor bei der Bestimmung ihres Repräsentativitätsmaßes.

4 Empfehlungssysteme

4.1 Informationsbedarf und Informationsfilterung

Ein Empfehlungssystem bietet die Möglichkeit, Informationen für einen bestimmten Nutzer zu filtern. Warum dies insbesondere für Bildersammlungen von Nutzen ist, wurde bereits in der Einleitung ausführlich besprochen. Man spricht auch davon, dass man versucht den persönlichen Informationsbedarf des Nutzers zu befriedigen. Der Informationsbedarf wird von Dirk Stelzer (Mertens und Back, 2001), im Lexikon der Wirtschaftsinformatik, folgendermaßen definiert:

Definition 4.1 *Der Informationsbedarf bezeichnet die Art, Qualität und Menge der Information, welche Aufgabenträger (Personen oder Organisationseinheiten) zur Erfüllung einer bestimmten Aufgabe benötigen.*

Um einen Informationsbedarf, ausgehend von einer bestimmten Person, zu lösen, werden in den meisten Fällen zwei unterschiedliche Ansätze verwendet: die Suche und die Empfehlung. Wie in Abbildung 4.1 illustriert, ist eine Suche sinnvoll, wenn sich die Anforderungen an die gewünschten Ergebnisse oft ändern, die Datenbasis an sich jedoch sehr statisch ist. Suchmaschinen stützen sich oft auf eine umfangreiche Indizierung der Eingangsdaten. Für georeferenzierte Datenquellen bietet gerade die räumliche Metainformation eine gute Möglichkeit der Indizierung: Daten können zusätzlich zu anderen Metainformationen nach ihrem Entstehungsort gruppiert werden. Im Allgemeinen sind Indexe

recht aufwendig zu erstellen, so dass ein Index nicht für jede Anfrage neu erstellt wird, sondern vor der Anfrage bereits vorhanden sein muss, damit die Anfrage effizient verarbeiten werden kann. Abgesehen von der Effizienz dieser Technik muss der Nutzer in der Lage sein, seinen Informationsbedarf in einer expliziten Form auszudrücken. Bei gängigen Suchmaschinen wird dem Nutzer dazu die Eingabe von Text ermöglicht. Bei der Suche geht es daher eher darum, einen sehr wechselhaften Informationsbedarf mit Hilfe einer Datenmenge zu befriedigen, die sich in überschaubaren Zeitabständen verändert.

Definition 4.2 *Eine Suchmaschine filtert eine Sammlung von Gegenständen nach Gegenständen, die zu einer explizit formulierten Suchanfrage passend sind. Es geht darum, einen kurzfristigen Informationsbedarf zu befriedigen.*

Allerdings wird ein Nutzer nicht immer in der Lage sein, seinen Informationsbedarf von sich aus hinreichend genau auszudrücken. Oder er befindet sich in einer Situation, in der er nicht explizit nach einer bestimmten Information sucht. Genau in diesen Situationen werden Empfehlungssysteme bevorzugt eingesetzt. Sie sind in der Lage auch dann Antworten zu geben, wenn keine explizite Anfrage vorliegt. Typisches Beispiel sind Nutzer von E-Commerce Plattformen, die ohne konkretes Ziel diese Plattform nutzen. Diese Art von Nutzern kann durch ein Empfehlungssystem zu Angeboten geleitet werden, die für ihn interessant sein könnten. Dazu stützt sich das Empfehlungssystem auf den langfristigen Informationsbedarf eines Nutzers, der sich aus seinem vergangenen Verhalten oder dem konkreten Kontext, in dem sich der Nutzer gerade befindet, ableiten lässt. Je nach Art des Empfehlungsdienstes kann dieser Kontext anders modelliert sein. Im Folgenden werden verschiedene Möglichkeiten dazu angesprochen. Konkrete Beispiele für einen solchen Kontext sind die momentan aufgerufene Webseite, die

soziale oder demographische Situation des Nutzers und, wie in dem Fall eines geografischen Empfehlungssystems, die geografische Position des Nutzers.

Definition 4.3 *Ein Empfehlungssystem filtert eine Sammlung von Gegenständen nach Gegenständen, die für den Nutzer von persönlichem Interesse sein können. Dazu nutzt das System Informationen über die Gegenstände und/oder den Nutzer. Von Interesse ist dabei auch der Kontext der Anfrage. Befriedigt wird ein langfristiger Informationsbedarf.*

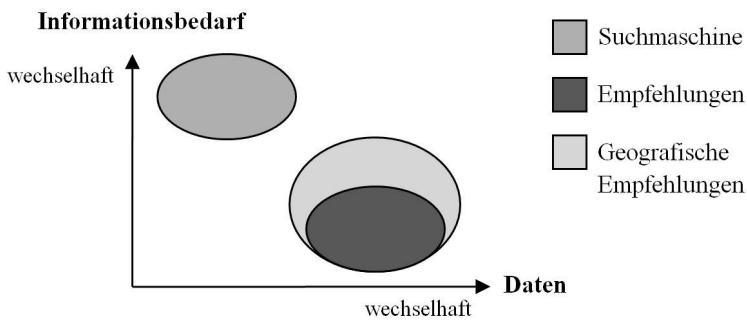


Abbildung 4.1: Befriedigung des Informationsbedarfs

Verschiedene Informationsfiltersysteme decken den Informationsbedarf des Nutzer auf ganz unterschiedliche Art und Weise ab. In den meisten Fällen findet man solche Systeme sicher im E-Commerce Bereich, wo die Systeme versuchen, den Nutzer bei Kaufsituationen zu unterstützen. Wie es am Beispiel Amazon zu sehen ist, wird oft eine Kombination verschiedenster Spielarten dieser Filtersysteme eingesetzt. Während die Suchmaschine dem Nutzer direkt und offensichtlich über die typische Eingabemaske präsentiert wird, arbeiten die verschiedensten Empfehlungssysteme im Hintergrund und werden dazu genutzt, jeden

Bereich der Webseite auf den momentanen Nutzer zu personalisieren. Betrachtet man beispielsweise einen bestimmten Artikel, den man über die Suchmaschine gefunden hat, werden durch eine Empfehlung sofort auch Alternativen oder Ergänzungen angeboten. Gerade auf den einzelnen Produktseiten lassen sich daher Empfehlungsprinzipien wie „wird oft zusammengekauft“ (Ergänzung) oder „Was kaufen Kunden, nachdem sie diesen Artikel angesehen haben?“ (Alternativen) finden. Die gängigste Art der Empfehlung ist allerdings „Das könnte Ihnen auch gefallen“ mit einer Auflistung von Gegenständen, die für sich gesehen eine gute Empfehlung darstellen und dem berechneten Informationsbedarf am nächsten kommt. Das primäre Qualitätsmerkmal für eine gute Empfehlung ist daher, den persönlichen Geschmack des Nutzers mit den berechneten Empfehlungen abzudecken. Meistens wird der Geschmack eines Nutzers durch explizite oder implizite Meinungsäußerungen errechnet und die Gegenstände mit diesem Nutzermodell verglichen. Eine häufige Kritik dieses Ansatzes ist allerdings, dass die ersten Ergebnisse einer Empfehlung sich meistens sehr ähnlich sind. Als Gruppe gesehen wirken sie daher sehr monoton (McNee u. a., 2006). Neuere Ansätze versuchen daher, diese Monotonie innerhalb einer Empfehlung zu verringern.

Während man bei Top-N Empfehlungen noch auf eine bessere Abdeckung des gesamten Empfehlungsraumes abzielt, ist in den letzten Jahren eine ganz verwandte Empfehlungsfrage zu diesem Problem aufgekommen, der sogenannten Mehr-Objekt-Empfehlungen (Schlieder, 2007). Diese Empfehlung versucht, eine ganz eigene Art des Informationsbedarfs zu befriedigen, nämlich eine Gruppe von Objekten zu finden, die möglichst gut ein bestimmtes Anfragekonzept beschreibt. Dabei soll eine Gruppe aus n unterschiedlichen Gegenständen bestehen.

Überträgt man diese Art von Empfehlungen auf den Anwendungsfall eines geografischen Empfehlungssystem kann zum Beispiel nach einer

Gruppe von Bildern gesucht werden, die möglichst gut die Eigenarten einer Stadt illustrieren. Das folgende Beispiel für die Stadt Bamberg soll dabei die Schwierigkeiten einer solchen Gruppenzusammenstellung verdeutlichen. Verwendet man eine einfache Suchmaschine, um nach Bildern mit Hilfe des Namens der Stadt zu suchen, trifft man oft auf sehr ähnliche Bilder in den ersten Suchtreffern. Abbildung 4.2 zeigt ein typisches Ergebnis für die Suche mit Hilfe des Suchbegriffs „Bamberg“. In diesem Fall wurde die Suche auf der Onlinegalerie Flickr durchgeführt. Wenn wir ein weiteres Qualitätskriterium für den Informationsbedarf festlegen, tritt ein weiteres Problem auf: Die Auswahl der Bilder soll insgesamt gesehen eine gute bildliche Konzeption der Stadt Bamberg darstellen. Das heißt, das System soll in der Lage sein, eine Auswahl von Bildern zu berechnen, in der möglichst viele typische Bilder enthalten sind; paarweise betrachtet sollen die einzelnen Bilder aber möglichst unterschiedlich sein. Daraus kann man ableiten, dass die Auswahl bestenfalls keine Duplikate bzw. redundante Informationen aufzeigen soll. In dem Beispiel für die Suchanfrage nach dem Begriff „Bamberg“ zeigen allerdings neun von 28 Bildern das „Alte Rathaus“. Für eine variabelere Bildauswahl wäre es sinnvoll, acht dieser Bilder mit Aufnahmen von anderen historischen Gebäuden der Stadt Bamberg zu ersetzen. Für eine gute Auswahl für Bamberg erwartet man zum Beispiel auch eine Ansicht des Bamberger Doms, Bilder der zahlreichen Kirchen oder der Altenburg.

Ein extremeres Beispiel bietet die Bildersuche über Google nach „Paris“ an, bei dem die Hälfte der Bildaufnahmen unter den ersten 28 Treffern den Eiffelturm zeigen. Eine Suchmaschine optimiert die Ergebnisliste alleine auf eine möglichst gute Übereinstimmung mit dem eingegebenen Suchwort, das Qualitätskriterium wird dabei Relevanz genannt. Da sowohl das „Alte Rathaus“ als auch der „Eiffelturm“ eine hohe Relevanz für die beiden Städte zeigen, tauchen sie vermehrt in den Sucher-



Abbildung 4.2: Beispiel für Informationsfilterung durch eine Suchanfrage nach „Bamberg“ auf Flickr

gebissen auf. Allerdings wird eine Eigenschaft vernachlässigt, die die Qualität einer Bildgruppe verbessert, die Variabilität.

Die Forschungsrichtung, die sich mit Informationsfilterung und Suchmaschinen beschäftigt, hat dafür einen eigenen Begriff geprägt und verwendet statt dem Begriff der Variabilität den Begriff der Diversität. In den meisten Arbeiten aus diesem Forschungsbereich geht es darum, eine Menge von Elementen zu finden, die zum einen relevant sind, zum anderen aber auch die Bandbreite der Ergebnisse aufzeigen und daher untereinander nicht allzu ähnlich sind. Es gibt verschiedene Ansätze, dies zu erreichen. Einige davon finden sich bereits 1964 in Arbeiten von Goffman (Goffman, 1964), der erwähnt, dass die Relevanz von Gegenständen abhängig von den bereits gewählten berechnet werden sollte. Diese erste Beschreibung von Diversität in Suchanfragen wurde auch von Carbonell und Goldstein (1998) wieder aufgegriffen. Sie appelliert für eine zusätzliche Art der Gewichtung von Objekten in einer Ergebnismenge; insbesondere in Situationen, in denen sich zahlreiche gute Ergebnisse mit hoher Redundanz finden lassen. In diesen Situationen lässt sich durch den bewussten Verzicht auf die maximale Relevanz, dennoch eine interessante Ergebnismenge finden. Um dies zu erreichen, definierten Carbonell und Goldstein ein zusätzliches Maß, das bestimmt, wie groß der Mehrwert ist, den ein Objekt zu einer bestehenden Menge von Objekten hinzuzufügt. In ihrem Fall nannten sie dieses zusätzliche Maß „novelity“, sprich das Maß an „Neuigkeit“, die dieses Objekt im Vergleich zu den anderen Gruppenmitgliedern aufweist. In den meisten Fällen wird dies ein Maß der Unähnlichkeit zweier Objekte sein und wird oft auch mit der negativen Ähnlichkeit vereinfacht dargestellt. Weist ein Objekt o zu einem anderen Objekt n die Ähnlichkeit $similarity(o, n)$ auf, so wird seine „Neuartigkeit“ $novelity(o, n)$ über die Funktion $novelity(o, n) = 1 - similarity(o, n)$ bestimmt. Umfangreichere Novelitymaße lassen sich zum Beispiel in (Clarke u. a., 2008)

finden. Angewendet werden die beiden Maße Relevanz und Novelty durch die Maximierung der Linearkombination von beiden. Deselaers u. a. (2009) verwendet ebenfalls eine zusätzliche Gewichtung durch ein Diversitätsmaß, um die Suche nach Bildern zu verbessern und Ergebnislisten mit vielen ähnlichen Bildern zu verhindern. Verglichen wird ein Greedy Ansatz, der in ähnlicher Form schon in Garbonell und Goldstein beschrieben wurde, mit zwei Ansätzen, die Verfahren aus der dynamischen Programmierung verwenden, um den Raum der möglichen Kombinationsmöglichkeiten besser abzudecken. Darüber hinaus wurden die Ergebnisse mit einem Clusteralgorithmus verglichen, der ohne zusätzliche explizite Definition eines Diversitätsmaßes auskommt. Clusterung als Vorverarbeitung vor einer Auswahl wird auch in meinem Ansatz verwendet, um Ergebnisse mit höherer Diversität zu erhalten. Allerdings auf Grundlage der vorhandenen geografischen Metainformationen, die in den zuvor erwähnten Ansätzen keine Rolle spielen.

Duplikate und redundante Informationen stören nicht nur die Qualität von Suchergebnissen, sie haben auch ähnlich negative Effekte bei der Berechnung von Empfehlungen. McNee und andere beschreiben zum Beispiel im Jahr 2006 den Effekt des „Ähnlichkeitslochs“ (McNee u. a., 2006), in dem man bei den Empfehlungen auf Amazon sehr schnell gefangen sein kann und dadurch nur noch recht ähnliche Empfehlungen angeboten bekommt. Besonders gegenstandsbasierte kollaborative Verfahren sind anfällig für solche Effekte. Rashid u. a. (2002) beschreibt sie als Cluster von ähnlichen Objekten, in denen sich dieser Ansatz verfängt. Ein Nutzer, der einen Star Trek Film gut bewertet, wird von Amazon mit hoher Wahrscheinlichkeit alle anderen Star Trek Filme empfohlen bekommen. Dies wird zwar zu einer hohen Trefferquote bei den Empfehlungen führen, da die meisten Fans einer Serie eine sehr hohe Vorliebe für Filme derselben Serie haben. Allerdings kann in diesen Fällen niemals der komplette Raum an möglichen Filmen ausge-

erschöpft werden, in dem sich auch Empfehlungen mit einem hohen Überraschungseffekt befinden. Anders ausgedrückt, stellt eine Empfehlung eines Star Trek Films für einen Star Trek Fan kein besonders hohes Novum dar. Ähnliche Effekte werden auch in anderen Arbeiten beschrieben. Ali und van Stam (2004) nennen dies den Portfolio-Effekt in ihrem Empfehlungssystem (TiVo). Herlocker u. a. (2004) beschreibt den Bedarf an „Novelty“ bei Empfehlungen in anderen Domänen abseits von Filmempfehlungen. Schlieder (2007) beobachtet fehlende Variabilität bei Empfehlungen auf Amazon. Als Beispiel führt er eine ergänzende Empfehlung für einen Reiseführer Bali an. In diesem Fall werden ausschließlich andere Reiseführer empfohlen, die ebenfalls Bali beschreiben. Bei einer guten Empfehlung wurde man eher Reiseführer über andere Ländern oder ergänzende Produkte aus dem kompletten Produktkatalog, die für eine Reise nach Bali sinnvoll sind, erwarten. Der Autor führt dabei Bücher über den Hinduismus oder Surfing als bessere Alternativen an, die sowohl den kulturellen wie auch den sportlichen Aspekt des Landes besser beleuchten würden. Diese Problematik ist vor allem für Top-N Empfehlungen bekannt, bei denen es darum geht, eine Liste aus n Objekten zu empfehlen, anstatt nur eine einzige Empfehlung zu finden.

Es gibt verschiedene Ansätze, um Top-N Empfehlungen divergenter zu machen. Ziegler u. a. (2008) verwendet zum Beispiel zusätzliches Hintergrundwissen in Form einer Taxonomie, die Produktsparten hierarchisch anordnet, um zu verhindern, dass zu ähnliche Objekte in die Ergebnismenge übernommen werden. Dazu verwendet er eine Heuristik, die sich an den osmotischen Prinzipien von Zellen orientiert und ein Maß, das die Neuartigkeit (Novelty) des Gegenstands bestimmt. Gegenstände einer bestimmten Sparte werden daher nur bis zu einer bestimmten Sättigung in die Ergebnismenge übernommen. Die Ähnlichkeit von Gegenständen lässt sich dabei als Abstand auf der gewählten Taxonomie

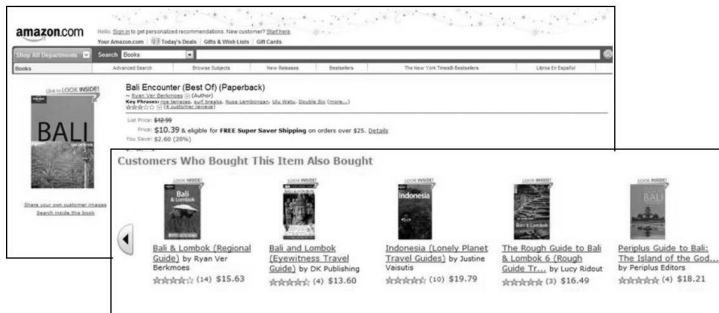


Abbildung 4.3: Ähnlichkeitsfalle bei gegenstands-basierten kollaborativen Filtern auf Amazon

bestimmen. McGinty und Smyth (2003) verweisen auf viele verschiedene Techniken wie man bestehende Empfehlungslisten abwechslungsreicher gestalten kann, ohne sich zu weit von den Vorlieben des Nutzers zu entfernen. Dazu übernimmt das Verfahren schrittweise neue Gegenstände aus den vorläufigen Empfehlungen ausschließlich dann, wenn sie sich von bereits ausgewählten Gegenständen unterscheiden. McGinty nennt diese Verfahren ähnlichkeitsbasierte vergleichende Empfehlungs-algorithmen. Zhang (2009) löst das Diversitätsproblem mit drei unterschiedlichen Ansätzen: Optimierung einer Zielfunktion, die aus einer Linearkombination aus Ähnlichkeit der Gegenstände zum Nutzer und Abstände der Gegenstände untereinander besteht (Zhang und Hurley, 2008), einer Partitionierung des Nutzerprofils und einem Clusterverfahren mit Dimensionsreduktion, wie es auch in (Zhang und Hurley, 2009) näher beschrieben wird.

Ein recht neuer Ansatz zum Thema Diversität beschäftigt sich mit der Idee, Ergebnislisten von georeferenzierten Dokumenten möglichst ausgeglichen über den Raum zu streuen. Man spricht von räumlicher Diversität (Clough u. a. (2006), Purves und Jones (2006)). Dies macht vor

allem dann Sinn, wenn die Suchanfrage einen geografischen Hintergrund hat, wie etwa „Schlösser in England“. Nutzerstudien (Tang und Sanderson, 2010) auf Amazons Mechanical Turk haben dabei gezeigt, dass Nutzer Ergebnismengen präferieren, die eine hohe räumliche Divergenz aufzeigen. Definiert wird in den meisten Fällen ein Divergenzmaß, das sowohl die Relevanz eines Objekts als auch dessen räumlichen Abstand zu allen bereits in die Ergebnismenge übernommenen Objekten berechnet. Das Objekt das momentan die höchste Divergenz aufzeigt, wird in die Ergebnismenge übernommen und die Divergenz der restlichen Objekte wird daraufhin neu berechnet. Beispiele für räumliche Divergenzmaße werden in Paramita u. a. (2009) beschrieben.

In Schlieder (2007) wird zum ersten Mal eine Mehr-Objekt-Empfehlung beschrieben, die ähnliche Qualitätsmerkmale wie die Divergenz und Relevanz im Information Retrieval Bereich aufzeigt. Die entsprechenden Maße für Typikalität und Variabilität lassen sich dabei wie folgt abstrakt beschreiben:

- **Typikalität:** Jedes einzelne Element in der Ergebnismenge soll für sich gesehen repräsentativ für ein bestimmtes Konzept sein. Im Fall einer geografischen Bildempfehlung wird dies ein räumliches Konzept wie typische Ansichten in einer Stadt sein. Zusätzlich soll die Gruppe als Ganzes gesehen möglichst nahe an einer prototypischen Vorstellung des Konzeptes liegen.
- **Variabilität:** Variabilität bedeutet, dass die Bilder paarweise gesehen möglichst unterschiedlich sein sollen. Man kann diese Eigenschaft auch als Streuung bzw. Diversity beschreiben, wie es sich vor allem im Information Retrieval Bereich etabliert hat.

Alle Qualitätsmerkmale zusammen sorgen dafür, dass die Menge insgesamt gesehen repräsentativ für ein bestimmtes Konzept ist und die einzelnen Bilder sich ähnlich aber nicht gleich sind.

Mehr-Objekt-Empfehlungen zu finden, die jeweils die erste Eigenschaft (Typikalität) maximieren, ist recht trivial, da man nur den typischsten Vertreter (oder verschiedene Instanzen davon) n mal in die Gruppe aufnehmen muss (wie n Aufnahmen des bedeutendsten Bauwerkes einer Stadt). Allerdings würde diese Gruppe eine sehr geringe Variabilität aufweisen, da die Mitglieder keinerlei Unterschiede aufzeigen. Würde man sich zu sehr auf eine hohe Variabilität konzentrieren, würde der eigentliche Kerngedanke der Bildgruppe zu sehr durch die extrem unterschiedliche Bildauswahl verschwimmen. Die Erstellung einer guten Mehr-Objekt-Empfehlung ist daher ein Optimierungsproblem der zwei genannten Eigenschaften.

4.2 Weitere Überlegungen zu Typikalität und Variabilität

Nach der Prototypentheorie kann man sagen, dass sich die Typikalität durch die Nähe zu einem Prototypen definieren lässt, während die Variabilität dessen Abstand beschreibt. In den bestehenden Ansätzen zur Divergenz ist zu erkennen, wie das Problem eine Menge von Objekten zu finden, meist auf ein Optimierungsproblem dieser beiden Werte zurückgeführt wird. Allerdings wird dadurch ein entscheidender Faktor vernachlässigt: der Nutzer. Die vorhandenen Präferenzen des Nutzers, der eine entsprechende Anfrage stellt, lassen sich schlecht in einfache Abstandsmaße fassen. Ein simples Beispiel soll kurz illustrieren, warum ein reiner Optimierungsalgorithmus nicht immer die besten Ergebnisse für einen bestimmten Nutzer erzeugen muss. Gehen wir davon aus, dass die Variabilität durch den durchschnittlichen Abstand zu einem Prototypen und die Typikalität durch den Abstand des Schwerpunktes zum Prototypen bestimmt wird. Abbildung 4.4 zeigt ein Beispiel in einem zweidimensionalen Eigenschaftenraum. Dabei ist das zwei-

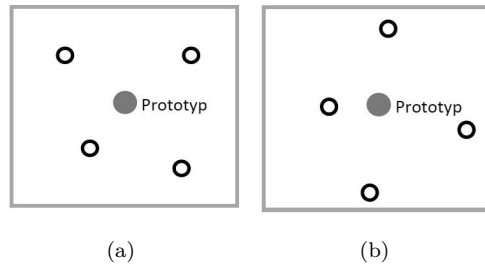


Abbildung 4.4: Multi-Objekt-Empfehlung in einem zweidimensionalen Eigenschaftensraum.

te Beispiel nur durch eine Drehung der Element um den Prototypen entstanden. Diese Drehung ändert nichts an den beiden Werten Typikalität und Variabilität, kann allerdings für einen bestimmten Nutzer einen bedeutenden Unterschied machen. Grundlegend ist es algorithmisch nicht greifbar, wie sich Nutzer in dem modellierten Gegenstandsraum verhält, wenn man die Ergebnisse nur von den Eigenschaften der Gegenstände abhängig macht. Daher folgt mein Ansatz dem Prinzip eines nutzerbasierten kollaborativen Filters. Dieser Ansatz ist genau in solchen Situationen sinnvoll, in denen sich das Nutzungsverhalten der Nutzer nicht im Vorfeld explizit modellieren lässt, sondern sich ihr Verhalten nur durch das Verhalten von ähnlichen Nutzer vorhersagen lässt. Neben den nutzerbasierten kollaborativen Empfehlungsansätzen lassen sich noch eine Vielzahl anderer Empfehlungsmethoden finden, die im nächsten Abschnitt kurz systematisiert vorgestellt werden sollen.

4.3 Systematisierung von Empfehlungssystemen

1992 wurde zum ersten Mal ein System mit dem Namen Tapestry veröffentlicht, welches Informationen effektiver nach Relevanz für den Nutzer selbst filtert, indem der Mensch in den Prozess mit eingebunden

wird (Goldberg u. a., 1992). Damals verwendete man bereits den Begriff kollaborative Filterung, der in späteren Arbeiten teilweise für andere personalisierte Filterungen übernommen wurde, auch wenn keine wirkliche Zusammenarbeit bzw. Kollaboration mehr stattfand. Heutzutage spricht man daher von Empfehlungssystemen als Überbegriff für Systeme, die eine Personalisierung der Inhalte ermöglichen. Um eine personalisierte Empfehlung auszusprechen, wurden eine Vielzahl von Ansätzen entwickelt. Ebenso vielzählig waren allerdings auch die Versuche, die Ansätze zu kategorisieren (Resnick und Varian, 1997) (Schafer u. a., 1999) (Terveen und Hill, 2001). Eine gute Übersicht findet man in Burke (2002). Burke beschäftigt sich dabei nicht zu sehr mit den Benutzerschnittstellen und den Interaktionsmöglichkeiten, die bei anderen Kategorisierungen eine maßgebende Rolle spielen, sondern mehr mit den Eingangsdaten und den verwendeten Methoden. Er unterscheidet dabei drei Parameter: (1) das Hintergrundwissen, das ein Empfehlungssystem zur Verfügung hat, (2) die Eingabedaten, die der Nutzer übermittelt und (3) die Verarbeitung dieser zwei Datenquellen. Damit benannte er fünf verschiedene Methoden: Kollaboration (Collaborative), inhaltsbasiert (content-based), demographisch (demographic), nutzenbasiert (utility-based) und wissensbasiert (knowledge-based). Tabelle 4.1 gibt eine Übersicht über die 5 Empfehlungskategorien nach Burke, U beschreibt dabei die Menge der Nutzer und I die Menge der Gegenstände.

Im Folgenden werden die inhaltsbasierten, nutzenbasierten und wissensbasierten unter dem Oberbegriff gegenstandsbasiertes Empfehlungssystem nur kurz vorgestellt, da sie in der Literatur recht wenig Einfluss besitzen. Interessant für die Wissenschaft wie auch für die praktische Umsetzung sind hauptsächlich die kollaborativen Ansätze. Hierbei handelt es sich um lernende Ansätze, die mit steigender Beteiligung der Nutzer bessere Ergebnisse liefern.

Methode	Beschreibung
Kollaboration	Bestimmt die ähnlichsten Nutzer aus U zum jeweiligen Nutzer u und extrapoliert ihre Bewertungen für einen bestimmten Gegenstand i
Inhaltsbasiert	Erzeugt eine Klassifikation, die auf das Bewertungsverhalten von Nutzer u passt und wendet sie auf i an
Demographisch	Findet Nutzer, die demographisch ähnlich zu Nutzer u sind und extrapoliert ihre Bewertungen für einen bestimmten Gegenstand i .
Nutzenbasiert	Wendet eine Nutzen-Funktion auf die Gegenstände an und sortiert die Gegenstände dementsprechend
Wissensbasiert	Gleiche die Gegenstände i mit den Bedürfnissen und Interessen von u ab

Tabelle 4.1: Kategorisierung von Empfehlungssystemen nach Burke
2002

Ausgehend von den Eingangsdaten ergibt sich eine Kategorisierung wie in Abbildung 4.5 dargestellt. Sowohl Benutzer als auch Gegenstände können durch geeignete Eigenschaften beschrieben werden. Je nachdem welche Daten zur Verfügung stehen, lassen sich entweder demographische oder gegenstandsbierte Ansätze wählen. Gibt der Nutzer ein Feedback über Gegenstände ab, entsteht die Kreuztabelle zwischen dem Benutzer des Systems und den Gegenständen. Liegen Nutzerbewertungen vor, lassen sich diese für kollaborative Filter nutzen, die es jeweils in zwei Ausprägungen gibt, je nachdem ob man die Zeilen oder Spalten dieser Matrix miteinander vergleicht. Man bezeichnet diese Ansätze dann dementsprechend nutzerbasierte kollaborative Filter (user-based collaborative filter) oder gegenstandsbierte kollaborative Filter (item-based collaborative filter).

4.4 Überblick über die einzelnen Ansätze

4.4.1 Kollaboration

Kollaborative Empfehlungssysteme sind vor allem durch die Forschungsgruppe um Paul Resnick bekannt geworden. Im Jahr 1994 stellte Resnick sein Empfehlungssystem (GroupLens) für Nachrichten einer Online-Plattform vor (Resnick u. a., 1994), die die Benutzer auf einer Skala von 1–5 bewerten konnten. Heutzutage ist das ein gängiges Bewertungsschema, das man zum Beispiel auf Amazon oder Netflix finden kann. Dieses Nutzerfeedback wird dann dazu genutzt, Ähnlichkeiten zwischen den Benutzern zu finden. Aus den Bewertungen von ähnlichen Nutzern wird dann eine Bewertung für einen, dem momentanen Benutzer bislang unbekannten, Artikel berechnet. Wegen der gemeinschaftlichen Art wie die Bewertungsdaten durch die Benutzer zusammengetragen werden, nennt man diese Art der Filterung kollaboratives Filtern.

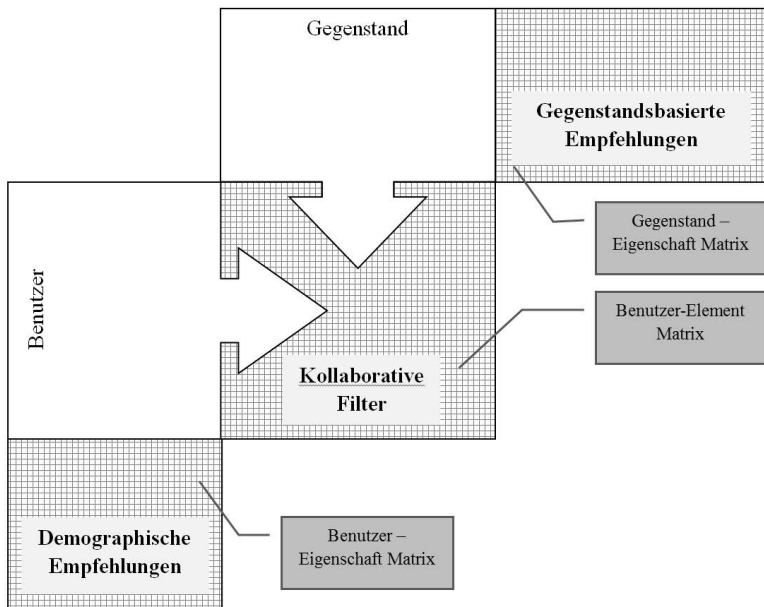


Abbildung 4.5: Übersicht über die Eingangsdaten und zugehörigen Empfehlungsansätze

Problem solcher Systeme ist meistens die niedrige Motivation der Nutzer, sich aktiv mit expliziten Bewertungen zu beteiligen (Rashid u. a., 2006). Eine Empfehlung kann allerdings nur dann ausgesprochen werden, wenn zuvor eine kritische Masse an abgespeicherten Bewertungen überschritten wurde. Kritisch bedeutet in dem Fall, dass genügend Nutzer gemeinsam mit anderen Nutzern eine Bewertung für eine gewisse Anzahl von Gegenständen abgegeben hat. Da die Ähnlichkeiten der Nutzer über die Bewertungen auf denselben Gegenständen berechnet werden ist ein Nutzer für das Empfehlungssystem erst sichtbar, wenn er mindestens einen Gegenstand bewertet hat, der schon zuvor bewertet wurde. Daraus ergeben sich mehrere grundlegende Probleme kollaborativer Ansätze:

- *New User Problem*: Wie schon erwähnt, lebt ein kollaboratives System von den Bewertungen der Nutzer. Meldet sich ein Nutzer neu an, können ihm noch keine Empfehlungen ausgesprochen werden. Viele Systeme verlangen daher von einem neuen Nutzer eine bestimmte Anzahl von Bewertungen, bevor dem Nutzer eine Empfehlung präsentiert wird. Bei Movielens sind es zum Beispiel 15 Bewertungen, bei anderen Systemen sogar bis zu 100. Ähnliche Probleme, wie komplett neue Nutzer, haben die sogenannten grauen Schafe („grey sheeps“ nach Su und Khoshgoftaar (2009)).
- *New Item Problem*: Ein ähnliches Problem hat das System, wenn neue Gegenstände in die Datenbasis übernommen werden. Gegenstände können ebenfalls nur dann empfohlen werden, wenn mindestens ein Nutzer eine Bewertung abgegeben hat. Daher ist es sinnvoll, dem Nutzer diese immer wieder zur Bewertung anzubieten. Verwandt mit diesem Problem ist auch, dass Synonyme die Genauigkeit der Empfehlungen erheblich stören. Dabei teilen sich die Bewertungen der Nutzer über diese Gegenstände auf, obwohl es sich um bei den unterschiedlichen Bewertungen um die

gleiche Aussage handelt. Ein Beispiel aus Amazon wäre: Derselbe Roman in unterschiedlicher Ausführung (Hardcover, Softcover, illustrierte Fassung, ...) erzählt dennoch dieselbe Geschichte.

- *Ramp on Problem*: Wird ein kollaborativer Filter in ein System mit bestehender Datenbasis übernommen, kommen sowohl das New Item wie auch das New User Problem zum Tragen und man hat eine typische Anlaufsituation, in der nur schwer Empfehlungen in guter Qualität möglich sind.
- *Lemming Effekt*: Gegenstände, die häufig von Menschen hoch bewertet werden, haben eine höhere Chance, von kollaborativen Filtern empfohlen zu werden. Dies umgeht man durch die Abwertung dieser sogenannten Bestseller. Auch von Nutzerseite kann es zu Effekten führen, die sich störend auf die Qualität der Bewertungen auswirken. So haben Menschen, die extrem viele Bewertungen abgeben, eine hohe Chance mit anderen Menschen ähnlich zu sein. Gerade neue Nutzer werden oft mit diesen Nutzern zusammenfallen und somit werden die Empfehlungen auch von diesen überdurchschnittlich aktiven Nutzern bestimmt. Hält sich nun ein Nutzer an diese Empfehlungen, hat dies zur Folge, dass der neue Nutzer ebenfalls immer ähnlicher zu dem bestehenden Nutzer wird. Man folgt sozusagen dem Kaufverhalten bestimmter Menschen wie ein Lemming, ohne seine eigene Nische zu finden. Dieser Effekt kann auch gezielt als eine Attacke auf die Empfehlungsergebnisse angewendet werden, der sogenannten „Lockvogel-Attacke“ (Shilling Attack nach Resnick und Varian (1997)), in der man gezielt viele bevorzugte Artikel hoch bewertet und verwandte Artikel besonders negativ.

In der Literatur unterscheidet man zwei grundsätzliche Arten von kollaborativen Systemen: benutzerbasierte und gegenstandsbasierte kollaborative Filter. Der GroupLens Ansatz beschreibt einen benutzerba-

sierten Ansatz, auf den auch diese Arbeit ausgerichtet ist. Technisch gesehen wird ein Nutzer durch seine abgegebenen Bewertungen repräsentiert, also dem entsprechenden Zeilenvektor in der Nutzer-Gegenstand-Matrix aus Abbildung 4.5. Eine erfolgreiche Alternative findet man auf der E-Commerce Plattform Amazon. Dort werden Empfehlungen auf Grundlage von gegenstandsbasierten Ansätzen generiert (Sarwar u. a., 2001) (Karypis, 2001). Dieser Ansatz berechnet die Ähnlichkeit der Gegenstände (Linden u. a., 2003). In Analogie zu den benutzerbasierten Verfahren werden ebenfalls Einträge der Nutzer-Gegenstand-Matrix verwendet, um die Gegenstände miteinander zu vergleichen. Allerdings stehen diese in der entsprechenden Spalte. Durch gegenstands-basierte kollaborative Filter lässt sich die Frage beantworten, welche Gegenstände oft zusammen bewertet wurden, wenn sich der Nutzer schon auf einen oder mehrere Gegenstände festgelegt hat. Dieser Ansatz ist daher weniger persönlich als das Ergebnis eines nutzerbasierten Ansatzes, da das Ergebnis nur von den ausgewählten Gegenständen abhängig ist. Aus diesem Grund konzentriert sich diese Arbeit darauf, einen benutzerbasierten Ansatz zu finden. In vielen Arbeiten werden kollaborative Ansätze allerdings auch in speicherbasiert (memory-based) und modellbasiert (model-based) kategorisiert (Su und Khoshgoftaar, 2009). Das klassische Verfahren gehört dabei zu den speicherbasierten Ansätzen, die kein explizites Modell lernen, sondern mit den gespeicherten Nutzerbewertungen nach Ähnlichkeiten unter den Nutzern suchen. Modellbasierte Systeme erstellen zunächst ein Modell ausgehend von den Nutzungsdaten, um dann eine Vorhersage mit Hilfe dieses Modells zu berechnen. Gängige Methoden, um dieses Modell zu berechnen, sind Bayes'sches Netze (Miyahara und Pazzani, 2000) oder auch Clusterverfahren, die auch in dieser Arbeit für die Berechnung eines Modells verwendet werden (Kapitel 5). Im Hinblick auf diese Kategorisierung kann der hier besprochene Ansatz als eine Mischung aus einem modellbasierten und einem speicherbasierten Verfahren angesehen werden.

In den letzten Jahren wurde der kollaborative Ansatz umfassend erforscht und Lösungen zu den verschiedensten Problemen diskutiert (Sarwar u. a., 2000) (Karypis, 2001) (McLaughlin und Herlocker, 2004) (Zhang und Pu, 2007) (Park und Tuzhilin, 2008) (Zhang und Hurley, 2008). Eine aktuelle und umfangreiche Übersicht über alle kollaborativen Ansätze findet man in Su und Khoshgoftaar (2009). Einige Arbeiten beschäftigen sich auch mit dem Unterproblem ganze Bildgruppen zu empfehlen (nach Burke (2002) eine Top-N-Empfehlung), meistens mit Hilfe von zusätzlichen semantischen Informationen der Objekte selbst (Ziegler u. a., 2004) (Konstan u. a., 2006) (Ziegler u. a., 2008).

In den meisten Arbeiten zu kollaborativen Verfahren wird von dem gleichen Anwendungsfall wie Resnick ausgegangen: Nutzer bewerten Gegenstände auf einer festen Skala und das System versucht, eine Wertung hervorzusagen. Dabei ist ein großer Schwerpunkt, die Qualität der Empfehlungen immer weiter zu verbessern.

Shi u. a. (2009) stellt erstmals ein Verfahren vor, das Ähnlichkeiten auf verschiedenen Spektren des möglichen Verhaltens berechnet. Durch diese Unterteilung ist es möglich, eine detaillierte Aussage über die Ähnlichkeiten von Benutzern zu bestimmen. Diese Aufteilung des Verhaltens auf verschiedene Spektren (oder von Shi auch Pools genannt) wurde bis heute noch nicht räumlich interpretiert. Anstatt die räumlichen Entscheidungen weltweit zu betrachten, untersucht meine Arbeit die Auswirkungen auf die Qualität der Empfehlung, wenn kleinere abgeschlossene Gebiete getrennt voneinander betrachtet werden.

Vorteile:

- Personalisierte Ergebnisse möglich, insbesondere mit dem nutzerbasierten Verfahren
- Gute theoretische und logische Grundlage
- Umfangreiche Forschungsarbeiten

- Neuer Strom an Bewertungen hilft dem System, die Empfehlungen für einen Nutzer immer wieder neu zu bewerten und zu verbessern. Qualität steigt mit der Anzahl der Bewertungen

Nachteile:

- Kaltstartprobleme wenn neuen Nutzern und neuen Gegenständen aufkommen
- Ergebnisse teilweise schwer für den Nutzer nachvollziehbar

4.4.2 Gegenstandsbasierte Filterung

Reine gegenstandsbasierte Ansätze berechnen die Ähnlichkeiten von Objekten auf Grundlage ihrer Produkteigenschaften. Diese Verfahren können daher sofort auf neue Gegenstände angewendet werden und haben dafür die meisten Parallelen zu den Techniken, die auch im Information Retrieval bekannt sind (Anand und Mobasher, 2003). Diese Ähnlichkeitsbeziehungen zwischen den Gegenständen sind allerdings unabhängig von den Nutzern, sodass die Ergebnisse weitgehend unpersönlich sind. Um eine Personalisierung zu erhalten, können Gegenstände empfohlen werden, die die meisten Ähnlichkeiten zu bereits konsumierten Gegenständen des Nutzers aufweisen. Diese Einschränkungen in der Personalisierung machen dann Sinn, wenn die Gegenstände nur eine kurze Gültigkeitsdauer aufweisen und sie nicht erst dann empfohlen werden, wenn genügend Meinungen über sie zur Verfügung stehen. Das in dieser Arbeit vorgestellte Verfahren weist eine ähnliche Eigenschaft, auf indem es neue Objekte mit bestehenden Entscheidungen gruppiert, die bereits von Nutzern bewertet wurden. Andere Ansätze versuchen, ein Nutzermodell auf Grundlage der konsumierten Gegenstände zu bestimmen, um daraufhin passende Gegenstände zu dem erzeugten Nutzerprofil zu finden. Maidel u. a. (2008) stellt ein solches Verfahren vor, das die Gegenstände in eine Ontologie einordnet und das

Nutzerprofil als eine Gewichtung dieser Ontologie darstellt. Dadurch lassen sich Abstände zwischen Gegenstand und den Vorlieben des Nutzers als hierarchische Distanz zwischen den Konzepten des Gegenstands und den Konzepten im Nutzerprofil berechnen. Nachteil dieses Verfahrens ist die Anfälligkeit Ähnlichkeitslöcher bzw. einen Lemming-Effekt zu erzeugen (Anand und Mobasher, 2003) (Anand u. a., 2007).

Folgt man der Kategorisierung von Burke, zählt Tapestry nicht zu den kollaborativen Filtern, wie die Benennung nach Goldberg vermuten lässt, sondern ist ein Beispiel eines nutzenbasierten Empfehlungssystems. Die Nutzer fügen über Annotationen neue Schlagwörter zu den Inhalten des Systems hinzu und definieren dann über einen booleschen Ausdruck (ähnlich SQL) eine Funktion, mit der das System entscheiden kann, ob eine bestimmte Information für einen Nutzer relevant ist oder nicht.

Vorteile:

- Keine Anlaufschwierigkeiten

Nachteile:

- Jeder Gegenstand muss in einem vorhandenen Modell beschrieben sein. Beispielsweise muss ein Gegenstand in eine umfangreiche Ontologie eingeordnet werden
- Persönlicher Geschmack des Nutzers und sein langfristiger Informationsbedarf kann nicht richtig erfasst werden

4.4.3 Demographische Filterung

Vom Grundgedanken sind demographische Filter recht ähnlich zu nutzerbasierten kollaborativen Filtern. Allerdings werden die Ähnlichkeiten nicht über die Bewertungen der Nutzer berechnet, sondern über die Eigenschaften der Personen, was im einfachsten Fall demographische

Daten sind. Dadurch wird das New User Problem umgangen, allerdings werden die Ergebnisse weniger differenziert als bei den kollaborativen Filtern. Neuere Ansätze beschränken sich daher nicht nur auf demographische Unterschiede, wie Alter oder Geschlecht, sondern versuchen, die Persönlichkeit der Nutzer zu bestimmen. Dabei bedient man sich vor allem Erkenntnissen aus der Psychologie, die bestätigen, dass eine starke Korrelation zwischen der Persönlichkeit eines Menschen und seinem Geschmack besteht. Besonders gilt dies für den Musikgeschmack (Kemp, 2004) (Rentfrow u. a., 2003).

Die meisten Ansätze verwenden dazu ein Persönlichkeitsquiz, das versucht, Alltagssituationen abzufragen, in denen sich Menschen emotional anders verhalten. Erforscht werden aber auch Interfaces, die aus der Spieltheorie stammen (Dunn u. a., 2009). Aus den Antworten des Nutzers wird seine Persönlichkeit abgeleitet, die mit den zu empfehlenden Gegenständen abgeglichen wird. Andere Verfahren versuchen abzuleiten, welche Emotionen ein bestimmter Gegenstand (wie ein Film oder ein Musikstück) hervorgerufen wird und fragen den Nutzer vor der Empfehlung nach seiner momentanen Gefühlslage. Ein solches System wird in (Hu und Pu, 2009a) beschrieben und Umsetzungen davon sind Whattorent und Yobo. Hu und Pu (2009b) zeigen durch Nutzerstudien weiterhin, dass die Akzeptanz der Nutzer und die Qualität der Ergebnisse vergleichbar mit den bestehenden kollaborativen Filtersystemen wie MovieLens sind. Mit Hilfe von persönlichkeitsbasierten Ansätzen könnte sich daher in Zukunft eine Alternative zu den kollaborativen Verfahren ergeben.

Vorteile:

- Demographische Fragen sind präziser erfassbar, als eine Vielzahl von Bewertungen für kollaborative Ansätze.

- Persönlichkeitsbasierte Verfahren versprechen eine gute Alternative oder Ergänzung zu den stark personalisierenden kollaborativen Verfahren zu werden.

Nachteile:

- Nutzerprofil bleibt nach der ersten Befragung meistens statisch bestehen.

4.4.4 Hybride Empfehlungssysteme

Wie bereits gesehen, hat jeder der drei Empfehlungsansätze seine Vor- und Nachteile. Zwar gibt es für jeden einzelnen Ansatz Möglichkeiten, die Nachteile zu verringern, allerdings kann man sie nicht wirklich verhindern. Aus diesem Grund geht man oft dazu über, verschiedene Empfehlungssysteme miteinander zu kombinieren, um so möglichst die Schwächen des einen Ansatzes durch die Stärken des anderen auszugleichen. Burke prägte 2002 erstmals den Begriff des hybriden Empfehlungssystems für eine derartige Kombination. Er gibt eine Systematisierung der möglichen Kombination von verschiedenen Empfehlungen (Burke, 2002) (Gunawardana und Meek, 2009). Tabelle 4.2 gibt eine Übersicht über die Kombinationsmöglichkeiten.

Sinnvolle Kombinationen sind Ansätze, die sich in der Startphase eines Empfehlungssystems auf Verfahren verlassen, die mit wenig Eingangsinformationen auskommen, wie beispielsweise inhaltsbasierte Verfahren, und nach den bekannten Anlaufschwierigkeiten auf stark personalisierende Verfahren umschwenken. Beispiele für derartige Kombinationen sind zum einen die vorliegende Arbeit und zum anderen Arbeiten wie Gunawardana und Meek (2009), die ebenfalls ein inhaltsbasiertes Verfahren mit einem kollaborativen Filter mit Hilfe einer Eigenschaftenkombination vereinen. Aber auch zu den anderen Kombinationsverfahren, wie die Gewichtung (Claypool u. a., 1999) oder der Vermittlung

gewichtet (weighted)	Die berechnete Punktezahl (bzw. Bewertungen) der einzelnen Empfehlungssysteme wird zu einer gemeinsamen Empfehlung kombiniert.
vermitteln (switching)	Das System vermittelt je nach Situation die Anfrage an ein zur Verfügung stehendes Empfehlungssystem.
gemischt (mixed)	Ergebnisse von verschiedenen Empfehlungssystemen werden gleichzeitig angezeigt.
eigenschaftskombiniert (feature combination)	Eigenschaften von verschiedenen Datenquellen werden für einen Empfehlungsalgorithmus zusammengeführt.
Kaskade (cascade)	Ein Empfehlungssystem verfeinert das Ergebnis, das ein anderes Empfehlungssystem zuvor berechnet hat.
eigenschaftenerweitert (feature augmentation)	Das Ergebnis des ersten Empfehlungssystems ist das Ergebnis des zweiten Empfehlungssystems.
Metaebene (meta-level)	Ein Model, das durch das erste Empfehlungssystem gelernt wurde, wird als Eingabe des zweiten Ansatzes verwendet.

Tabelle 4.2: Kombinationsmethoden nach Burke 2002

(Billsus und Pazzani, 2000), lassen sich ähnliche Arbeiten finden. Burke selbst verwendet für sein Entree Restaurant Empfehlungssystem ein wissensbasiertes Empfehlungssystem, kaskadiert mit einem kollaborativen Verfahren.

4.5 Bestehende Empfehlungssysteme für Bilder

Eines der wenigen bestehenden Beispiele für ein Bildempfehlungssystem für Bilder ist Photoree. Als Bildquelle dient dabei die Onlinegallery Flickr. Zusätzlich werden von dem Nutzer Informationen darüber gesammelt, welche Bilder der Benutzer gerne hat und welche nicht. In der Beschreibung des Services heißt es: „Photoree is to photos what Last.fm is to music and StumbleUpon to websites“. Technisch gesehen ist Photoree eine typische Implementierung eines benutzerbasierten Empfehlungssystems und gründet somit auf denselben Prinzipien, wie der in dieser Arbeit vorgestellte Algorithmus. Der Unterschied ist allerdings, dass Photoree kein geografisches Feedback verarbeitet, sondern nur die expliziten Bewertungen durch den Nutzer. Die Meinungen werden in einem simplen Ratingsystem ausgedrückt, das zwei wertende Zustände für ein Bild kennt (uprate, downrate). Aus den bereits abgegebenen Wertungen werden dann neue Bilder berechnet. Was dieses Beispiel allerdings verdeutlicht, ist, dass rein bewertungsabhängige kollaborative Filter meistens nicht sehr einsichtige Ergebnisse abgeben, wenn die Suche nicht weiter eingeschränkt werden kann. Zudem ist dieser Ansatz sehr zeitaufwendig für den Benutzer, da er zunächst ausreichend viele Bewertungen generieren muss. Im Fall von Photoree muss der Benutzer 100 Bilder bewerten, bevor für ihn eine persönliche Empfehlung ausgesprochen werden kann. Dies ist ein typisches Muster, um die Startprobleme eines kollaborativen Filters zu umgehen.

Andere Systeme, wie JustClick (Fan u. a., 2009), verwenden zwar den Begriff personalisierte Bildempfehlung, meinen allerdings inhaltsbasierte Verfahren, die Ähnlichkeiten der Bilder auf verschiedenen Eigenschaften des Bildes berechnen, um den Nutzer beim Durchsuchen der Bilder zu unterstützen.

4.6 Ortsbasierte Empfehlungssysteme

Geografisches Hintergrundwissen für ein Empfehlungssystem zu nutzen, ist nicht komplett neu. Allerdings wird es in den meisten Fällen ausschließlich als zusätzlicher Filter verwendet, um Ergebnisse aus klassischen Ansätzen in Bezug zu der momentanen Position zu setzen. Die meisten Arbeiten dazu lassen sich unter dem Stichwort „location-based recommendation“, also ortsbezogenen Empfehlungen, finden. Bae-Hee u. a. (2006) führt die Distanz von Objekten zur momentanen Position als zusätzlichen Kontext ein. Insgesamt arbeitet er mit drei Kontexten, die alle ihre eigene Rangliste für die Datenbasis liefern. Die gewichtete Summe aller Kontexte ergibt dann das endgültige Ergebnis. Horozov u. a. (2006) verwendet einen einfachen kreisförmigen Puffer, abhängig vom momentanen Standpunkt des Benutzers, um die Menge der Benutzer zu reduzieren, mit denen er dann eine benutzerbasierte Filterung nach GroupLens durchführt. Interessant an dieser Arbeit ist die Beobachtung, dass Menschen, die in der gleichen Nachbarschaft leben, auch ähnliche Restaurants in einer anderen Stadt besuchen. So scheint gleiches Verhalten an einem Ort auch das Verhalten an einem anderen zu bestimmen. Zumindest gilt dies bei Horozov für „Leben“ und „Essen“. In meinem Ansatz werden die geographischen Informationen nicht nur als zusätzliche Filterung der Daten verwendet, sondern als Eingabedaten für das Empfehlungssystem selbst.

Neuere Ansätze gründen auf Daten aus ortsbasierten Netzwerken wie Gowalla und können sich daher bereits auf ein sehr viel detailliertes Wissen über die zu empfehlenden Orte stützen (Berjani und Strufe, 2011). Die Idee dieser Netzwerke ist es, dass Nutzer an zuvor beschriebenen Orten einchecken können. Spots an einem bestimmten Ort sind in diesen Netzwerken demnach explizit bekannt, während der hier vorgestellte Ansatz ein Modell des Ortes aus den Daten selbst noch erzeugen muss.

4.7 Geographische Empfehlungssysteme

Die Arbeit, die erstmals den Begriff geographisches Empfehlungssystem einführt, verfasste Schlieder (2007). In dieser Arbeit wird die grundlegende Ideen für ein solches System diskutiert und dessen Nutzen aufgezeigt. In der Arbeit wird der Begriff geografisches Empfehlungssystem als ein System beschrieben, das aus einer Datenquelle von georeferenzierten Objekten eine Auswahl als Empfehlung zurückliefert. Zusätzlich wird eine Mehr-Objekt-Empfehlung definiert, die eine Auswahl von Objekten als sinnvoll ansieht, wenn sie von außen betrachtet möglichst homogen erscheint, paarweise gesehen allerdings möglichst heterogen ist. Die Kombination von beiden Ansätzen beschreibt ein System, das in der Lage ist, für eine Stadt mehrere interessante Plätze vorzuschlagen oder eine Diashow für ein bestimmtes Gebiet zu erstellen.

Definition 4.4 *Ein geographisches Mehr-Objekt-Empfehlungssystem schlägt eine Menge von georeferenzierten Objekten vor, die typisch für die gewählte Anfrage sind. Die einzelnen Empfehlungen wurden aber auch variabel genug gewählt, um eine breitgefächerte Antwort zu geben.*

4.8 Zusammenfassung

Im Forschungsbereich der Empfehlungssystemen lässt sich eine Vielzahl von verschiedensten Ansätzen finden, die alle ihre Vorteile, aber auch ihre Nachteile haben. Das gängigste Verfahren ist dabei sicherlich der nutzerbasierte kollaborative Empfehlungsansatz, da er auch die besten Ergebnisse verspricht. Für die Umsetzung für einen bestimmten Anwendungsfall lassen sich die jeweiligen Schwächen gezielt durch einen hybriden Ansatz reduzieren.

Im Sinne von Burke, ist der in dieser Arbeit vorgestellte Ansatz ein Meta-Level Empfehlungssystem, das einen gegenstandsbasierten mit einem kollaborativen Ansatz kombiniert. Zunächst wird ein geografisches Modell aus den Daten berechnet und dieses Modell für den nächsten Empfehlungsschritt verwendet. Für den ersten Empfehlungsschritt wird dabei ein geografisches Clusterverfahren benutzt, das im letzten Empfehlungsschritt zu Ergebnissen mit hoher geografischer Diversität führt. Zusätzlich minimiert die Verwendung eines Clusterverfahrens die Auswirkungen des New Item Problems, da neue Gegenstände meistens zu schon bestehenden Bildgruppen hinzugefügt werden, die die maßgebende Rolle für den kollaborativen Empfehlungsschritt spielen. Mein eigener Ansatz berechnet eine Mehr-Objekt-Empfehlung nach Definition 4.4 wobei die Anfrage an das System durch eine geografische Region bestimmt wird. So kann eine Anfrage für die Stadt Bamberg beispielsweise durch ein umschließendes Raumpolygon bestimmt werden, das der Stadtgrenze entspricht.

5 Typische Bildgruppen

Der eigene Lösungsansatz setzt sich aus zwei Teilschritten zusammen, die in den kommenden zwei Kapiteln beschrieben werden. Dieses Kapitel beschäftigt sich zunächst mit der grundlegenden Fragestellung, welche Konzeptualisierungen typisch für ein bestimmtes Gebiet sind. Weiterhin wird untersucht, ob sich eine typische Konzeptualisierung durch Bilder von einer Konzeptualisierung durch andere Medien, wie beispielsweise Text, unterscheidet. Bedingt durch die Darstellungsmöglichkeiten des gewählten Mediums, wird man verschiedenen Objekten oder Gebäuden einer Stadt in unterschiedlicher Häufigkeit beobachten können. Im Falle von Fotografien trifft man dabei auf festgehaltene Ansichten einer Stadt von verschiedenen geografischen Positionen. Eine Fotografie weist bereits durch die Art ihrer Entstehung einen starken Ortsbezug aus. Während man einen Text über ein bestimmtes Gebäude von jedem Ort der Welt aus schreiben kann, muss ein Fotograf sich an einen bestimmten Ort bewegen, bevor er die entsprechende Ansicht festhalten kann.

Alle Ansichten der Nutzergemeinschaft werden genutzt, um ein prototypisches Modell des Gebietes zu berechnen. Gewichtet wird dieses Modell nach dessen Typikalität, die mit Hilfe eines Popularitätsmaßes eines Ortes umgesetzt wird. Zunächst wird ein Anwendungsfall beschrieben, der die Suche nach den charakteristischen Aufnahmeorten in Bildersammlungen motiviert. Dieses Modells eines Ortes wird im nächsten Kapitel mit einem kollaborativen Filter verbunden. Damit

wird die Empfehlung von typischen Ansichten in eine Empfehlung von persönlichen Ansichten überführt.

5.1 Anwendungsfall: typische Bildgruppen einer bestimmten Region

Es gibt viele Situationen im Leben, in denen man gerade an den typischsten Entscheidungen von anderen Nutzern interessiert ist. Vor allem im touristischen Bereich, in dem meistens die Besuchszeit eines Ortes knapp bemessen ist, orientiert man sich gerne an den jeweiligen Highlights in einem bestimmten Gebiet. Touristische Highlights, sind Orte, die eine bestimmte Ausstrahlung aufzeigen und daher von den Reisenden auch vermehrt aufgesucht werden. Nicht selten sind solche Orte ein geschicktes Zusammenspiel von gezieltem Marketing und den positiven Rückmeldungen von Menschen, die bereits dort waren. Ein extremes Beispiel für diesen Schneeballeffekt stellt sicherlich die Legende um den Omelette Shop in Jodhpur (Indien) dar ¹⁶. Ein kurzer Artikel im Lonely Planet hat dabei dem kleinen Laden mit seinen saftigen Omeletts zu Weltruhm verholfen. Seither ist der kleine Laden auf zahlreichen Bildern von Touristen zu sehen und Ramkishan Gawlani, der Besitzer des Ladens, verbraucht am Tag 1000 Eier um die Nachfrage nach seinen Omletts zu stillen. Mittlerweile ist allerdings das einzige Alleinstellungsmerkmal des kleinen Ladens die Gewissheit der Kunden, das "originale" Omelette gegessen zu haben, das im Lonely Planet erwähnt wurde. Früher war dieser Laden der einzige, der überhaupt Omeletts verkaufte; aber mittlerweile gibt es überall im Stadtgebiet Nachahmer, die versuchen, von der Bekanntheit dieses Ortes mit zu profitieren. Allerdings zeigt sich auch hierbei, dass ein Tourist meis-

¹⁶ <http://www.sueddeutsche.de/reise/lonely-planet-in-indien-eierweise-ruhm-1.945601>

tens sehr gezielt seine Umgebung konsumiert und sich sehr stark durch die Meinung anderer beeinflussen lässt. Bei diesem Beispiel kommt verstärkend hinzu, dass man nach einem Omelett schon gesättigt ist und nicht auf jedes Angebot eingehen kann. Diese Beobachtung trifft natürlich nicht nur auf das Essen zu. Dieses sehr illustrative Beispiel zeigt zwei Eigenschaften von populären Orten:

- Popularität ist messbar. Im Falle eines Restaurants ist Popularität unter anderem an der Menge der zahlenden Kunden erkennbar. Im Falle von touristischen Orten werden wir im Folgenden versuchen über die Anzahl der Bildaufnahmen seine Popularität zu messen.
- Popularität zeigt eine zeitliche Dynamik. Der zweite große Aspekt, ist die zeitliche Veränderung von Orten und nicht zuletzt auch deren Umfeld. Die zeitliche Komponente wird in Zukunft verstärkt eine Rolle spielen, da dann raum-zeitlich verortete Bilder auch über größere zeitliche Distanzen zur Verfügung stehen werden. Dadurch lassen sich dann auch Veränderungen in der Popularität in einer Region erkennen, zum Beispiel die zunehmende oder abnehmende Popularität des kleinen Omelett-Geschäftes in Jodhpur.

Der erste Anwendungsfall stellt ein System dar, das in der Lage ist, die momentane Verteilung der Popularität um den eigenen Standpunkt herum zu berechnen. Dazu berechnet das System die Popularität von Orten anhand der Bildaufnahmen, die dort zu finden sind. Durch den Zeitstempel in den Metadaten der Bilder könnte diese Verteilung auf bestimmte Zeitabschnitte eingegrenzt werden. Von dieser Möglichkeit soll allerdings zunächst abgesehen werden. Berechnet wird daher nur die Popularität, die insgesamt von den Orten erzeugt wurde. Dieses Verfahren wird von einem ortsbezogenen Dienst für mobile Geräte dazu verwendet, die n populärsten Orte auszuwählen und dem Nutzer auf

einer Karte zu präsentieren. Damit behält der Nutzer dieses mobilen Endgerätes die Übersicht, welche bekannten Orte bzw. Landmarken in seiner Umgebung liegen. Das System kommt dabei ohne eine vorherige Modellierung des Ortes aus und beruht alleine auf den Erfahrungen von anderen Nutzern, die sie in Form von Bildern zurückgelassen haben. Informationsquellen für eine solche Anwendung sind schon heute mittels öffentlicher Bildmanagementsysteme in vielfältiger Art vorhanden. Eine kurze Beschreibung einiger dieser System findet man in Kapitel 3.

“Ich würde gerne 10 Bilder und deren Aufnahme sehen, die typisch für das Stadtgebiet von Bamberg sind“ ist eine kennzeichnende Fragestellung, die das System beantwortet.

In dem System treten folgende Akteure auf:

- **Aktiver Nutzer:** ist der Nutzer, der momentan eine Anfrage stellt
- **Fotograf:** ist die Gruppe von Benutzern, die georeferenzierte Bilder für das System zur Verfügung stellen. Das System, das die Bilder selbst verwaltet, kann getrennt von dem Empfehlungssystem vorliegen bzw. von einem anderen Service bereitgestellt werden.

Es werden folgende Anwendungsfälle von dem System verwirklicht:

- **Empfehlung anfragen:** Das ist die einzige Aktion, an der der Nutzer des Systems aktiv beteiligt ist. Der Nutzer legt für eine gültige Anfrage ein bestimmtes Gebiet fest, für das er eine Berechnung von Vorschlägen durchführen möchte und teilt diese Entscheidung dem System mit.
- **Bilder einstellen:** Benutzer stellen Bilder in eine Bildergalerie oder ein anderes Repository unter einer Nutzerkennung ein. Die Nutzerkennung ist für spätere Analyseschritte notwendig. Die Bilderbasis kann vom System getrennt sein, muss aber über eine öffentliche Schnittstelle zugänglich sein. Wichtig ist, dass die Bil-

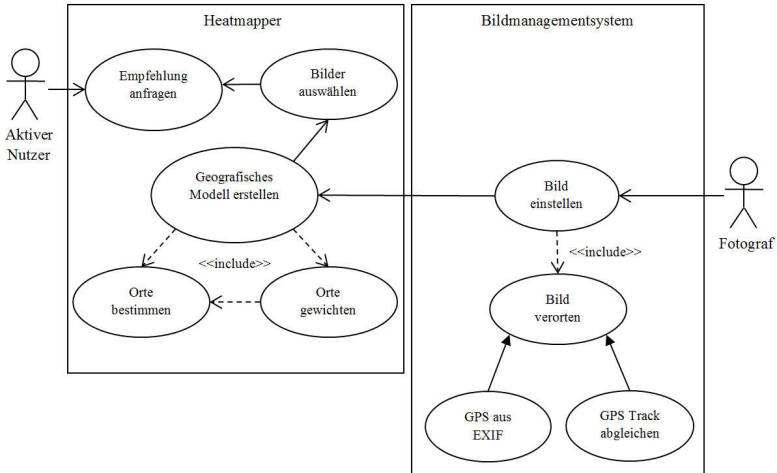


Abbildung 5.1: Anwendungsdiagramm für Bildempfehlungen für typische Bilder

der verortet vorliegen, daher schließt die Aktion “Bild einstellen“, die Aktion “Bild verorten“ mit ein.

- **Bild verorten:** Zusätzlich zu der primären Bildinformation gibt der Fotograf die Position seines Standpunktes an. Um ein Bild zu verorten gibt es verschiedene Möglichkeiten. Zwei davon sind die direkte Verortung durch den Fotoapparat und somit die Extraktion der GPS Informationen aus den EXIF Daten oder der nachträgliche Abgleich mit GPS Informationen aus externen Geräten, wie in Kapitel 2 beschrieben.
- **geografisches Modell erstellen:** Aus den räumlichen Metainformationen in den Bilddaten wird ein geografisches Modell des Gebietes erzeugt, das zu jedem Ort seine Popularität angibt. Diese Aktion schließt die folgenden Aktionen mit ein:

- Orte bestimmen: Zunächst muss das System festlegen, was als eigenständiger Ort gelten soll. Mit diesem Schritt beschäftigt sich Abschnitt 5.4.1.
- Orte gewichten: Als nächster Schritt werden die berechneten Orte nach dem verwendeten Popularitätsmaß gewichtet. Ein Popularitätsmaß wird in Abschnitt 5.4.2. entwickelt.
- Bilder auswählen: Auf der Grundlage des berechneten geographischen Modells kann nun eine Auswahl von Bildern für das gewählte Gebiet getroffen werden.

Abbildung 5.2 zeigt eine prototypische Umsetzung dieses Anwendungsfalls für ein mobiles Endgerät, das mit einem mobilen Browser über eine Webschnittstelle auf einen solchen Service Zugriff hat. Auf diesen Service kann über einen beliebigen Browser mit Hilfe von <http://www.kinf.wiai.uni-bamberg.de/tp5/m.heatmap.php> zugegriffen werden. Hilfreich bei der Umsetzung einer solchen Anwendung sind vor allem die technischen Möglichkeiten, die eine Positionsbestimmung von Endgeräten ermöglichen. Moderne Browser Plug-Ins bieten schon jetzt mittels verschiedenster Techniken (Zugriff auf GPS, WLAN-Triangulierung, . . .) die Möglichkeit, auf eine möglichst genaue Ortsbestimmung zuzugreifen. Bei der beschriebenen Anwendung wurde Google Gears verwendet. Zukünftige Browser werden die momentane Position des Nutzer über die neuen Möglichkeiten von HTML 5 abfragen können. Dadurch kann das System die momentane Umgebung des Gerätes als geografische Einschränkung für die Anfrage verwenden. Im November 2012 wurde die Entwicklung von Goolge Gears offiziell zugunsten von HTML 5 eingestellt, dass mittlerweile von allen modernen Browsern unterstützt wird. Aus diesem Grund ist die automatische Lokalisierung in der aktuellen Implementierung des Heatmappers nicht mehr einsatzfähig.



Abbildung 5.2: Mobiler Client für einen ortsbezogenen Dienst für typische Ansichten um den eigenen Standpunkt.

5.2 Bildgruppen in der Wirklichkeit

Bildgruppen, die typisch für ein bestimmtes Gebiet sind, findet man vor allem redaktionell erstellt in der Vermarktung von Städten über Poster, Flyer oder die offiziellen Internetauftritte. Hierbei geht es zum einen darum, auf einen Blick alle Highlights, zum anderen aber auch die vielseitigen Facetten der jeweiligen Stadt zu zeigen. Eines dieser Sammelwerke sind Patchwork Postkarten, die wir aus mehreren europäischen Städten untersucht haben. Dadurch konnten die typischsten Orte und deren Häufigkeiten bestimmt werden. Durch diesen Datensatz war auch ein Vergleich zwischen den automatisch generierten Bildgruppen und den real existierenden Bildgruppen möglich (siehe Kapitel 8). Jede der gesammelten Postkarten hatte die Auflage mehrere bildliche Eindrücke der Stadt zu zeigen. Zudem sollte sie den Eigenschaften einer Mehr-Objekt-Empfehlung entsprechen. Es gibt durchaus auch Postkarten, die verschiedene Detailansichten von einem einzigen Gebäude zeigen, diese waren für den gewünschten Datensatz allerdings weniger interessant.

Eine Bildgruppe kann als spezielle Ausprägung der zuvor eingeführten Tagging-Relation gesehen werden (Definition 2.2). Dabei besteht das Ausgangsobjekt nicht aus einem einzelnen Element, sondern aus einer Menge von Bildern. Diese Art der Verschlagwortung nach Schlieder (2007) Mehr-Objekt-Tagging genannt und ebenfalls durch eine dreistellige Relation angegeben. Dabei ist zu beachten, dass der erste Parameter in der Relation in diesem Fall eine Menge von Objekten darstellt

Definition 5.1 *Ein Mehr-Objekt-Tagging ist eine dreistellige Relation aus einer Menge von Objekten, dem dazugehörigen Schlagwort und dem Urheber dieser Relation: $\text{tagging}(\{\text{Objekt}_1, \dots, \text{Objekt}_N\}, \text{Schlagwort}, \text{Benutzer})$.*

In den Beispielen in Abbildung 5.3 finden sich die Städtenamen “Nürnberg” und “Bamberg” auf den Postkarten, während die Bilder eine facettenreiche Illustration dieser Orte bieten. Im Vergleich zur Definition nach Gruber (2005), die nur von einem Objekt pro Tag ausgeht, lässt sich argumentieren, dass ein Mehr-Objekt-Tagging viel ausdrucksstärker ist als das Einzel-Objekt-Tagging. Es lassen sich folgende semantische Eigenschaften mit Hilfe von mehreren Objekten ausdrücken:

- Man kann typische Instanzen eines gemeinsamen Konzepts darstellen. Typisch für das Konzept “Bamberg” sind zum Beispiel Ansichten von Gebäuden, wie dem “Alten Rathause” oder dem “Dom”.
- Ein Schlagwort kann mit Hilfe von mehreren Bildern mit einer anderen Konzeptualisierung verbunden werden. So kann das Schlagwort “Bamberg” mit Bildern des Weltkulturerbes illustriert werden oder aber mit Bildern, die Bamberg als fränkische Bierhauptstadt darstellen. Die semantische Aussage eines Schlagwortes hängt demnach von der Auswahl der Objekte für das Mehr-Objekt-Tagging ab.
- Neben den typischen Ansichten lassen sich auch die Variationen der einzelnen Teilobjekte besser darstellen. So sind sechs Bilder des Eiffelturms zwar alle für sich gesehen typische Ansichten für Paris, zeigen aber nicht die Vielfalt von Paris.

Schaut man diese drei Eigenschaften genauer an, lässt sich leicht erkennen, dass es sich um eine Umsetzung von Typikalität und Variabilität in der Domäne von Bildgruppen handelt. Die Typikalität wird dabei durch den Bezug zu einem geografischen Gebiet bestimmt, sowie durch Konzepte wie “Weltkulturerbe” oder “Bierstadt”. Die Variabilität durch die Vermeidung von Ansichten auf dasselbe Objekt.

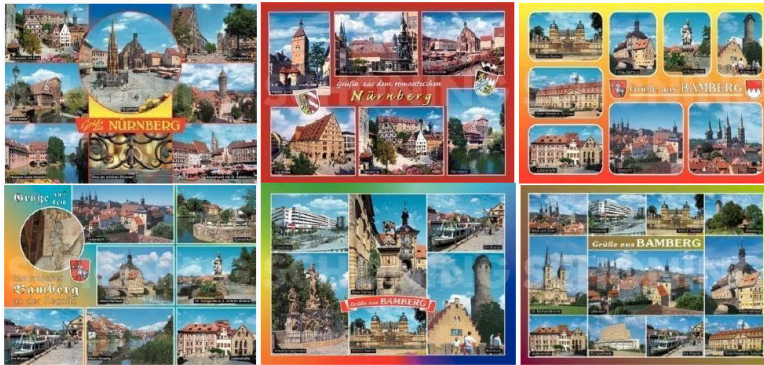


Abbildung 5.3: Postkarten aus verschiedenen Städten (Schoening)

Multi-Objekt-Tagging wird momentan von den Bildgalerien so gut wie nicht unterstützt. Was sich finden lässt, sind Ordnerstrukturen, die von den Nutzern meistens dazu genutzt werden, ihre Bilder nach Orten oder Events zu unterteilen und von automatischen Verfahren auch dabei unterstützt werden können (Naaman u. a. (2004), Seite 54). Im späteren Verlauf meiner Arbeit wird zusätzliches geografisches Hintergrundwissen in Form einer Partonomie dazu genutzt, die Bilder automatisch einem Ort zuzuordnen und dadurch automatisch eine Mehr-Objekt-Tagging Relationen mit dem Ort als Schlagwort zu erzeugen. Dadurch werden die unterschiedlichen Konzeptualisierungen des geografischen Ortes der einzelnen Nutzer sichtbar und ermöglichen es, auf deren Grundlage, ihre Ähnlichkeiten in diesem Ort zu bestimmen.

Tabelle 5.1 gibt eine Übersicht über die Gebiete für die Postkarten gesammelt wurden. Auf den insgesamt 72 Postkarten sind 466 Bilder zu sehen, was einen Durchschnitt von 6,5 Bildern pro Postkarte ergibt. Pro Stadt sind dabei im Durchschnitt 20 verschiedene Sehenswürdigkeiten zu sehen, von denen dann jeweils eine Untermenge für die Postkarte ausgewählt wurde.

	Warwick	Cardiff	Barcelona	Paris	Potsdam	Edinburgh	Bamberg	Freiburg	Insgesamt
Bilder	29	33	38	71	100	33	104	57	466
Sehens- würdig- keiten	22	16	21	20	26	14	24	18	161
Postkarten	7	5	6	8	12	7	15	12	72

Tabelle 5.1: Übersicht über die gesammelten Postkarten

Im Fall von Bamberg zeigen die Postkarten Gebäude der historischen Altstadt, die als UNESCO Weltkulturerbe bekannt ist. Interessant ist, dass diese Auszeichnung als Weltkulturerbe nicht auf Grund eines einzelnen Gebäudes vergeben wurde, sondern als Ansammlung von mehreren Gebäuden von gleichem historischem Interesse, aufgeteilt in die drei historisch gewachsenen Regionen (Bergstadt, Inselstadt und Gärtnerstadt). Wären alle Gebäude gleichbedeutend (Anforderung 1) würde man auf den Postkarten eine Gleichverteilung der abgebildeten Gebäude erkennen können und man könnte eine zufällige Auswahl dieser Abbildungen wählen, um eine typische Sicht auf die Stadt zu erhalten. Allerdings findet man einige wenige Gebäude auf fast allen Postkarten und es muss zumindest eine Vorstellung davon vorhanden sein, was wirklich typische Ansichten sind.

Tabelle 5.2 zeigt die häufigsten Eindrücke der Stadt Bamberg. Zusätzlich zu dieser Beobachtung lassen sich einige Gebäude nur auf ganz wenigen Postkarten finden, was für den Gedanken der Variabilität spricht (Anforderung 3). Eine vollkommen andere Konzeptualisierung (Anforderung 2) des Ortes lässt sich auf den Postkarten leider nicht finden und alle Postkarten haben paarweise gesehen eine hohe Übereinstim-

mung bei der Auswahl der Bilder. Auf keinem der Bilder ist eine Brauerei oder ein Bierkeller abgebildet, was eher dem Konzept der Bierstadt Bamberg gerecht werden würde. Auch vollkommen andere Stile wurden bei der Bildauswahl nicht beachtet. So wäre es durchaus interessant, eine Postkarte nur mit Bildern zu gestalten, die mit der sowjetischen Schnappschusskamera Lomo entstanden sind.

Abbildung 5.4 gibt an, wie sich die insgesamt 24 abgebildeten Gebäude auf die Postkarten von Bamberg verteilen. Die Abbildung liest sich wie folgt: 12 der Gebäude ließen sich nur auf ein bis drei Postkarten finden, sechs Gebäude auf vier bis sechs Postkarten und so weiter. Nur ein einziges Gebäude war auf fast allen Postkarten abgebildet: das “alte Rathaus“. Die Kurve, die durch diese Häufigkeiten aufgezeigt wird, kommt einer Potenzreihe sehr nahe und gibt einen ersten Hinweis, dass eine Auswahl von Bildern, um ein geografisches Konzept zu illustrieren, auf der einen Seite sehr typische Eindrücke enthält, auf der anderen Seite aber auch eine bestimmte Variabilität aufweist. Die Verteilungen auf den Postkarten für die restlichen Städte ergeben ein ähnliches Bild.

Rang	1	2	3	4	5
Name	Altes Rathaus	Dom	Villa Concordia	Klein Venedig	Michaelsberg
Bild					
Anzahl	14	11	9	9	8

Tabelle 5.2: Verteilung der Sehenswürdigkeiten in Bamberg

Um die Anzahl der Bilder zu bestimmen, die dasselbe Objekt in einer Stadt zeigen, wurden von Hand gleichbedeutende Abbildungen auf den verschiedenen Postkarten gruppiert und benannt. Diese Vorgehensweise

ist bei größeren Bildgruppen allerdings nicht mehr möglich und es muss im Folgenden ein automatisches Verfahren gefunden werden. Bei dieser Gruppierung soll vor allem der Ortsbezug der Bilder eine Rolle spielen. Weiterhin muss diskutiert werden, ob die reine Anzahl an Bilder als Popularitätsmaß ausreichend ist. Die übergeordnete Frage ist dabei, ob man aus unstrukturierten Bildsammlungen durch einen kollaborativen Ansatz zusätzliche Semantik erzeugen kann.

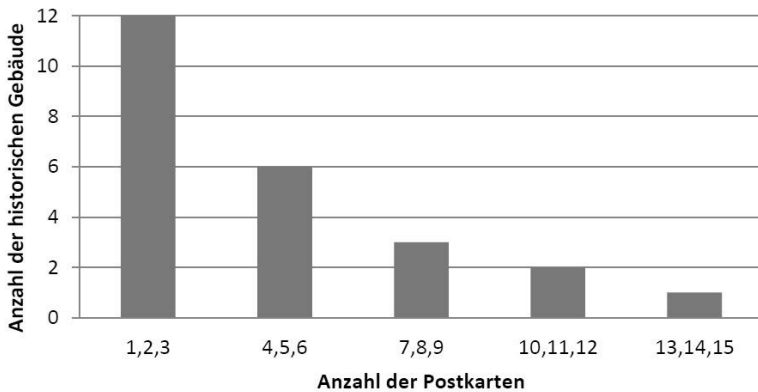


Abbildung 5.4: Typische Verteilung der verschiedenen Abbildungen auf Postkarten

5.3 Kollaborative Semantik in unstrukturierten Bildsammlungen

Bildsammlungen, die man in öffentlichen Onlinegalerien findet, basieren meistens auf dem Grundsatz des sozialen Netzes, das heißt eine Grundordnung ist nicht vorgegeben, sondern entsteht durch die Nutzer des Systems selbst. Dies steht im Gegensatz zu den Techniken aus dem semantischen Netz, bei dem versucht wird, Dateninformationen

sehr strukturiert anzubieten, zum Beispiel mit Hilfe einer vorgegebenen Ontologie. Im sozialen Netz kann der Nutzer sein Wissen über ein Objekt in Form von Metadaten frei zum Ausdruck bringen. Dies geschieht mit Hilfe des Taggings als freie Indexierung der Inhalte. Das Ergebnis dieser freien Indexierung durch den Nutzer wird als Folksonomy bezeichnet eine Taxonomy, die durch die Nutzergemeinschaft („Folk“, englisch für Volk) entsteht. Diesem Ansatz gegenüber steht die Wissensmodellierung durch eine Ontologie. Hierbei muss die zu speichernen Metainformationen in ein festgelegte Modell eingeordnet werden. Es gibt Ansätze, den einen Mechanismus zur Wissensrepräsentation in den jeweils anderen übersetzen. Für viele Ansätze ist eine ontologische Modellierung zu strikt oder eine Folksonomy zu frei und unstrukturiert [(Christiaens, 2006), (Al-Khalifa und Davis, 2007), (Rattenbury und Naaman, 2009), (Hagemann und Vossen, 2009)].

Im Fall von georeferenzierten Bildsammlungen trifft man vor allem im Privatbereich auf Folksonomies, da diese dem Benutzer keine feste Begrifflichkeit vorschreiben und daher für ihn einfacher zu handhaben sind. Nur bei sehr professionellen Bildagenturen findet man hauptsächlich Indizierungstechniken, die man eher dem Konzept des semantischen Netzes zuordnen würde. Diese Arbeit beruht hauptsächlich auf Datensätzen, die in öffentlichen Web2.0 Galerien zu finden sind und somit ebenfalls mit Folksonomies arbeiten. Es stellt sich daher die Frage, welche semantischen Informationen man aus diesen kollaborativen Datensätzen gewinnen kann. Ein kollaborativer Beitrag ist dabei zum einen Text in Form von Tags und zum anderen geografische Positionen in Form von Geotags. Widmen wir uns zunächst der Analyse der Geotags.

Aus der Sicht der semantischen Analyse ist es interessant, ob es Stellen im geografischen Raum gibt, die häufiger mit dem Konzept einer Stadt in Verbindung gebracht werden, ähnlich der Häufigkeit der Bil-

der auf den gesammelten Postkarten aus dem letzten Abschnitt. Um diese Frage auf Grundlage von unstrukturierten Bildsammlungen zu untersuchen, wurde im Rahmen dieser Arbeit eine Applikation, mit dem Namen Heatmapper, entwickelt. Diese Anwendung ermöglicht es, die angesprochenen Bildquellen auf diese Fragestellung hin zu untersuchen. Untersucht werden dabei die Informationen aus den Geotags der Bilder, das heißt die Analyse beruht ausschließlich auf den geografischen Koordinaten und der Nutzerkennung, unter der die Bilder veröffentlicht wurden.

Um den Heatmapper an real existierenden Datensätzen zu testen, wurde er im Februar 2008 auf Bilder von der Webseite Panoramio angewandt (Schlieder und Matyas, 2009). Die Bilder befanden sich in einem rechteckigen Gebiet um die Stadtmitte von den vier europäischen Städten Amsterdam (NED), Bamberg (GER), Cardiff (GBR) und Dublin (IRL), die sich schon grundlegend in der Einwohnerzahl unterscheiden: Amsterdam 750.000, Dublin 510.000, Cardiff 320.000 und Bamberg 70.000. Der komplette Datensatz umfasst 12.066 Bilder, die von 3.369 verschiedenen Nutzern veröffentlicht wurden. Für jede Stadt wurde die Anzahl der Bilder und die Anzahl der verschiedenen Benutzer bestimmt. Zusätzlich wurden zwei verschiedene statistische Durchschnittsmaße (Mittelmaß und Median) über die Anzahl der Bilder, die ein Benutzer für ein Gebiet eingestellt hat (Bilder pro Benutzer - bpb), bestimmt, siehe Tabelle 5.3.

Der Median liegt relativ niedrig bei nur zwei Bildern in den kleineren Städten und bei drei Bildern in der Stadt Amsterdam. Das bedeutet, dass über die Hälfte der Nutzer nur zwei Bilder für das System bereitstellen, was an den Eigenarten, der Plattform liegen kann. Nach einem genaueren Blick auf die Datenbasis selbst, dokumentieren die Benutzer von Panoramio gern Reisen in mehrere Städte, wobei sie jeweils nur ein Bild als eindeutige Markierung für eine Stadt verwenden. Ein Nutzer

	Bilder	Benutzer	BPB	Median BPB
Amsterdam	7460	2172	3,44	3
Bamberg	1035	141	7,34 (3,61)	2
Cardiff	597	176	3,39	2
Dublin	2974	880	3,38	2

Tabelle 5.3: Datensatz und Mittelmaße der Bilder pro Benutzer

hat zum Beispiel jeweils ein einzelnes Bild für Amsterdam, London und New York hochgeladen. Auffallend war die hohe durchschnittliche Anzahl Bilder pro Benutzer (bpb) in Bamberg. Sie war mehr als doppelt so hoch wie bei allen anderen Gebieten. Grund war ein einzelner Nutzer, der alleine über 500 Bilder veröffentlicht hat. Höchstwahrscheinlich handelt es sich dabei um einen Einwohner der Stadt Bamberg, da die Anzahl seiner Aufnahmen über die Zeit zunahm. Im Juni 2011 hatte dieser Nutzer 4529 Bilder für Bamberg in Panoramio veröffentlicht. Lässt man diesen Benutzer außen vor, ergibt sich ein durchschnittlicher bpb von 3,61, was zwar immer noch der höchste Wert von allen vier Städten, aber vergleichbar mit dem bpb in anderen Städten ist.

Ein Ergebnis dieser empirischen Studie ist die Erkenntnis, dass die Entscheidung ein Bild in eine öffentliche Bildsammlung zu laden nicht davon abhängt, ob eine ähnliche Ansicht desselben Objekts schon vorhanden ist. Wäre dies der Fall, dann würde der Raum gleichmäßig mit Bildern gefüllt sein und wir könnten nicht beobachten, dass man in solchen kollaborativen Datensätzen über 1.500 Bilder vom Spire of Dublin oder, um ein extremeres Beispiel zu nennen, 25.000 Bilder des Eiffelturms findet. Es kann davon ausgegangen werden, dass bereits jetzt eine 25.001ste Ansicht des Eiffelturms auf Flickr zu sehen sein wird. Im Umkehrschluss lässt sich daraus ableiten, dass jeder Nutzer für sich gesehen seine Umwelt nach seinen Vorlieben dokumentiert und diese

Erfahrung mit anderen über seine eingestellten Bilder mitteilt. Nun wird es im nächsten Schritt darum gehen, die räumlichen Entscheidungen zu verstehen, die während der Entstehung eines Bildes getroffen wurden.

5.4 Räumliche Entscheidungen bei georeferenzierten Bildern

Wir verwenden den Begriff fotografische Sicht für eine Untergruppe von fotografischen Werken, die ganz bewusst von einem Ort aus erstellt worden sind, um ein bestimmtes nicht-bewegliches Objekt zu fotografieren. Gemeint sind daher vor allem Gebäude- oder Landschaftsaufnahmen und weniger künstlich inszenierte Studioaufnahmen. Studioaufnahmen schaffen durch ihre Inszenierung einen völlig neuen Ortsbezug oder einen gewollten Eindruck von Ortlosigkeit, wie es zum Beispiel bei Passbildern der Fall ist. Der eigentliche Entstehungsort ist bei diesen Aufnahmen unwichtig. Ähnliche Definitionen finden sich in der Anthropologie und geografie unter dem Begriff des Nicht-Ortes [(Auge, 1995), (Relph, 1976)], wozu zum Beispiel Orte des modernen Transportnetzes, wie Flughäfen oder Bahnhöfe zählen. Das sind Orte, die sich ortsunabhängig für denjenigen, der sie betritt, gleich anfühlen und keinen individuellen Charakter mehr aufweisen. Allerdings sind fotografische Ansichten von Nicht-Orten dennoch interessant, da sich die Bilder ähnlich sehen, aber durch den Ortsbezug eine andere Information liefern. So ist es durchaus wertvoll zu wissen, ob es sich bei dem gezeigten Flughafen um New York oder Amsterdam handelt.

Jede fotografische Sicht beinhaltet zwei getrennte räumliche Entscheidungen, die der Fotograf treffen muss, um diese auf einem Bild festzuhalten. Die erste Entscheidung ist sein eigener Standpunkt. Die nächste räumliche Entscheidung ist die Richtung, in der er die Kamera ausrich-

tet. Die Richtung wird meistens von einem konkretem Objekt definiert, das in einer Richtung von der momentanen Position liegt. Daher kann man die beiden räumlichen Entscheidungen auch mit Hilfe von zwei geografischen Positionen angeben (siehe auch Abbildung 5.5).

Definition 5.2 *Der Point of View (POV) ist die geografische Position, an der sich die Kamera befindet, während das Bild erstellt wird. In den allermeisten Fällen ist dies auch der Standpunkt des Fotografen. Auf diesem Bild ist daher seine momentane Sicht von dieser Position aus (bzw. ein Ausschnitt davon) zu sehen.*

Definition 5.3 *Der Point of Interest (POI) in einem Bild ist die geografische Position des Objekts, auf dem der Hauptfokus im Bild liegt, was nicht das Objekt in der Bildmitte sein muss.*

Zwei fotografische Sichten sind dann äquivalent, wenn diese beiden räumlichen Entscheidungen sich sehr ähnlich sind.

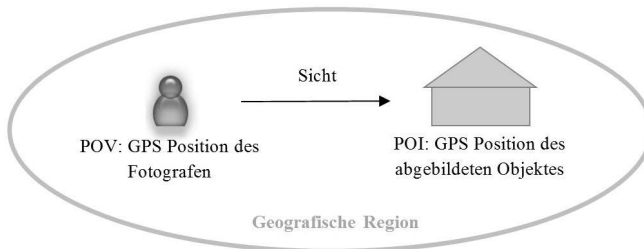


Abbildung 5.5: Räumliche Entscheidungen

Moderne digitale Bildformate bieten die Möglichkeit, Informationen über das Bild (Größe, Pixelseitenverhältnis, Orientierung, ...), aber auch den Erstellungsvorgang (Kameratyp, Hersteller, eingestellte Brennweite, eingestellte Blende, Datum, Uhrzeit, ...), zu speichern.

Diese Informationen über die eigentlichen Primärdaten (dem Bild) werden Metadaten oder Metainformationen genannt. In den meisten Bildformaten werden Metadaten direkt in das Bild selbst eingebunden. Dafür gibt es im Moment zwei wichtige Standards, den Exchangeable Image File Format (EXIF) von Japan Electronic and Information Technology Industries Association (JEITA) und den Extensible Metadata Platform (XMP) von Adobe.

XMP verwendet dabei eine XML Syntax und ist ein offener und erweiterbarer Standard für die Metadaten eines Bildes. Der Kern des Standards kann daher um bestehende Standards erweitert werden. So lässt sich zum Beispiel der komplette EXIF Standard in XMP einbetten (man erkennt die EXIF-Tags an dem Präfix „exif“). Zusätzlich bietet XMP die Möglichkeit, alle Metadaten in eine getrennte Datei, dem sogenannten „sidecar file“, abzulegen und damit das Bild von seinen beschreibenden Informationen zu trennen. Im Gegensatz dazu sind die EXIF Informationen immer direkt in die Datei eingebettet und zudem noch auf JPG und TIFF Dateien beschränkt. Insgesamt lässt sich daher sagen, das XMP flexibler, einfacher einzusetzen und zeitgemäßer ist.

Mit dem Aufkommen von Lokalisierungsverfahren für den Fotografiebereich wurden beiden Formate um spezielle Datenfelder für räumliche Informationen ergänzt. In beiden Formaten kann seither nur der Standort des Fotografen mit in die Kamera annotiert werden. Um eventuell das fotografierte Objekt aus den Metadaten eines Bildes zu bestimmen kann, falls vorhanden, die Blickrichtung ausgewertet werden. Lösungsvorschläge, Objekte im Bild zu identifizieren, werden in Kapitel 7 kurz vorgestellt.

Sowohl die Entscheidung des POV als auch des POI werden nach persönlichen Vorlieben getroffen, daher man kann sich für den Vergleich des Nutzerverhaltens auf einen der beiden festlegen. Die Information,

die letzten Endes in den Metadaten zu finden ist, entspricht dem POV, den wir zuvor definiert haben. Technisch gesehen ist also im Moment nur der Vergleich mit Hilfe des POVs möglich. Ein spielerischer Ansatz, der beide Positionsangaben erzeugt, wird im GeoGame „CityExplorer“ (Matyas u. a., 2008) verwendet. Hierbei wird vom Nutzer verlangt, jeweils zwei Positionen für eine Aufnahme anzugeben: den Standpunkt, der automatisch beim Erstellen des Bildes protokolliert wird, und eine Position, die möglichst nahe am Objekt ist. Auch die Onlinegalerie Locr.de hat diese logische Trennung schon vollzogen, in dem sie dem Nutzer ermöglicht neben seinem Standpunkt einen weiteren Punkt auf einer Karte zu markieren, den sie selbst als „Objektstandort“ bezeichnen: „Der Objektstandort zeigt den Blickwinkel des Fotografen und die Position des Motivs, das du fotografiert hast“. Abbildung 5.6 zeigt ein Beispiel dieses Features auf Locr.de.

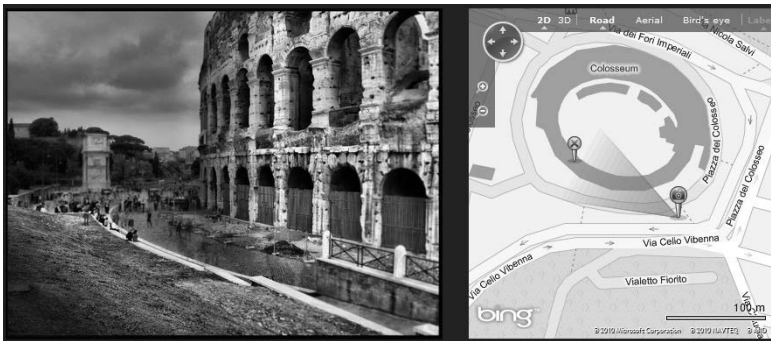


Abbildung 5.6: Trennung von POI und POV auf Locr.de

Aufgrund der besseren Datenbasis in den vorhandenen Bildsammlungen werden die folgenden Überlegungen auf der Basis des Fotografenstandpunktes oder auch POV geführt. Dieselben Metriken und Algorithmen können auch mit Hilfe der gewählten POIs berechnet werden.

Beide räumliche Entscheidungen (POV und POI) sind für die folgenden Überlegungen gleichwertig.

5.4.1 Gruppierung von räumlichen Entscheidung

Im Folgenden wird es notwendig sein, eine räumliche Entscheidung nicht als einen einzelnen räumlichen Punkt zu sehen, sondern als eine Gruppierung von einzelnen räumlichen Entscheidungen. Eine räumliche Entscheidung besitzt daher eine gewisse räumliche Ausdehnung und ist mehr eine Region als ein einzelner Punkt. Es ist also sinnvoll, anstatt von einzelnen POVs von einem Cluster von POVs (CPV) zu sprechen. Diese räumliche Ausdehnung hat die folgenden zwei Gründe:

- *GPS-Ungenauigkeit:* Es kann nicht davon ausgegangen werden, dass der gleiche Standpunkt des Fotografen auch auf die gleiche GPS-Position abgebildet wird. Bei guten Bedingungen weist das GPS-Signal eine Genauigkeit von ungefähr 15 Metern auf. Erst mit Differential-GPS kann man eine höhere Genauigkeit erreicht werden. Bei den heutigen Geräten, die für die Fotografie eingesetzt werden, kann allerdings nicht immer von der genauesten GPS-Lösung ausgegangen werden. Die Hersteller von GPS-Kameras halten sich eher bedeckt mit Aussagen über den verbauten GPS-Sensor in den Kameras. Auch bei spezieller Geotagging Hardware ist das nicht immer der Fall. So bietet Sony keine genaue Spezifikation des GPS Chipsatzes für seinen GPS CS1 Datenlogger. Vergleicht man dieses Gerät zum Beispiel mit einem Antaris4 GPS Chipsatz (Wintec WBT-300) kann man die deutlich schlechtere Empfangsqualität erkennen und somit einen deutlichen Versatz zum eigentlichen Straßenverlauf. Abbildung 5.7 zeigt einen Vergleich zwischen diesen beiden GPS-Loggern, die für den Privatbereich als spezielle GeoTagger angeboten werden.

Beide Geräte wurden von derselben Person getragen und auf einem Weg auf offenem Feld getestet. Es zeigen sich Unterschiede bis zu 30 Meter auf, im Durchschnitt um die 10 Meter, wobei der Antaris4 Chipsatz fast exakt auf dem Weg verläuft.

- *Mehrere Blickwinkel:* In Zukunft kann man zwar davon ausgehen, dass die GPS-Geräte durch neue Technologien immer genauer werden, dennoch macht es Sinn, den Nutzerstandpunkt vielmehr als eine Region zu betrachten anstatt als einen genauen geographischen Punkt. Selbst wenn der Benutzer sich nicht genau an derselben Position befindet, wird ihm in den meisten Fällen dennoch ein sehr ähnlicher Blick auf dasselbe Objekt ermöglicht. Bei einer leichten Veränderung der Position verändern sich demnach auch die möglichen Ansichten nur leicht. Erst wenn man zwei weit entfernte Standpunkte einnimmt, sind jeweils andere Objekte sichtbar.



Abbildung 5.7: Vergleich von zwei Bewegungsspuren, die mit zwei Datenloggern aufgenommen wurden, Sony-GPS CS1 in Weiß und Wintec WBT-300 in Rot

Ein CPV sollte daher die möglichen GPS-Fehler bzw. die unterschiedliche Genauigkeit der GPS-Geräte auffangen und zudem noch möglichst die Regionen ergeben, von denen bevorzugt dasselbe Objekt von mehreren Blickwinkeln fotografiert wird. Verwendet wird daher ein geografisches Clusterverfahren, das diese CPVs aus den räumlichen Anhäufungen von gewählten Fotografenstandpunkten berechnet. Es wird

angenommen, dass dort, wo vermehrt Fotoaufnahmen erstellt werden, eine geringe Anzahl (im Idealfall eins) von nichtbeweglichen Objekten für diese Häufung verantwortlich ist. Der verwendete Clusteralgorithmus sollte folgende Anforderungen aufweisen:

- Berechnung einer dynamischen Anzahl von Clustern. Die Anzahl k der Cluster ist vorher unbekannt und es kann daher nicht auf Standardverfahren wie K-Means zurückgegriffen werden.
- Die maximale Ausdehnung der Cluster sollte beschränkt sein. Würde ein Cluster zu groß werden, würde es gegen Anforderung 2 (nur leichte Varianz in der Position) verstoßen und auch POVs mit aufnehmen, die grundsätzlich verschiedene Ansichten ermöglichen.

Als eine mögliche Lösung dieser Bedingung wurde ein Clusteralgorithmus adaptiert, der als Variante des k-Means Ansatzes gesehen werden kann. Allerdings ohne die Bedingung, sich zuvor auf die Anzahl der Cluster festzulegen. Ashbrook und Starner (2002) verwendeten einen ähnlichen Algorithmus, um unstrukturierte GPS-Tracks nach persönlichen Orten zu untersuchen, an denen sich eine Person mit Vorliebe aufgehalten hat. In beiden Fällen wird ein Suchradius verwendet, dessen Größe zuvor festgelegt und während der Clusterung nicht verändert wird. Der Algorithmus zur Gruppierung von POVs zu einer Menge von CPVs lässt sich wie folgt beschreiben:

- Wähle einen POV aus der Menge aller POVs. Dieser Punkt kann zufällig gewählt werden und dient als Startpunkt für den Clusteralgorithmus. Die geografische Position des POVs gibt den Mittelpunkt des Suchradius an.
- Finde alle POVs, die in dem momentanen Radius liegen. Alle POVs werden zu einer temporären Ergebnismenge CPV_{temp} hinzugefügt.

- Berechne den Mittelpunkt des Suchradius mit Hilfe der POVs in CPV_{temp} und wiederhole Punkt 2 mit dem neuen Mittelpunkt so lange, bis sich die Menge der POVs nicht mehr ändert. Dadurch wandert der Suchradius in ein lokales Maximum.
- CPV_{temp} wird in die Ergebnismenge aufgenommen und alle POVs aus der Menge der POVs entfernt. Danach beginnt der Algorithmus ab Punkt 1 mit den restlichen POVs wieder von vorne.

Um eine schnellere Clusterung, gerade bei vielen Eingangsdaten, zu erreichen, wurde folgende Vereinfachung verwendet, die ohne die innere Schleife zwischen Punkt 2 und Punkt 3 auskommt. Bei dieser Variante wird jeder Punkt nur ein einziges Mal in Betracht gezogen, was eine lineare Laufzeit ermöglicht. Benötigt wird dazu eine Liste, in der alle vorläufigen CPVs (CPV_{pre}) abgelegt werden. Diese Liste ist zu Beginn noch leer und ergibt nach Ablauf des Algorithmus die Ergebnismenge mit den berechneten CPVs.

- Kontrolliere jeden *POV*, ob er in einem Suchradius eines CPVs in CPV_{pre} liegt.
 - Falls kein CPV gefunden werden konnte, wird ein neuer *CPV* erstellt und in die Liste CPV_{pre} mit aufgenommen. Der Mittelpunkt des Suchradius entspricht dabei der Position des POVs.
 - Falls ein *CPV* gefunden worden ist, wird der *POV* zu ihm hinzugefügt und der Mittelpunkt wird mit Hilfe von allen räumlichen Entscheidungen in dem Cluster neu berechnet.

Abbildung 5.8 zeigt das Ergebnis dieses Algorithmus, angewandt auf 1204 Bilder aus dem Stadtgebiet von Bremen, in dem 555 Cluster mit einem verwendeten Suchradius von 80 Meter gefunden werden konnten. Insgesamt waren 334 Benutzer an diesem Datensatz beteiligt. Der größte Cluster, der am Rathaus von Bremen liegt, enthält 140 Bilder von

70 verschiedenen Nutzern, während sich am hinteren Ende der Verteilung 298 Cluster anreihen, die nur ein Bild von einem einzigen Nutzer enthalten.

Zusätzlich zu der räumlichen Clusterung bietet die Heatmapper Anwendung auch die Möglichkeit, die textlichen Informationen der Bilder in einem Cluster nach Häufigkeit zu sortieren. In dem Fall von Panoramio wurden die verwendeten Tags und die Begriffe in den Titeln in Form einer Tagcloud dargestellt, um dem Benutzer eine Vorstellung über die prominentesten Ansichten in einem Cluster zu vermitteln. Im Fall von Bremen lassen sich dadurch recht schnell die fünf Sehenswürdigkeiten erfassen, die in den ersten fünf CPVs gezeigt werden: Rathaus, Dom, Böttcherstrasse, Schnoor und die Windmühle am Wall.

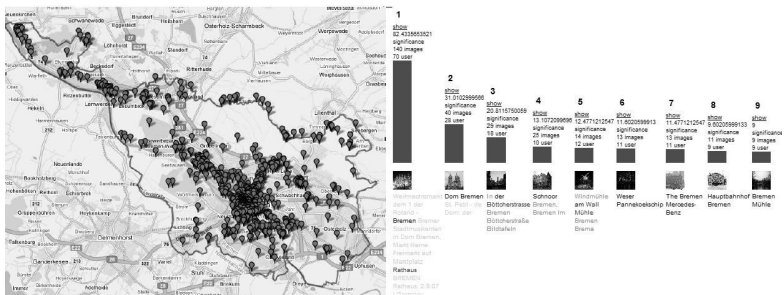


Abbildung 5.8: Heatmap von Bremen

Die Nachteile, Bildempfehlungen auf Grundlage der CPVs zu berechnen, lassen sich unter anderem in Abbildung 5.9 erkennen. Generell kann man sagen: wählt man den POS für eine Empfehlung, können die POIs (also die abgebildeten Objekte auf den Bildern) doppeldeutig werden. Diesen Effekt hat man aber auch, wenn man sich nur auf die Popularität des POI stützt. Die populärste Ansicht kann man daher nur berechnen, wenn man den populärsten POI vom populärsten Standpunkt aus fotografiert.

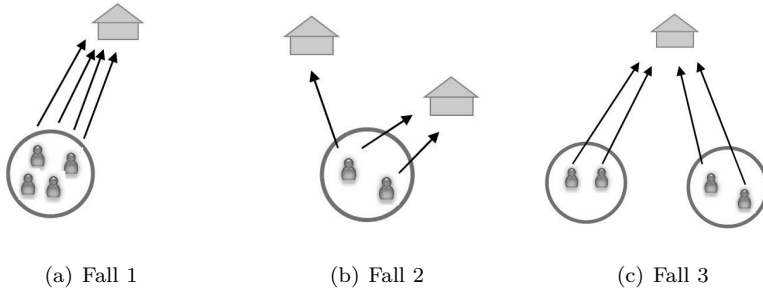


Abbildung 5.9: GPS-Position als Indikator für kongruente räumliche Entscheidung

Wünschenswert wäre der Idealfall, dass alle Bilder von einem Ort aus dasselbe geografische Objekt zeigen; der POV bestimmt also eindeutig den POI (Abbildung 5.9, Fall 1). Im Normalfall werden jedoch mindestens zwei Sichten von einem Ort aus aufzufinden sein, wobei wir von einer prominenten Richtung ausgehen können. Im Fall der Bilder aus Bamberg wurde von allen berechneten CPVs, durchschnittlich auf 76% aller Bilder, jeweils ein bestimmtes Objekt illustriert und nur 24% der Bilder in einem CPV weisen andere Motive auf (Abbildung 5.9, Fall 2). Ein Beispiel für die analysierten Bildgruppen stellt Abbildung 5.10 dar.

In der Theorie sollte man vor allem bei sehr exponierten Gebäuden eine Vielzahl von POVs feststellen können. Diese Thematik in Abbildung 5.9(c) illustriert. Einige Objekte werden von vielen verschiedenen Stellen aus sichtbar sein oder von weit entfernten Stellen. Ein Beispiel wäre eine Burganlage, die stark exponiert auf einer Bergspitze sitzt. In der Theorie bedeutet es, dass diese Anlage von zu vielen verschiedenen Stellen fotografiert wird, um alle Standorte mit einem einfachen Clusteralgorithmus greifen zu können.

Nach einer genauen visuellen Untersuchung der Ergebnisse aus Bamberg wurde jedoch deutlich, dass gerade Fall 3 nicht sehr oft in den untersuchten Bildsammlungen auftritt. Die Altenburg zum Beispiel findet man zwar in weit entfernten CPVs. Allerdings findet man bestimmte weit entfernte Ansichten auch in gemeinsamen Clustern wieder, die eine besonders schöne Ansicht auf die Burg und den Bergkamm ermöglichen. Außerdem lässt sich feststellen, dass deutlich mehr Bilder zu finden sind, die direkt in der Burg gemacht wurden. Die meisten Benutzer scheinen viel eher zu dokumentieren, was sie in ihrer nächsten Umgebung gesehen haben. In der gesamten Stadt gibt es hauptsächlich vier Stellen, an denen weitläufige Blicke in Bilder dokumentiert sind. Und diese Blicke werden immer von den gleichen Stellen aufgenommen. Dass selbst große Gebäude nur von einer begrenzten Anzahl von Blickwinkeln fotografiert werden, hat auch architektonische Gründe. Gebäude bieten dem Betrachter meistens diejenige Fassade an, die auch am attraktivsten erscheint. Dies wird in den meisten Fällen, die Eingangsseite sein, durch die man das Gebäude üblicherweise betritt und die somit meistens einen repräsentativen Charakter aufweist. Räumliche Relationen, wie vorne oder hinten, lassen sich so auch auf Gebäude übertragen und vorne wird meistens auf die Gebäudeseite angewendet, an der der Eingang liegt. Denkt man zum Beispiel an große religiöse Bauten, wie die Kathedrale Notre Dame oder den Kölner Dom, lässt sich schnell eine einzige Position feststellen, von der aus bevorzugt fotografiert wird. Der räumliche Kontext eines Gebäudes unterstützt ebenfalls bestimmte Sichtachsen. So befindet sich zum Beispiel „vor“ den Kirchen meistens ein großer Platz, von dem man die Kirche in ihrer Gesamtheit sehen kann, während sie ansonsten von dicht stehenden Gebäuden umrahmt sind. Auch bei anderen repräsentativen Bauten kann man diese inszenierten Sichtachsen sehen, die bevorzugt von Fotografen aufgegriffen werden.

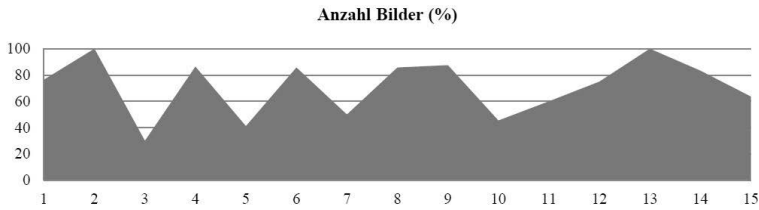
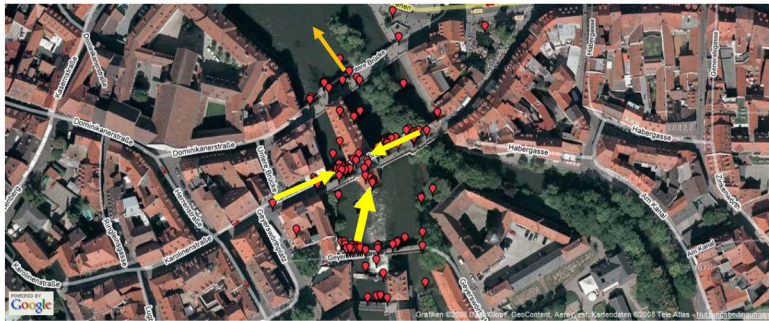


Abbildung 5.10: Anzahl der Bilder, die in den größten Gruppen von Bildern dasselbe Objekt zeigen.

Abbildung 5.11 zeigt ein Cluster, das visuell untersucht worden ist. Die gelben Pfeile geben dabei die hauptsächlichen Richtungsvektoren in den Bildern an. Bestimmt werden die Bilder in diesem Cluster eindeutig von dem Gebäude in seiner Mitte, dem Alten Rathaus. Viele der Bilder zeigen das Rathaus von einer benachbarten Brücke aus. Von dieser aus eröffnet der Fluss Regnitz einen freien Blick auf den freihängenden Südflügel des Gebäudes, wobei links und rechts die Brücken zu den Flussufern zu sehen sind. Diese Ansicht ist eine der bekanntesten und lässt sich schon auf historischen Aufnahmen finden. Erst am nördlichen Ende des Clusters finden sich Aufnahmen, die vom Rathaus weg, ebenfalls über den Fluss hinweg einen guten Blick auf die Häuserfront am Fluss, die man Klein Venedig nennt, abbilden.

5.4.2 Popularität eines CPV

Von nun an betrachtet man nicht länger die einzelnen räumlichen Entscheidungen eines Nutzers im Sinne eines POVs, sondern nur noch die berechneten Bildgruppen durch das räumliche Clusterverfahren, das im letzten Abschnitt beschrieben wurde. Definitionsgemäß besteht ein CPV aus mindestens einer räumlichen Entscheidung, die innerhalb des kreisförmigen Suchradius liegt und damit mindestens ein Bild enthält.



Altes Rathaus (78/111)



Klein Venedig (7/111)



Blick zum Michaelsberg (3/111)



Kreuzigungsgruppe (4/111)



Abbildung 5.11: Analyse der Ansichten in einem CPV. Bilder stammen aus dem populärsten Cluster in Bamberg.

Wie schon bei der Untersuchung der Postkarten zu erkennen war, beruht eine Auswahl von Bildern auf einer Vorstellung, wie populär eine gewisse Sicht ist. Um diesem Gedanken entgegenzukommen, berechnet der Heatmapper ein Maß, das die Popularität eines Ortes bestimmt. Intuitiv würde man davon ausgehen, dass die Anzahl der Bilder ein guter Kandidat für diese Einschätzung sein kann. Bei mehreren Versuchen, sind allerdings Fälle von Nutzern aufgekommen, die für denselben Standpunkt extrem viele Bilder (> 20) eingestellt haben. Der Grund war in allen Fällen ein lokales Ereignis, das den Nutzer ermutigt hat, zum einen viele Bilder zu machen und zum anderen seine Position nicht weit genug zu ändern, um in einem anderen CPV mit aufgenommen zu werden. Ein Beispiel sind Open-Air-Veranstaltungen, wie Konzerte. Events sind leider keine Orte in unserem Sinne und haben nur einen zeitlich-begrenzten Bezug zu dem Ort, an dem sie stattgefunden haben. Ein Event findet zwar an einem bestimmten Ort statt, beinhaltet aber meistens eine inszenierte Funktion, die nicht mit der eigentlichen Funktion des Ortes übereinstimmt. Diese Inszenierung macht es schwierig, Bilder vom gleichen Ort am Tag des Events mit Bildern zu vergleichen, die außerhalb des Events stattgefunden haben. Verwendet man demnach nur die Anzahl der Bilder als einziges Maß, um die Popularität eines Ortes zu bestimmen, würde man bestimmte CPVs überbewerten. Wählt man hingegen die Anzahl der Nutzer, würde viel Information einfach verloren gehen und die reine Anzahl der Benutzer ist zudem nicht aussagekräftig genug. Es lassen sich viele CPVs finden, die zwar die gleiche Anzahl von Personen anzieht, sich aber in der Anzahl der aufgenommen Bilder stark unterscheiden. Zum einen weisen manche Gebäude zum Beispiel wegen ihren Details in der Fassade schon eine höhere Attraktivität auf. Zum anderen bieten manche Stellen auch mehr als einen interessanten Blick auf unterschiedliche Gebäude oder Gegenstände an, was die Popularität des Ortes ebenfalls erhöht. Ein sinnvolles Popularitätsmaß sollte daher diese Unterschiede in der Aus-

strahlung eines Ortes zusätzlich einfangen können. Es muss also ein Kompromiss zwischen der reinen Anzahl der Benutzer und der reinen Anzahl der Bilder geben. Die Idee bei diesem Kompromiss ist es, dass der Anteil, den ein weiteres Bild desselben Nutzers zur Popularität eines CPVs liefert, immer geringer wird. Somit hat das erste Bild, das jemand zu einem CPV hinzufügt, den höchsten Beitrag, während weitere Bilder einen immer kleineren Beitrag liefern. Die Gewichtung der Beträge hängt also von der schon vorhandenen Bilderanzahl eines Nutzers ab.

C sei nun die Menge der CPVs, die durch den Clusteralgorithmus berechnet worden ist. Die Menge C besteht aus einzelnen CPVs, die Bilder von k verschiedenen Nutzern enthalten. Jeder dieser Nutzer hat eine unterschiedliche Anzahl von Bildern beigetragen $n(u)$, jedoch mindestens ein Bild; daher gilt $n(u) > 0$. Eine mögliche Definition der Popularität eines CPV lässt sich folgendermaßen berechnen:

$$p(c) = \sum_{u \in U_n} (1 + \log n(u)) \quad (5.1)$$

Mit Hilfe dieses Maßes erhöht das erste Bild eines Nutzers die Popularität des Clusters um den Faktor 1. Das zweite Bild eines Nutzers hat dann bereits einen sehr viel geringeren Beitrag von $\log(2) \approx 0.3$ und der n -te Beitrag von $\log(\frac{n}{n-1})$. Mit anderen Worten: Ein CPV, in dem 2 Bilder von zwei verschiedenen Nutzern vorhanden sind, hat eine höhere Popularität als ein CPV, in dem beide Bilder von einem einzigen Nutzer eingestellt wurden. Im zweiten Fall würde sich eine Popularität von $1 + \log(2)$ ergeben und im ersten Fall von $2 > 1 + \log(2)$. Die Berechnung von $p(c)$ basiert auf dem Zehnerlogarithmus $\log_{10}$.

In Tabelle 5.4 wurden die CPVs aus Bamberg nach dem Popularitätsmaß $p(c)$ sortiert. Was sich gut erkennen lässt, ist die unterschiedliche Sortierung zwischen der Anzahl der Nutzer und der Popularität. An

CPV	Bilder	Benutzer	Popularität	Beispielbild
Altes Rathaus	111	59	68,5	
Klein Venedig	19	18	18,3	
Dom	22	14	16,0	

Tabelle 5.4: die drei größten Cluster of Points of View (CPV) in Bamberg

den ersten drei CPVs von Bamberg lässt sich zwar nicht feststellen, dass dies auch für die Anzahl der Benutzer und dem Popularitätsmaß gilt, aber wir werden im Folgenden Beispiele sehen, bei denen das Popularitätsmaß andere Ergebnisse liefert als die reine Sortierung nach der Anzahl der Benutzer. Für Bamberg erhält man zum Beispiel ab dem Rang 5 eine andere Sortierung. Weitere Beispiele für andere Städte werden im nächsten Abschnitt ähnliche Ergebnisse aufzeigen.

5.4.3 Die Heatmap als geografisches Modell einer Region

Durch die Gruppierung der verorteten Bilder mit Hilfe eines geografischen Clusteralgorithmus und der anschließenden Gewichtung durch das Popularitätsmaß wurde aus den zuvor unstrukturierten Daten ein geografisches Modell erzeugt. Die Methode wurde auf vier verschiedene europäische Städte angewendet. Nachfolgend werden die Ergebnisse für Amsterdam, Bamberg, Dublin und Cardiff diskutiert.

Für alle vier Städte wurde der Verlauf der Popularität in Relation zum zugehörigen Rang der räumlichen Entscheidung als Kurve illustriert. Diese Popularitätskurve in den Diagrammen 5.13 – 5.16 ist in blauer Farbe eingezeichnet. Die Diagramme beinhalten zusätzlich eine Kurve, die mit Hilfe der kleinsten Fehlerquadrate an die Messwerte angenähert wurde. Diese Annäherungskurve ist in roter Farbe hervorgehoben. Das Fehlermaß R^2 liegt bei allen fünf Städten bei über 0,94, was eine gute Annäherung der beiden Kurven bedeutet. Es liegt beinahe eine perfekte Korrelation vor. Diese Beobachtung bestätigt uns in der Annahme, dass die Rang-Popularitätskurve den Gesetzmäßigkeiten einer Potenzreihe folgt. Es ist ebenfalls erstaunlich, dass auch bei räumlichen Entscheidungen die gleichen Beobachtungen zu sehen sind wie bei anderen Benutzer-generierten Daten. Die Angleichung durch Potenzreihen lässt sich zum Beispiel auch bei der Häufigkeit von Begriffen in Folksonomies feststellen. Wertet man zum Beispiel die Tags der Bilder aus der Region Bamberg aus, ist es nicht erstaunlich, dass Wörter wie „Bamberg“ oder „Rathaus“ viel häufiger vorkommen als andere. Mehr dazu in Abschnitt 5.5. Die einfache Ableitung auf andere Datensätze, wie der bildlichen Repräsentation von räumlichen Entscheidungen, war allerdings nicht selbstverständlich und daher überraschend.

Betrachtet man die Popularitäts-Rang-Diagramme etwas genauer, fällt auch hier das Muster des Langen Schwanzes (engl. Long Tail nach Anderson (2006)) auf, das schon in der Einleitung erwähnt wurde. Laut Theorie müssten sich im Kopf der Verteilung Ansichten befinden, die den Bestsellern aus den e-Commerce-Plattformen entsprechen und somit Empfehlungen von typischen Eindrücken möglich machen. In der Tat befinden sich im oberen Bereich der Verteilung hauptsächlich die bekanntesten Sehenswürdigkeiten der jeweiligen Stadt. Dies ist zwar nicht sehr überraschend, zeigt aber dennoch, dass eine einfache räumliche Heuristik als Clusterverfahren ausreicht, um qualitativ gute Er-

gebnisse zu erzielen. Komplexere Offline-Verfahren sind vorerst nicht notwendig. Bilder aus den oberen CPVs zeigen prototypische Ansichten der Städte. Für die folgenden Beispiele wurde jeweils ein Bild aus den oberen fünf CPVs für Amsterdam, Bamberg, Cardiff und Dublin ausgewählt und diese zeigen Ansichten, die von Ortskundigen sofort erkannt werden und als klischeehafte Abbildungen gewertet werden. Aus einem anderen Blickwinkel könnte man diese Abbildungen auch als semantischen Kern für das Konzept dieser Städte betrachten. Semantisch gesehen definiert der Heatmapper das Konzept einer Region, die durch ein einfaches Polygon begrenzt wird. Das Konzept dieses Raumes wird dabei durch eine Auswahl von Bildern instanziiert.

Es folgen nun die Ergebnisse der Analyse für die vier europäischen Städte. Für jede Stadt ist zunächst das vorgestellte Diagramm zu sehen, das die Verteilung der Popularität der unterschiedlichen Orte in der Stadt zeigt, Bilder der fünf populärsten CPVs und eine kurze Diskussion über die Ergebnisse und Beobachtungen, die aus dieser Analyse gewonnen werden konnten. Für die Benennung der CPVs wurde analysiert, welches Toponym am häufigsten in den Schlagwörtern des Clusters vorkam. Da sich an der Webseite Panoramio hauptsächlich lokale Nutzer beteiligen, reflektierten die Tags meistens die jeweilige Landessprache. Die Toponyme wurden, soweit es Sinn macht, ins Deutsche übersetzt.

Betrachtet man die Anzahl der Bilder, die von den einzelnen Benutzern eingestellt wurden, ist es interessant zu beobachten, dass wie bei den meisten kollaborativen Datensätzen eine ziemlich deutliche 80–20 Verteilung festzustellen ist. Das bedeutet, dass 80% der Bilder, die in dem Datensatz zu finden sind, von 20% der Nutzer stammen. Abbildung 5.12 illustriert diese Verteilung mit Hilfe einer Lorenzkurve, in der die Anzahl der Bilder mit steigender Anzahl der betrachteten Nutzer (jeweils in Prozent) angegeben werden. Dieses Verhalten ließ sich in allen vier Städten finden. Es lässt sich auch beobachten, dass Nutzer,

die nur wenige Bilder einstellen, zumindest Ansichten von einem der bedeutenden Gebäude auswählen.

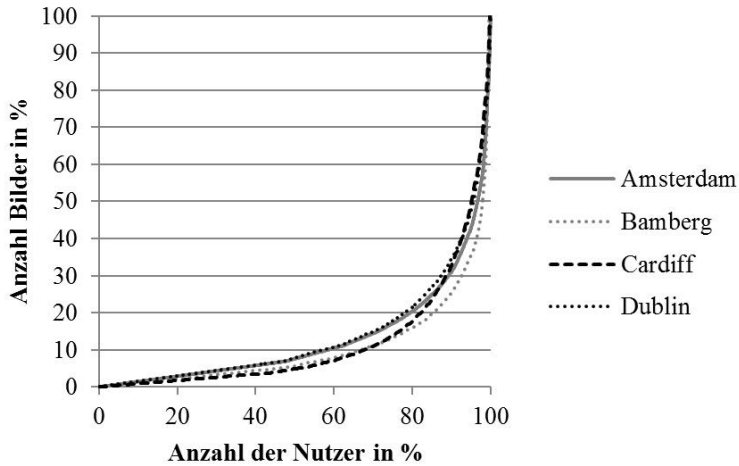


Abbildung 5.12: Lorenzkurve für die Mitarbeit der Nutzer in den vier Testgebieten

Amsterdam

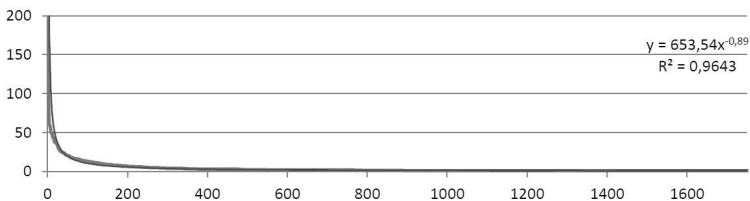


Abbildung 5.13: Verteilung der Popularität in Amsterdam

Am Beispiel von Amsterdam lässt sich deutlich erkennen, dass die verwendete Methode fähig ist, die populärste Stelle in der Stadt zu finden.

An erster Stelle befindet sich der Dam, der zentrale Ausgangspunkt für Besichtigungstouren durch die Stadt. Neben der Offenheit des Platzes selbst, der viel Raum für gute Aufnahmen bietet, befindet sich dort auch der königliche Palast als eines der Hauptmotive, die von dort aus aufgenommen werden. Im zweiten CPV lässt sich gut der zweite Problemfall, bei der Analyse der POVs illustrieren: Die Bilder, die dort aufgenommen werden, teilen sich fast gleichmäßig auf die zwei benachbarten Gebäude auf. Zum einen findet sich dort das Anne-Frank Haus, zum anderen das Westerwerk, eine bekannte protestantische Kirche in Amsterdam. Der dritte CPV zeigt die hohe Diversität der Ergebnisse. The Grashopper ist eine feste Größe im Nachtleben von Amsterdam, zentral in der Nähe von Dam und Hauptbahnhof gelegen scheint es mit der Kombination aus Restaurant, Sportsbar und Coffeeshop viele Nutzer der Webseite Panoramio anzusprechen. Hier zeigt sich auch die Stärke des Heatmappers, wirklich die jeweiligen Präferenzen der Nutzergemeinschaft zu reflektieren. Die Plätze vier und fünf ergeben sich wieder aus historischen Plätzen und Gebäuden in Amsterdam.

Bamberg

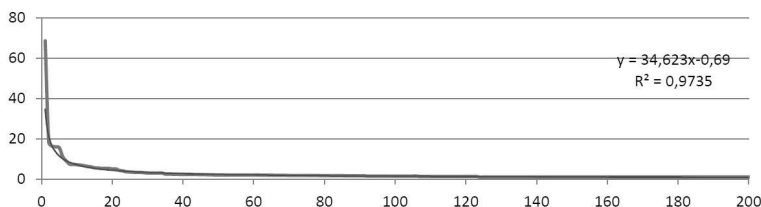


Abbildung 5.14: Verteilung der Popularität in Bamberg

Für Bamberg ergeben sich die erwarteten stereotypischen Ansichten der Stadt. Überraschend ist es allerdings, selbst weitenfernte Orte wie die Altenburg unter den ersten fünf Rängen zu finden. Die Altenburg

CPV	Bilder	Benutzer	Popularität	Beispielbild
Dam	240	189	200	
Anne Frank Haus / Westerwerk	77	50	60,4	
The Grashopper	67	55	58,6	
Waag	67	54	58,2	
Muntplein	62	64	51,4	

Tabelle 5.5: Die fünf größten clusters of points of view (CPV) in Amsterdam

liegt gute vier Kilometer und mehrere Höhenmeter von der Innenstadt entfernt und wird von Tagestouristen im Normalfall nicht besucht.

Cardiff

Cardiff illustriert sehr gut den Problemfall drei, da eigentlich angenommen werden kann, dass Cardiff Castle der bedeutendste Ort in der Stadt ist. Aber hier lässt sich auch der Unterschied zwischen dem POV und dem POI deutlich machen. Das eigentliche Gebäude, das auf den meisten Bildern als das Cardiff Castle ausgeschrieben wird, wird von einer hohen Burgmauer umzäunt (siehe Bild). Es ist also von der Straße aus

CPV	Bilder	Benutzer	Popularität	Beispielbild
Rathaus	111	59	68.5	
Klein Venedig	19	18	18.3	
Dom	22	14	16.0	
Altenburg	14	10	10.9	
Rosengarten	12	8	8.9	

Tabelle 5.6: Die fünf größten clusters of points of view (CPV) in Bamberg

nicht sichtbar. Um die Mauer zu überwinden muss man die Öffnungszeiten der Anlage einhalten und zusätzlich den Eintritt bezahlen. Somit sind die eigentlich interessanten Bildstandpunkte zeitlich und finanziell beschränkt. Der CPV, der in der Rangliste auf Platz 3 erscheint, enthält hauptsächlich Standpunkte vor der Burgmauer, wahrscheinlich von Personen, die an einer der beiden Hürden gescheitert sind. Innerhalb der Stadtmauern lässt sich durch den Heatmapper ein weiterer CPV finden, der die typische Ansicht auf die kleine normannische Burg (The Norman Keep) zeigt und sich auf Platz sechs befindet. Einen weiteren typischen Teil der Burg, das viktorianische Herrenhaus mit seinem herausragenden Glockenturm, findet man auf Platz acht. Die meisten

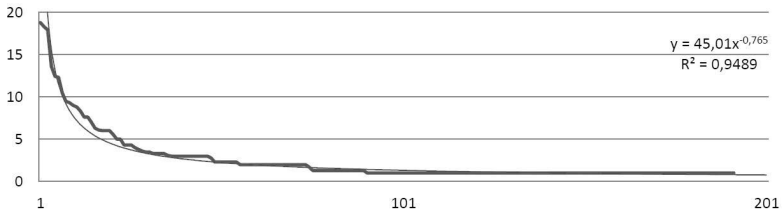


Abbildung 5.15: Verteilung der Popularität in Cardiff

Aufnahmen aus den restlichen hochplatzierten CPVs zeigen Gebäude der neugestalteten Hafengegend, die das kulturelle Zentrum der Stadt bilden. Auf Platz eins findet man das Millenium Stadion, eines der größten Stadien von ganz Großbritannien, gefolgt vom Pierhead Building, dem Wahrzeichen der Stadt und auch als walisischer Big Ben bekannt. Das Millenium Centre, der Sitz der walisischen Nationaloper, nimmt Platz vier ein.

Dublin

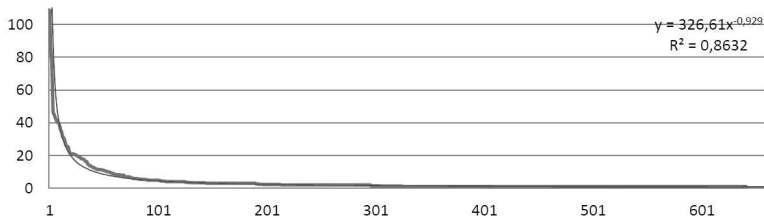


Abbildung 5.16: Verteilung der Popularität in Dublin

Wikipedia schreibt über die „Half Penny Bridge“, dass es die am meisten fotografierte Sehenswürdigkeit von Dublin ist. Nach der Analyse durch den Heatmapper findet man die „Half Penny Bridge“ zwar auf Platz drei. Der beliebteste Standpunkt, um ein Bild in Dublin zu ma-

CPV	Bilder	Benutzer	Popularität	Beispielbild
Millenium Stadium	34	16	18.8	
Pierhead Building	24	15	17.6	
Cardiff Castle	21	15	16.5	
Millenium Centre	24	13	15.4	
Norwegische Kirche	15	13	13.6	

Tabelle 5.7: Die fünf größten clusters of points of view (CPV) in Cardiff

chen, ist für die Panoramio-Nutzer auf bzw. an der O’Connel Brücke, die einer der belebtesten Strassen von Dublin über den Fluß führt. Auf derselben Straße befindet sich seit Dezember 2002 auch das neue Wahrzeichen von Dublin mit dem einfachen Namen “The Spire“. Mit der Temple Bar und dem Trinity College schließt die Top 5 aus Dublin mit zwei touristischen Highlights ab.

5.5 Kollaborative Semantik in Texten

Unter dem Begriff kollaborativer Semantik wird die Erstellung von zusätzlichem Wissen auf Grundlage von gemeinschaftlich erstellten Textsammlungen verstanden. Es stellt sich daher die Frage, inwieweit sich

CPV	Bilder	Benutzer	Popularität	Beispielbild
O'Connell Bridge	131	92	100.5	
Dublin Spire	121	78	86.7	
Half Penny Bridge	93	69	74.6	
Temple Bar	59	51	53.1	
Trinity College	55	44	47.2	

Tabelle 5.8: Die fünf größten clusters of points of view (CPV) in Dublin

die Ergebnisse aus der Analyse der geografischen Konzeptualisierungen mit Hilfe der Ortsangaben in Bildern im Gegensatz zu der Konzeptualisierung durch Texte unterscheiden. Dafür wurden verschiedene Textquellen untersucht.

- Zunächst wurde die Häufigkeit von Ortsbegriffen völlig unabhängig von Bildern analysiert. Dabei war insbesondere interessant, wie oft sie in Zusammenhang mit einem der vier Städtenamen (Amsterdam, Bamberg, Cardiff und Dublin) zu finden sind. Diese Art der Untersuchung wird auch Kookkurrenzanalyse genannt.
- Zusätzlich zu den geografischen Informationen in den untersuchten Bildern weisen die meisten von ihnen mehrere Tags als zusätzliche textliche Informationen auf. Diese bilden eine Folkso-

nomy für die Ansichten, die in den Bildern der entsprechenden Regionen zu finden sind. Interessant waren hierbei die Häufigkeiten der verschiedenen Begriffe. Die Fragestellung ist dabei, ob die verschiedenen Orte innerhalb einer Stadt ähnlich oft in Texten erwähnt werden, wie sie Fotografen, die sich im geografischen Raum bewegt haben, tatsächlich fotografieren. Für Bamberg würde man zum Beispiel Tags wie „AltesRathaus“ oder „Dom“ unter den häufigsten Schlagwörtern vermuten.

- Beim dritten Ansatz, eine Konzeptualisierung in Texten zu finden, wurde ausgehend von einem Textkorpus nach Assoziationsregeln gesucht. Im Gegensatz zur Kookurrenzanalyse haben die Beziehungen zwischen zwei Begriffen eine Richtung. Sind zwei Begriffe $B1$ und $B2$ über eine Assoziationsregel verbunden, dann kann man dies mit einer Kante $B1 \rightarrow B2$ in einem Graph darstellen. Diese Kanten weisen ein Gewicht auf, das die Häufigkeit der Fundstellen von $B1$ angibt in denen mit einer bestimmten Signifikanz auch über $B2$ gesprochen wird. Wir werden dazu am Ende des Kapitels noch ein anschauliches Beispiel sehen. Der untersuchte Korpus setzte sich aus Texten zusammen, die sich mit dem Ort „Bamberg“ beschäftigen. Als Quelle dienten dabei frei verfügbare Onlinetexte.

Die Ergebnisse der drei Ansätze werden im Folgenden mit den Ergebnissen der Bildanalyse verglichen und diskutiert. Es ergeben sich dabei folgende Hypothesen: Tags ergeben eine andere Priorisierung (gemessen an der Häufigkeit ihres Auftretens), als die Analyse der verorteten Bilder ergeben haben. Dies kann mehrere Gründe haben. (1) Zum einen kann man sagen, dass manche Orte es zwar ermöglichen, interessante Fotos zu machen, zum Beispiel, weil ein beeindruckendes Gebäude zu sehen ist. Aber die Namen der Dinge bzw. Orte, die fotografiert werden, müssen nicht zwangsläufig jedem geläufig sein. Daher lassen sich

Texte, die diese Orte auch bei ihrem Namen nennen, nur sehr spärlich finden. (2) Zum anderen lassen sich manche Orte nur schwer fotografieren (schwer zugänglich, Eintrittskosten, ...). Eventuell sind sie auch visuell nicht herausragend genug, um fotografiert zu werden. Nichtsdestotrotz haben sie einen sehr bekannten Namen. Daher trifft man auf mehr Textstellen, in denen diese Orte erwähnt sind, als Bilder, in denen sie abgebildet werden. (3) Vor allem Tags haben die Eigenschaft, dass sie meist mehrsprachig sind, aus zusammengesetzten Wörtern bestehen oder Rechtschreibfehler haben, was es zusätzlich erschwert, ein gemeinsames zu beschreibendes Toponym eindeutig zu identifizieren. Zwar lassen sich einige Methoden finden, die diese Eigenart versuchen zu umgehen, allerdings können sie nicht immer eine eindeutige Zuordnung finden. Beispiele sind das Stemmingverfahren, bei dem man Wörter auf einen Wortstamm reduziert, um leichter Synonyme zu finden, oder auch die Verwendung von Synsets: Wortgruppen, in denen neben Synonymen auch die Zuordnung zu einem übergeordneten Konzept zu finden sind. Zusätzlich werden auch Beziehungen zu anderen Wortgruppen angegeben.

5.5.1 Untersuchung von Konzeptualisierung einer Region in Texten

Um festzustellen, ob Text eventuell eine andere Konzeptualisierung der einzelnen Regionen zulässt, wurde die Häufigkeit der Begriffe, die diese Regionen beschreiben, untersucht. Wie schon bei den Bildern wurde daher von einer übergeordneten Region ausgegangen, die wiederum durch die vier Städte festgelegt waren. Diese übergeordnete Region wird in den Texten nicht mehr durch geografische Koordinaten bestimmt, sondern durch den Städtenamen definiert. Der Begriff „Bamberg“ entspricht daher dem Polygon, das das Gebiet der Stadt umschließt. Begriffe, die mit dem Städtenamen in irgendeiner Form in Beziehung stehen, entsprechen in der geografischen Untersuchung den einzelnen Fotogra-

fenstandpunkten, die wir zuvor auf bestimmte Orte in der jeweiligen Stadt zurückgeführt hatten. Dadurch sollen zwei Hypothesen untersucht werden:

- *Hypothese 1:* Die Häufigkeit von Textstellen, in denen der Stadtname zusammen mit einer Ortsbezeichnung auftritt, ergibt eine deutlich andere Häufigkeitsverteilung, als das bei Bildern aus dem Gebiet dieser Stadt der Fall ist.
- *Hypothese 2:* Es gibt Begriffe, die viel häufiger in Zusammenhang mit dem Städtenamen auftreten, als es über Bilder und ihrer Position nachvollziehbar ist.

Um beide Hypothesen zu testen, wurden als Maß für die Popularität eines Textes die Anzahl der Google Treffer für jedes Begriffspaar aus den Tabellen 5.5–5.8 bestimmt. Dabei wurden sowohl die Treffer der einzelnen Begriffe aufgenommen, als auch die Häufigkeit, mit der beide Begriffe zusammen vorkommen. Die folgenden Abbildungen verdeutlichen die Häufigkeitsverteilungen. Die Listen sind dabei nach der Häufigkeit der Kookkurrenzen mit dem Ortsnamen sortiert. Als Anfrage für die Kookkurrenz wurde folgende Syntax verwendet *Stadtname + Ortsbezeichnung*, was nur Dokumente in die Trefferliste mit aufnimmt, in der beide Begriffe auch tatsächlich vorkommen.

Interessant ist vor allem das Ergebnis von Cardiff. Bei der Untersuchung der Texte bestätigte sich die Einschätzung aus Abschnitt 5.4.3, dass „Cardiff Castle“ durchaus das bedeutendste Bauwerk in Cardiff ist und mit Abstand am häufigsten mit „Cardiff“ zusammen in Texten zu finden ist (1.010.000 Fundstellen im Vergleich zu den 81.700 Fundstellen des zweithäufigsten Begriffs „Millenium Stadium“). Es bestätigt sich die Theorie, dass sich viele der Fotografen von den hohen zeitlichen (Öffnungszeiten) und räumlichen (Eintrittskosten) Beschränkungen in ihrem Verhalten beeinflussen lassen und daher eher Bilder ohne zusätzliche Einschränkungen machen. Während die Motivauswahl eines

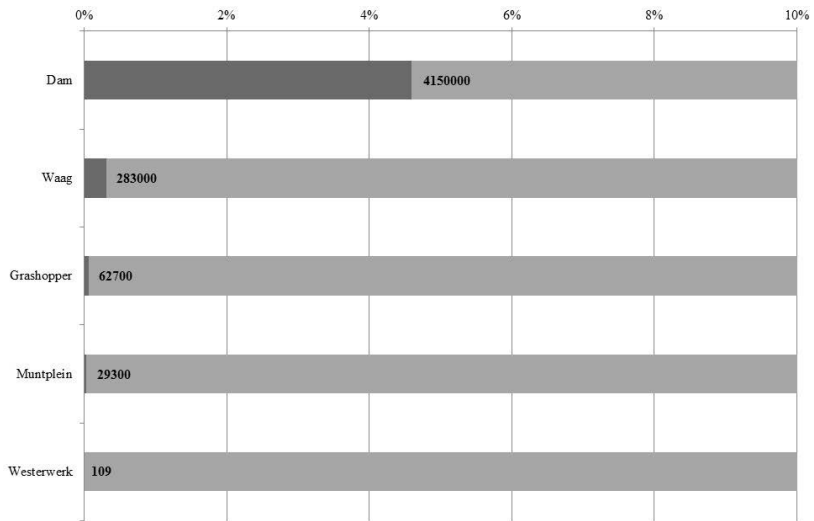


Abbildung 5.17: Häufigkeitsverteilung für Begriffe aus Amsterdam



Abbildung 5.18: Häufigkeitsverteilung für Begriffe aus Bamberg

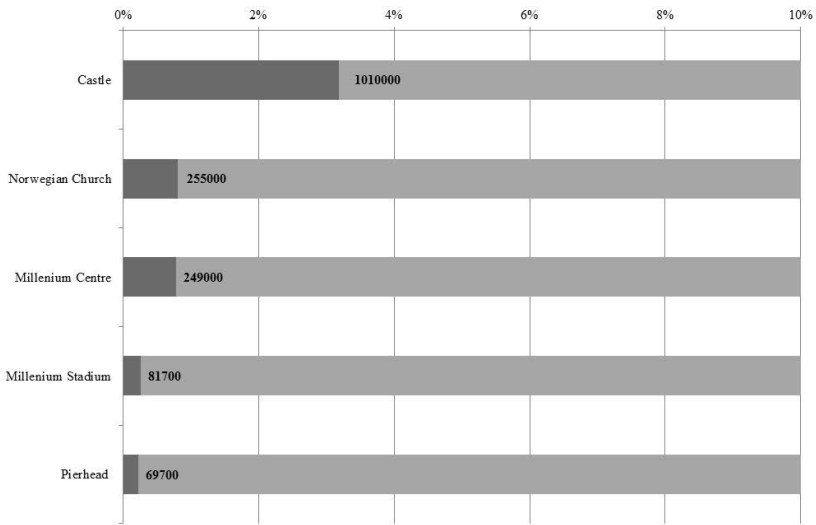


Abbildung 5.19: Häufigkeitsverteilung für Begriffe aus Cardiff

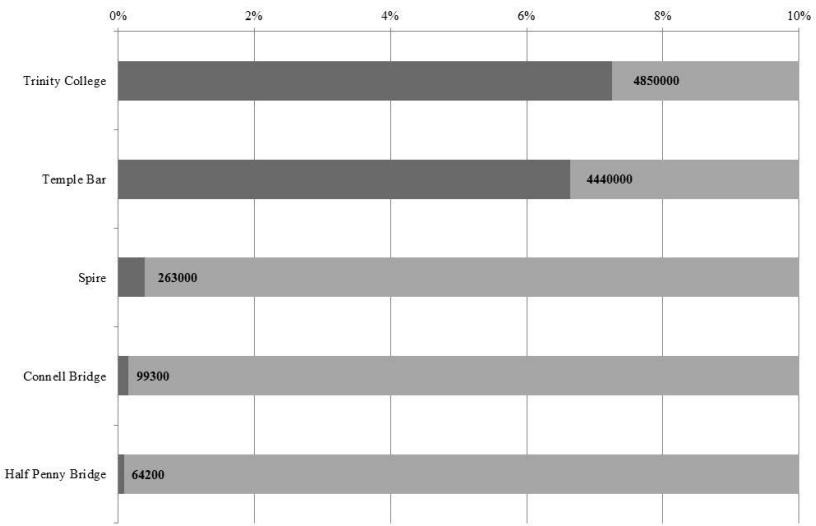


Abbildung 5.20: Häufigkeitsverteilung für Begriffe aus Dublin

Fotografen von Umwelteinflüssen beeinflusst wird, ist der Text weitgehend frei davon. Relativ zur Grundgesamtheit wird „Cardiff Castle“ in Texten öfters erwähnt, als auf Bilder festgehalten.

Auffallend verhalten sich die Häufigkeiten der beiden Eigenbegriffe „Trinity College“ und „Temple Bar“, für die im Zusammenhang mit „Dublin“ mehr Treffer möglich waren als ohne diesen Ortsbegriff. Die Anzahl wurde für die Analyse allerdings auf die Zahl der Einzeltreffer beschränkt, um die Darstellung der relativen Häufigkeiten nicht zu verletzen. Allerdings scheinen sie als Begriffe enger mit „Dublin“ verbunden zu sein als andere Begriffe. Ähnlich verbunden sind die Begriffe „Bamberg“ + „Altenburg“ ($\sim 3,7\%$) und „Amsterdam“ + „Dam“ ($\sim 2,5\%$).

Überraschend ist die Überprominenz der Begriffspaars „Bamberg“ + „Altenburg“, da die Altenburg in Marketingmaterialien der Stadt eher eine untergeordnete Rolle spielt und etwas abseits der Stadt liegt. Dennoch findet sich in öffentlichen Texten, die von Google indexiert wurden, eine hohe Anzahl von Fundstellen. Dabei lassen sich verschiedene Beweggründe, die Altenburg in einem Text zu erwähnen, aus den Fundstellen ablesen: Die Altenburg wird als eine der Sehenswürdigkeiten für die Region Bamberg beschrieben, sie beherbergt auf dem höchsten der insgesamt sieben Anhebungen von Bamberg ein gutes Restaurant und ist nicht zuletzt ein sehr beliebter Ort, um Hochzeiten zu feiern. Im Vergleich zu Bauten wie dem Bamberger Dom oder dem zentralen Rathaus ist es dennoch erstaunlich, so viele Treffer vorzufinden.

Tabelle 5.9 zeigt die Rangverschiebungen in der Rangfolge der Top 5, die sich durch die bildliche Konzeptualisierung ergibt (vgl. Abschnitt 5.3) und der dazu passenden Rangfolge durch die Häufigkeit in Textstellen. Besonders die Auswertung für Dublin zeigt sehr deutlich, dass man recht schnell die erste Hypothese bestätigen kann. Im Durchschnitt verschiebt sich jeder Begriff alleine in der Top 5 um beinahe 2 Ränge.

	Dam	Waag	Grashoper	Muntplein	Wester- werk
Amsterdam	1	4	3	5	2
	Alten- burg	Dom	Old town hall	Little Venice	Rose- garden
Bamberg	4	3	1	2	5
	Castle	Millenium Stadi- um	Norve- gian Church	Millen- ium Centre	Pierhead
Cardiff	3	1	5	4	2
	Trinity College	Temple Bar	Spire	Connell Bridge	Half Penny Bridge
Dublin	5	4	2	1	3

Tabelle 5.9: Rangverschiebung zwischen Text und Bildanalyse

Um die zweite Hypothese zu kontrollieren, wurden einige Stichproben mit Begriffen durchgeführt, die nicht in den Auswertungen der Bilder aufgekomen sind. Für Bamberg finden sich zum Beispiel Begriffspaare, wie „Bamberg“ + „Universität“, die mit 1.530.000 Treffern noch deutlich höher liegen als alle bisherigen Begriffe. Auch für Amsterdam liegen Begriffe wie „Van Gogh Museum“ oder „Rijksmuseum“ noch deutlich vor den Begriffen der Bildanalyse, allerdings weit hinter dem Begriff Dam, der den Begriff Amsterdam in beiden Untersuchungen prägt. Würde man daher die Textanalyse mit allen Begriffen durchführen, die in den geografischen Clustern der Heatmap identifiziert wurden, würde man deutlich größere Rangverschiebungen erzielen, als dies bei den untersuchten fünf Rängen der Fall ist.

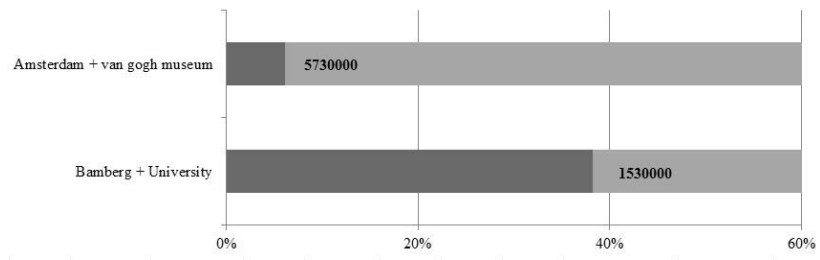


Abbildung 5.21: Häufigkeitsverteilung für „Bamberg“ + „Universität“ und „Amsterdam“ + „van Gogh museum“

5.5.2 Untersuchung einer Folksonomy

Die nächste Untersuchung analysiert die Texte, die bei den Bilder als Zusatzinformation zu finden sind. Insgesamt wurden um die 3.000 Bilder aus Panoramio, Flickr und Locr mit denselben geografischen Einschränkungen für Bamberg betrachtet, wie sie auch für die Berechnung der Heatmap verwendet wurden. Es handelt sich also um dieselben Bilder, wie bei der Suche nach einer bildlichen Konzeptualisierung. Al-

lerdings wird in dieser Untersuchung nicht der räumliche Standpunkt untersucht, sondern die Häufigkeit der Schlagworte in den Tags und den einzelnen Wörter aus den Bildtiteln (Bilder aus Panoramio beinhalten kaum Tags, daher wurden die einzelnen Wörter im Bildtitel wie Tags behandelt). Abbildung 5.22 zeigt die Schlagwörter, die einen bestimmten Ort oder Gebäude in der Stadt bezeichnen. Alle anderen Tags wurden aus dem Schaubild für eine bessere Übersichtlichkeit entfernt. Dieses Verfahren benötigt allerdings spezifisches Ortswissen, sodass die Analyse an dieser Stelle exemplarisch für Bamberg durchgeführt wurde. Was allerdings auch bei den anderen untersuchten Städten zu beobachten war, ist das erhöhte Aufkommen von Wörtern, die das Bild in einen übergeordneten geografischen Zusammenhang stellen. So war in allen Städten der Ortsname an erster Stelle zu finden. Bei Bamberg befand sich an erster Stelle der Begriff „bamberg“ in den Metadaten von insgesamt 1.509 Bildern (also in ungefähr jeden zweiten Bild). Das am meisten verwendete Tag bei den Bilder aus Amsterdam war folglich „amsterdam“, in Cardiff „cardiff“ und in Dublin „dublin“. Auf den nächsten Rängen waren Bezeichnungen für die Region oder das Land zu finden. In Bamberg belegten die Ränge zwei bis acht „germany“, „bavaria“, „deutschland“, „francia“, „franken“, „bayern“ und „oberfranken“. Dies lässt darauf schließen, dass die meisten Nutzer Tags eher dazu nutzen, um ihre Bilder räumlich zu gruppieren. Nur bei wenigen Bildern wird konkret der Inhalt des Bildes beschrieben. Es zeigt auch deutlich, dass in Tags das gleiche Konzept oft in mehreren Sprachen angegeben wird. Beispiele aus dem Datensatz sind „franken“ und „francia“ oder „bavaria“ und „bayern“.

Betrachtet man nun die Gesamtverteilung der Tags lassen sich recht schnell einige der drei angesprochenen Punkte mit Beispielen belegen. Das Diagramm in Abbildung 5.22 beinhaltet alle Tags, die mindestens 12 mal im Datensatz gefunden wurden. Das erste geografische Topo-

nym, das ein Gebäude beschreibt, verweist, wie auch schon bei den Postkarten und der Heatmap, auf Bilder des Rathauses. Schon bei der nächsten Fundstelle tritt Fall (1) ein und wir entdecken das Toponym „seehof“, das in den Bildern kaum eine Rolle spielt, aber als Begriff sehr oft gewählt wurde (insgesamt 76 mal). Zusätzlich befindet sich der Begriff „schlossseehof“, der dasselbe Toponym beschreibt, auf Platz 6. Schloss Seehof liegt recht weit außerhalb der Stadt. Dies führt zu der Annahme, dass Menschen, die es besuchen, zwangsläufig den Namen gehört haben müssen, da man nicht zufällig während einer Stadttour an ihm vorbeiläuft. Man muss sich also bewusst entschließen, zum Schloss Seehof zu fahren. Im Folgenden befinden sich drei Synonyme für den Bamberger Dom (bambergerdom, cathedral und kaiserdome) in der Verteilung bevor wieder der Fall (1) auftritt und der Tag „bahnhof“ in der Liste zu finden ist. Am Bahnhof selbst wurden nicht sehr viele Bilder aufgenommen, allerdings weisen fast alle den Tag „bahnhof“ auf (35 Stück). Weiter am Ende der Verteilung trifft man wieder vermehrt auf die Ortsbegriffe, die man schon in den beiden anderen Datenbeständen als typisch ausgemacht hatte, wie die Altenburg, Klein Venedig, den Rosengarten die Alte Hofhaltung und so weiter.

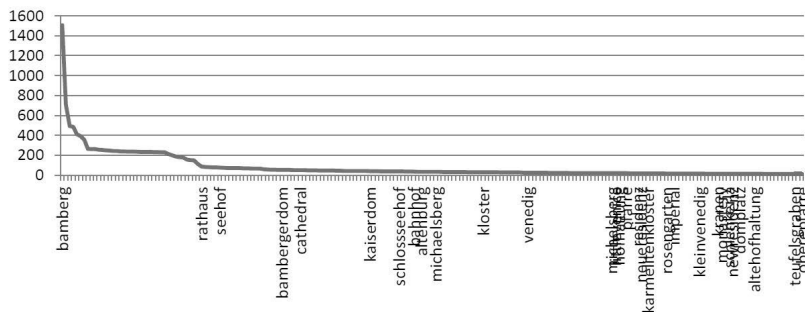


Abbildung 5.22: Auswertung der Tags auf den Bildern von Flickr, Panoramio und Locr

Es wurde versucht, eine automatische Identifikation von Toponymen mit Hilfe des freien Ortsnamenverzeichnisses oder auch Gazetteers Geonames zu erreichen. Dazu wurde das Schlagwort über den Geonames Service gesucht und überprüft, ob die zugehörige geografische Position in der momentanen Region enthalten ist. Dabei konnten allerdings einzelne bedeutende Gebäude, wie das Alte Rathaus, nicht mehr identifiziert werden. Daher ist diese Methode für unsere Analyse eher unbrauchbar. Tabelle 5.10 fasst dennoch die Ergebnisse kurz zusammen. Alleine Altenburg, Schillerplatz und Residenzschloss weisen den angestrebten Detailgrad auf, den auch die Einzelbilder der Postkarten aufgezeigt haben. Ansonsten finden sich viele Teilorte von Bamberg in den Ergebnissen sowie Flussbezeichnungen.

Bamberg	1520	regnitz	5	rothof	2
Bavaria	494	Gundelsheim	5	stegaurach	1
oberfranken	258	main-donau-kanal	3	michelsberger	1
Regnitz	82	Bughof	3	stegaurach	1
Seehof	76	Main	3	main-donau	1
altenburg	36	Memmelsdorf	3	unteraurach	1
Bayern	35	weltkurerbe	2	rhein-main-d...	1
bamberg	8	bamberg's	2	dörfleins	1
Baviera	7	schillerplatz	2	donau	1
Hallstadt	6	residenzschloss	2	bischberg	1
Europa	6	dörfleins	2	regnitz	1
oberhaid	6	mühlendorf	2	kreuzberg	1
Gaustadt	6	lichteneiche	2	windows	1
wildensorg	5	bischberg	2	kemmern	1

Tabelle 5.10: Häufigkeiten von Toponymen, die über Geonames identifiziert wurden

5.5.3 Kookkurrenzanalyse eines Textkorpus

Eine weitere Methode, um die Konzeptualisierung von Begriffen in Texten kenntlich zu machen, ist die Kookkurrenzanalyse. Man spricht von einer Kookkurrenz zwischen zwei Begriffen, wenn diese im selben Satz vorkommen oder höchstens durch eine bestimmte Anzahl von Wörtern voneinander getrennt sind. Die Idee ist, dieses Verfahren auf des Toponym Bamberg anzuwenden, um einen weiteren Vergleich von bildlicher Konzeption mit der Konzeption in Texten zu erhalten. Angewendet wird diese Methode allerdings nicht auf Bilder, sondern auf einen ganzen Textkorpus, der Texte von verschiedenen Onlinequellen beinhaltet, die mit dem Begriff Bamberg in Verbindung stehen. Zunächst wird versucht, einen guten Überblick über aufkommende Schlüsselbegriffe innerhalb dieser Texte zu geben. Dies wird auch als Screening (Chilla 2005) bezeichnet. Dazu werden sowohl qualitative, wie auch quantitative Methoden der Textanalyse miteinander verbunden. Zum einen wird versucht mit Textmining-Verfahren sinntragende Wortgruppen in den Texten zu finden und zum anderen die semantische Vernetzung über association rules zu berechnen. Die Berechnung der association rules beruht dabei auf der Kookkurrenz der sinntragenden Wortgruppen.

Für die Analyse wurde das Programm Sascet verwendet, das am Lehrstuhl für Kulturinformatik an der Universität Bamberg entwickelt wurde. Im zugrundeliegenden Textkorpus finden sich Texte aus verschiedenen Enzyklopädien (Wikipedia, Encyclopædia Britannica, Catholic Encyclopedia, ...), sowie die offizielle Darstellung der Stadt und auf den Seiten der UNESCO und schliesslich auch Texte von verschiedenen Reiseführern über Bamberg (Lonely Planet, Tripadvisor). Für die Analyse mit Sascet konnten allerdings nur englische Begriffe verwendet werden, so dass die Ergebnisse, in Abbildung 5.23 dargestellt, auch nur englische Begriffe beinhalten.

Für die Auswertung dieser Textquellen lässt sich mit Hilfe von Sascet ein Graph, wie er in Abbildung 5.23 zu sehen ist, erzeugen. Es lässt sich recht schnell sehen, dass nur recht abstrakte Begriffe mit Bamberg selbst verbunden sind. So wird die Stadt an sich nur selten mit einem herausragenden Gebäude in Verbindung gebracht, sondern gerne als sehr facettenreiche Stadt dargestellt. Eine herausragende Facette ist dabei sicher die Sicht der Stadt als Bierstadt. Mit neun Brauereien mit alter Tradition in Bamberg und insgesamt der höchsten Konzentration von Brauereien in der ganzen Welt, ist diese Verbindung mehr als naheliegend. Interessant ist auch die Verbindung zu Rom: Bamberg wird gerne mit Rom verglichen, da die Stadt auch auf sieben Hügeln erbaut wurde. Für eine bildliche Illustration dieser Begriffe würde man, wie auch schon bei der Postkarte, die versucht, einen guten Überblick über die typischen Sehenswürdigkeiten der Stadt zu bieten, mehrere Bilder benötigen. Das Konzept der Bierstadt würde durch eine Bildgruppe aus neun Bildern der historischen Brauereien gut illustriert werden können, während das Konzept „fränkisches Rom“ durch Ansichten auf die sieben Hügel ebenso gut dargestellt werden kann. In touristischen Aufnahmen wird man derartige Bildgruppen allerdings sehr selten finden, da solche Aufnahmen sehr viel Hintergrundwissen und Planung erfordern. Andersherum argumentiert würden gerade solche Bildgruppen aus der Masse anderer touristischen Bildgruppen herausstechen.

Abseits des Clusters um Bamberg lassen sich dann auch wieder Begriffe, wie „cathedral“, „old town hall“, „castle altenburg“, „new residenz“ oder „rose garden“ finden. Interessant ist, dass bestimmte Orte oft als Einheit zu finden sind, wie das Alte Rathaus und der Dom, die in den Texten gern als Beispiel für die architektonischen Höhepunkte der Stadt gewählt werden, oder die Neue Residenz und der Rosengarten, der streng genommen ein Teil der Neuen Residenz ist. Andere Begriffe, die in dem Graphen zu sehen sind, bedürfen einer genaueren Untersuchung

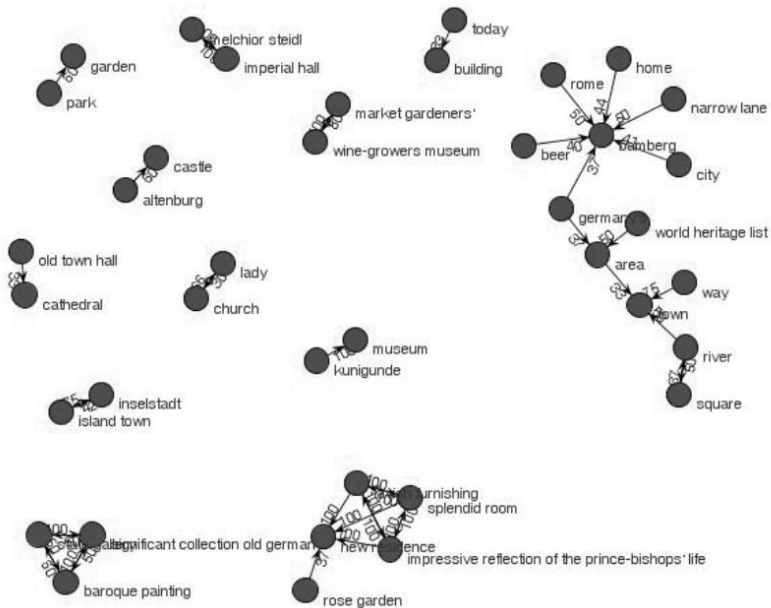


Abbildung 5.23: Kookurrenzanalyse für Texte über Bamberg. Rote Knoten stehen für einen Begriff, eine Kante für besonders häufiges gemeinsames Auftreten der beiden Begriffe.

der Textfunde, um eine genauere Interpretation zu geben. So gibt die Verbindung von *today* und *building an*, dass die meisten historischen Gebäude der Stadt heutzutage einen ganz anderen Verwendungszweck haben, als noch zu ihrer Entstehungszeit. Diese zeitlichen Nutzungsbedingungen lassen sich allerdings schlecht mit Bildern darstellen, da sich durch die neue Funktion das Aussehen der Gebäude meistens nicht verändert hat; dies ist daher nur in Textform vermittelbar.

5.6 Zusammenfassung

Wir haben gesehen, wie man mit einfachen heuristischen Verfahren eine bildliche Konzeptualisierung aus verorteten Bildsammlungen berechnen kann. Die Heatmap eines Gebietes entsteht dabei ausschließlich aus geografischen Metainformationen und dem kollaborativen Charakter der Ausgangsdaten. Bilder werden geografisch gruppiert und dadurch werden die einzelnen Cluster von der Nutzergemeinschaft durch ihre Bilder bewertbar. Als Ergebnis entsteht ein semantisches Modell, das jedem Ort eine bestimmte Popularität zuordnet. Es hat sich auch erwiesen, dass sich zum einen eine andere Konzeptualisierung durch Texte bestimmen lässt und durch Texte eine ganz andere Gewichtung auf die einzelnen Aspekte einer Stadt gelegt wird. Zum anderen ist eine Konzeptualisierung durch Texte sehr schwer ohne menschliche Unterstützung berechenbar, da man die Ergebnisse aus den Frequenzanalysen erst noch für sich interpretieren muss.

Auch wenn das Konzept einer Stadt aus der Gesamtsumme aller Entscheidungen berechnet werden kann, verwendet jeder Benutzer für sich gesehen eine eigene Konzeptualisierung. Dies zeigt sich vor allem am langen Schwanz der Rang-Popularitäts-Kurve. Während sich im Kopf der Verteilung der semantische Kern des räumlichen Konzepts befindet, sind im langen Schwanz die Unterschiede zwischen den Nutzern

auszumachen. Wir werden im nächsten Kapitel diese Beobachtungen dazu nutzen, ein Maß zu erarbeiten, das die Ähnlichkeit von Nutzern über ihre Konzeptualisierung des Raumes entsprechend gewichtet. Mit Hilfe der ähnlichsten Nutzer lassen sich daraufhin Aussagen über die persönliche Konzeptionalisierung in einer neuen Region vorhersagen.

6 Persönliche Bildempfehlungen

In Kapitel 6 wurde ein Verfahren beschrieben, das typische Aufnahmen aus einer zugrundeliegende Nutzergemeinschaften auslesen kann. Diese Bildgruppen weisen durch die räumliche Gruppierung in Bildcluster auch eine räumliche Diversität auf. Dennoch fehlt der besondere Bezug zu der Person, die eine Anfrage an das Empfehlungssystem stellt. Wir wollen nun die persönlichen räumlichen Erfahrungen des Nutzers mit dem Verhalten von anderen Nutzern vergleichen. Dazu muss der Benutzer allerdings selbst aktiv werden und dem System seine räumliche Präferenzen mitteilen. Es wird, ausgehend von einem benutzerbasierten kollaborativen Ansatz, ein Verfahren entwickelt, das diese räumlichen Präferenzen für die Berechnung von Nutzerähnlichkeiten nutzt und daraus neue Vorschläge generiert. Die grundlegenden Überlegungen zu diesem Kapitel wurden in Matyas und Schlieder (2009) veröffentlicht. Zunächst wurde der passende Anwendungsfall entwickelt, der den Einsatz eines kollaborativen Filters möglich macht. Dieser Anwendungsfall wird im folgenden Abschnitt beschrieben.

6.1 Anwendungsfall: Personalisierte Bildgruppen

In diesem Anwendungsfall ist der Nutzer nicht nur an Bildern interessiert, die typisch für das Gebiet sind, sondern auch dem persönlichen räumlichen Verhalten des Nutzers entsprechen. Damit kann der Nutzer zum Beispiel einen persönlichen Eindruck von seinem nächsten Ur-

laubsziel erhalten und Bilder von Orten entdecken, die er dort selbst fotografieren würde.

Eine passende Anfrage an das System könnte lauten: „Ich würde gerne 10 Bilder von Dublin sehen, die mich interessieren könnten“. Abbildung 6.1 zeigt, wie diese Anfrage mit Hilfe eines kollaborativen Ansatzes beantwortet werden soll.

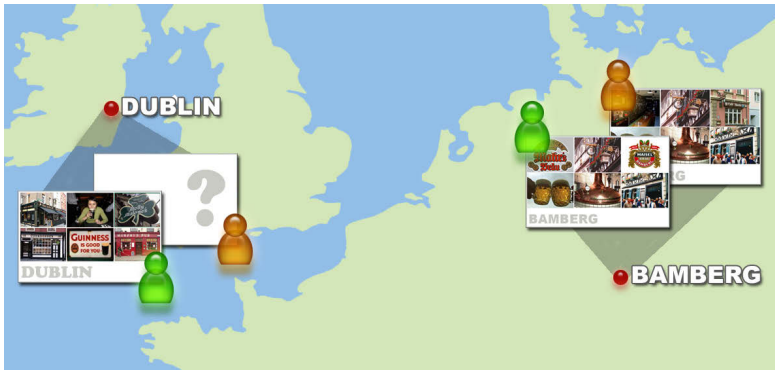


Abbildung 6.1: Kollaborativer Ansatz für ein geografisches Empfehlungssystem

Angenommen der orangene Nutzer will eine Empfehlung für Dublin erhalten, dann greift das System auf seine bisherigen räumlichen Entscheidungen zurück. In diesem Fall sind das sechs Entscheidungen, die der Nutzer in Bamberg getroffen hat. Das System vergleicht nun diese Entscheidungen mit den Entscheidungen von anderen Nutzern aus Bamberg und findet heraus, dass der grüne Nutzer ein recht ähnliches Verhalten aufgezeigt hat. Glücklicherweise hat dieser Nutzer bereits Entscheidungen in Dublin getroffen. Diese Aufnahmen werden dem orangenen Nutzer als Empfehlung vorgeschlagen. In der Regel wird man nicht nur einen ähnlichen Nutzer in Bamberg finden und somit nicht nur eine eindeutige Menge von Entscheidungen für Dublin. Mög-

liche Lösungsansätze, die Ähnlichkeit der Nutzer zu bestimmen und die gefundenen Entscheidungen zu einer neuen Empfehlung zu vereinen, werden in den nächsten Abschnitten entwickelt.

Für den Anwendungsfall können folgende Akteure festgelegt werden:

- *Aktiver Benutzer*: Ist der Benutzer, der momentan eine Anfrage stellt; entspricht also dem grünen Nutzer aus dem vorherigen Beispiel. Im Gegensatz zum Nutzer aus dem ersten Anwendungsfall (Kapitel 5.1. Typische Bildgruppen einer bestimmten Region) muss dieser Nutzer selbst aktiv werden und Bilder von sich zur Verfügung stellen, die das System analysieren kann. Erst wenn der Nutzer wirklich aktiv war, kann das System eine Empfehlung für ihn aussprechen.
- *Fotograf*: Jede Person dieser Gruppe kann theoretisch zum aktiven Nutzer werden, wenn er seine eigenen Bilder für das System zugänglich macht. Sinnvoll ist es, diese Gruppe weiter zu unterteilen. In eine interne Gruppe, für die ebenfalls eine Empfehlung ausgesprochen werden kann und in Nutzer, die mit ihren Aufnahmen nur zusätzliches Hintergrundwissen beitragen, um etwa die Popularität eines Ortes zu bestimmen. Der Vorteil dieser Trennung liegt darin, dass die Aufnahmen der zweiten Gruppe für eine Empfehlung ausgeschlossen werden können. Dadurch werden nur Bilder von ausgewählten Nutzern empfohlen. Zum Beispiel könnte das System Bilder aus Flickr und Panoramio während der Modelerzeugungsphase verwenden. Die Auswahl der Bilder könnte dann auf Aufnahmen von Panoramio beschränkt bleiben.

Es werden folgende Anwendungsfälle von dem System verwirklicht:

- *Empfehlung anfragen*: Der aktive Nutzer stellt eine Anfrage an das System, in dem er eindeutig eine Region seiner Wahl und die Anzahl der Bilder bestimmt. Das System berechnet ihm dann eine

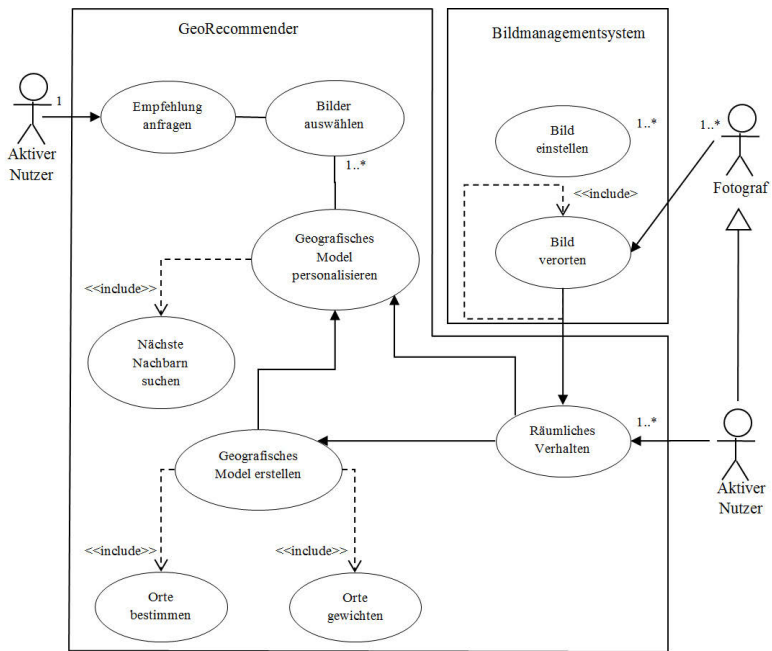


Abbildung 6.2: Anwendungsdiagramm Bildempfehlungen

personalisierte Bildsammlung, in der diese Anzahl von Bildern enthalten ist.

- *Bild einstellen:* In diesem Fall können zwei getrennte Datenquellen betrachtet, werden in die verschiedene Benutzer Bilder einstellen können. Wichtig ist, dass der aktive Benutzer selbst Bilder eingestellt hat, bevor er eine Empfehlung abfragt. Ansonsten ist das System höchstens in der Lage eine Empfehlung wie im Kapitel 5 zu erzeugen.
- *Bild verorten:* Die Positionen der Bilder werden mit Hilfe von GPS oder über eine Karte ermittelt. Räumliches Verhalten: Jedes georeferenzierte Bild wird vom Empfehlungssystem als räumliches Verhalten aufgefasst. Räumliches Verhalten wurde in Kapitel 5 genauer beschrieben.
- *Geografisches Modell erstellen:* Aus dem räumlichen Verhalten wird ein geografisches Modell des Gebietes erzeugt, das zu jedem Ort seine Popularität angibt. Diese Aktion schließt die folgenden Aktionen mit ein:
 - *Orte bestimmen:* Zunächst muss das System festlegen, was als eigenständige Orte gelten soll. Mit diesem Schritt beschäftigt sich Abschnitt 5.4.
 - *Orte gewichten:* Als nächster Schritt werden die berechneten Orte nach dem verwendeten Popularitätsmaß gewichtet. Ein Popularitätsmaß wird in Abschnitt 6.3.2. entwickelt
- *Geografisches Modell personalisieren:* Ausgehend von dem Modell das die Orte der Region nach ihrer Popularität gewichtet hat, werden die Orte nach einem neuen Maßstab neu gewichtet. Dieses Gewicht soll die Relevanz für den Nutzer aufzeigen.

- *Nächste Nachbarn suchen:* Um eine neue Gewichtung zu erreichen, müssen die ähnlichsten Nutzer gefunden werden. Dazu wird ein Ähnlichkeitsmaß verwendet.
- *Bilder auswählen:* Nachdem die verschiedenen Orte nach Relevanz für den Nutzer sortiert wurden, muss das System ein Bild als Repräsentant für einen Ort auswählen. Diese Auswahl ergibt dann die eigentliche Empfehlung.

Die Anwendung selbst kann zum Beispiel mit einer Online-Bildergalerie zusammenarbeiten. Diese übernimmt dann die entsprechenden Anwendungsfälle, die als „Bild einstellen“ und „Bild verorten“ in Abbildung 6.2 ausgewiesen sind. Bestehende Systeme dieser Art gibt es bereits unter Flickr, Panoramio, Picassaweb, Locr und viele mehr. Benutzer stellen Bilder für dieses System zur Verfügung und identifizieren sich dem Empfehlungssystem gegenüber mit derselben Nutzerkennung, damit eine eindeutige Zuordnung des Nutzers in beiden Systemen möglich ist. Aus dieser Bildmenge werden dann die Empfehlungen generiert. Zusätzlich können als zusätzliche Hintergrundinformationen auch andere georeferenzierte Datenquellen verwendet werden. Die meisten von ihnen bieten, ganz im Sinne von Web2.0, eine öffentliche Schnittstelle an, mit deren Hilfe man auf öffentliche Daten zugreifen kann. Diese externen Metadaten werden für den internen Analyseschritt verwendet. Auch wenn die Abbildung 6.2 nur ein Bildmanagementsystem aufzeigt, können an dieser Stelle mehrere verschiedene georeferenzierte Bildsammlungen zum Einsatz kommen. In allen Fällen werden die persönlichen Geo-Tags der Bilder analysiert (ein Tripel aus Nutzerkennung, Bildkennung und GPS-Position).

Kollaborative Ansätze leben von der aktiven Teilnahme des Nutzers, das heißt er gibt für einige Objekte eine persönliche Bewertung als Feedback an das Empfehlungssystem zurück. Wir werden im nächsten Abschnitt sehen, wie sich diese Bewertungen durch einen nutzerbasier-

ten kollaborativen Filter nutzen lassen, um Objekte nach persönlichen Vorlieben zu filtern. In Kapitel 6.3 werden wir diesen generischen Ansatz auf den konkreten geografischen Anwendungsfall umsetzen. Ein Nutzer in dieser Definition eines geografischen Bildempfehlungssystems ist folglich dann aktiv, wenn auf die geografischen Positionsdaten in seinen Bildern zugegriffen werden kann. Die einzelnen Teilschritte werden dabei für den hier beschriebenen Anwendungsfall adaptiert.

6.2 Generischer kollaborativer Empfehlungsansatz

Dieser Abschnitt gibt einen detaillierten Einblick in die Funktionsweise und die Methodik eines kollaborativen Filters, wie er in den Arbeiten von Resnick (1994) geprägt würde. Die Abgrenzung zu anderen Empfehlungsverfahren findet man in Abschnitt 4, zusammen mit den Stärken und Schwächen der einzelnen Systeme. Führen wir uns die Abbildung 4.5 in Abschnitt 4 noch einmal vor Augen, dann stellen wir fest, dass kollaborative Filter vom Nutzer verlangen, dass die Nutzer bereits bestimmte Gegenstände bewertet haben. Auf dieser Grundlage berechnet das System eine Vorhersage für eine Bewertung für Objekte, die der einzelne Nutzer selbst noch nicht bewertet hat. Ein benutzerbasierter kollaborativer Empfehlungsansatz durchläuft dabei mehrere Teilschritte:

- Nutzer geben Feedback über einzelne Objekte ab
- Berechnung der Nutzerähnlichkeiten anhand ihres Nutzerfeedbacks
- Auswahl einer Teilmenge der Nutzer
- Aggregation der Erfahrungen der nächsten Nachbarn

Schritt 1: Jedes Empfehlungssystem verlangt vom Benutzer ein Mindestmaß an Feedback. Dieses Feedback kann das Ausfüllen von For-

mularen mit persönlichen Angaben sein oder auch nur die Auswahl eines Gegenstandes, der einen momentan am meisten interessiert. In nutzerbasierten kollaborativen Systemen geht es darum mit den unterschiedlichen Nutzungspräferenzen der Benutzer umzugehen. Dazu muss er oder sie diese Präferenzen auf einem Teil der Sammlung auch zum Ausdruck bringen. In den meisten benutzerbasierten Verfahren geht man davon aus, dass der Nutzer eine explizite Bewertung mit Hilfe einer bestimmten Bewertungsskala abgibt. In den meisten Fällen wird dies auf einer Skala von 1 bis 5 verlangt.

$$R_{x,i} \tag{6.1}$$

gibt dabei die abgegebene Wertung des Nutzers x für den Gegenstand i an.

$$R_x \tag{6.2}$$

ist der Vektor, der aus allen Bewertungen des Nutzers x auf allen möglichen Gegenständen besteht und

$$\overline{R_{x,i}} \tag{6.3}$$

ist die durchschnittliche Bewertung dieses Nutzers auf allen Gegenständen, die er seither bewertet hat. Diese Erhebungsphase wird auch Ramp-up genannt.

Schritt 2: Die Ähnlichkeiten der Benutzer werden je nach Art des Nutzerfeedbacks unterschiedlich berechnet. Grundsätzlich gibt es für jeden Benutzer x einen Bewertungsvektor R_x , in dem alle Entscheidungen zu bestimmten Gegenständen enthalten sind. Diese Vektoren können mit Hilfe von verschiedenen Verfahren verglichen werden. Resnick verwen-

deten 1994 Pearsons Korrelationsfaktor, der die lineare Abhängigkeit der zwei Featurevektoren berechnet.

$$sim_{pearson}(x, y) = \frac{\sum_{i \in I} (R_{x,i} - \bar{R}_x)(R_{y,i} - \bar{R}_y)}{\sqrt{\sum_{i \in I} (R_{x,i} - \bar{R}_x)^2} \sqrt{\sum_{i \in I} (R_{y,i} - \bar{R}_y)^2}} \quad (6.4)$$

Alternativ wird auch die Cosinusähnlichkeit verwendet, die den Vorteil hat, die Vektoren beim Vergleich automatisch zu normalisieren, da nur der Winkel zwischen ihnen entscheidend ist und nicht ihre Länge. Daher ist das Ergebnis das Gleiche, wenn die Gemeinsamkeiten zwischen den Einzelentscheidungen beider Nutzer relativ konstant bleiben, auch wenn einer von beiden insgesamt immer höher bewertet als der andere.

$$sim_{cosine}(x, y) = \frac{\sum_{i \in I} R_{x,i} R_{y,i}}{\sqrt{\sum_{i \in I} (R_{x,i})^2} \sqrt{\sum_{i \in I} (R_{y,i})^2}} \quad (6.5)$$

Schritt 3: Auf Grundlage der berechneten Ähnlichkeiten wird eine Menge von Nutzer gewählt, die für den Benutzer x die wahrscheinlich beste Empfehlung aussprechen kann. Man nennt diese Menge auch die nächsten Nachbarn (nearest Neighbors) oder auch die Nutzergemeinschaft des Nutzers x .

- Top N: Es werden die ersten n Benutzer gewählt, die höchste Ähnlichkeit aufweisen (Resnick u. a., 1994).
- Threshold: Benutzer werden nur ausgewählt, wenn der Wert ihrer Ähnlichkeit eine bestimmte Bedingung erfüllt. Der berechnete Ähnlichkeitswert liegt zum Beispiel oberhalb einer bestimmten Schranke (Shardanand und Maes, 1995). Dabei ist zu beachten, dass man abhängig von dem gewählten Korrelationsmaß nicht zwangsläufig nur die ähnlichsten Nutzer wählen muss. Im Falle des Pearson-Koeffizienten können auch sehr unähnliche Nutzer

(nahe bei -1) gewählt werden, um zu einem guten Ergebnis zu kommen.

- Kombination aus beiden Ansätzen: Es werden die ersten n Nutzer gewählt, deren Ähnlichkeit zusätzlich eine festgelegte Bedingung erfüllen.

Schritt 4: Als letzter Schritt folgt die Aggregation der Präferenzen, die von den nächsten Nachbarn festgelegt wurden. Diese nächsten Nachbarn werden trotz ähnlichem Verhalten in der Vergangenheit, nicht die gleiche Meinung über einen bestimmten Gegenstand i haben. Es ist daher ein weiterer Kombinationsschritt notwendig. Üblicherweise berechnet man das Mittelmaß der Wertungsschwankungen der nächsten Nachbarn, gewichtet mit deren Ähnlichkeiten. Das Ergebnis addiert man auf den Mittelwert der Wertungen des aktuellen Nutzers. Dieser berechnete Wert schätzt eine mögliche Bewertung des Gegenstands i durch den Nutzer x ab. Das heißt, es wird der Durchschnitt der eigenen Wertungen entweder erhöht oder gesenkt, je nachdem wie die Bewertungen der nächsten Nachbarn in der Vergangenheit mit dem Nutzer x übereingestimmt haben. Das Ergebnis der kollaborativen Filterung liegt bei den meisten Verfahren wieder innerhalb der festgelegten numerischen Skala, die schon für Bewertungen der Nutzer verwendet wurde. Führt man diesen letzten Schritt mit allen unbewerteten Gegenständen durch, entsteht dadurch ein Ranking nach vorhergesagten persönlichen Vorlieben.

$$\widehat{R_{x,i}} = \overline{R_x} + \frac{\sum_{y \in U_x} (R_{y,i} - \overline{R_y}) \text{sim}(x, y)}{\sum_{y \in U_x} |\text{sim}(x, y)|} \quad (6.6)$$

Diese Schritte werden wir nun für die Vorhersage von georeferenzierten Bildgruppen umsetzen. Als Grundlage dient dazu die Gewichtung nach Popularität der gruppierten räumlichen Entscheidungen, die wir mit Hilfe des Heatmapping Verfahren berechnet haben. Ziel dieser Ge-

wichtung ist es, die persönlichen Entscheidungen, die sich im langen Schwanz der Popularitätsverteilung befinden, möglichst hoch zu bewerten. Dadurch wird das bestehende Ranking verändert und es wird eine persönlichere Rangfolge generiert. Im Kopf dieser Verteilung befinden sich dann die relevanten Orte für den entsprechenden Nutzer.

6.3 Geografischer Empfehlungsansatz

In den nächsten Abschnitten werden wir nun das generische Verfahren aus Abschnitt 6.2 für eine geografische Empfehlung anpassen und geeignete Maße erarbeiteten, die dem geografischen Kontext gerecht werden.

6.3.1 Schritt 1: Benutzer-Feedback

Erstes Ziel bei der Umsetzung des generischen Ansatzes eines benutzerbasierten kollaborativen Filters für geografische Daten war es, ein geeignetes Feedback des Benutzers zu finden, an dem sich seine Präferenzen ablesen lassen. Wie zuvor erwähnt wird in der Regel eine explizite Bewertung von Nutzer verlangt. Wie diese Präferenzen vom Nutzer abgefragt werden ist dabei eine sehr wichtige Designentscheidung, wenn man ein kollaboratives System verwenden will. Die Bewertungsskalen werden von den Nutzern nicht immer selbstverständlich angenommen oder richtig verstanden. Es zeigt sich oft ein deutlicher Unterschied zwischen den künstlich generierten Beispielen in wissenschaftlichen Arbeiten und dem täglichen Umgang mit solchen Empfehlungsansätzen.

Um dies zu verdeutlichen, wollen wir uns zunächst die folgende Statistik (Abbildung 6.3) etwas genauer betrachten, die von den Betreibern der Onlineplattform Youtube 2010 erhoben wurde. Sie beschreibt die Häufigkeit einer bestimmten Bewertung auf einer 5-Punkte-Skala ver-

wendet wurde. Dieser Fall verdeutlicht recht anschaulich die Probleme, die auftreten können wenn explizite Bewertungen auf einer festen Skala verwendet werden. Zunächst lässt sich deutlich erkennen, dass die meisten Benutzer überwiegend fünf Sterne vergeben. Obwohl auch differenzierte Aussagen möglich sind, verwenden Youtube-Nutzer diese Skala um ihre vollständige Zustimmung zu äußern. Fünf Sterne sind somit gleichbedeutend mit der Aussage „Ja, dieses Video ist gut“. Ganz selten wird ein Stern benutzt um auszusagen, dass ein Video besonders schlecht ist. Diese Videos werden dann meistens mit einem einzigen Stern ausgezeichnet. Zwei, drei oder vier Sterne werden so gut wie nie verwendet. Ein Empfehlungssystem wird mit dieser Datengrundlage lediglich fünf Sterne vergeben können oder gar keine Aussage treffen können. Detailliertere Aussagen sind daher nicht möglich. Im Laufe der Umgestaltung von Youtube im Jahr 2012, wurde die Bewertungsskala auf zwei Werte reduziert („Mag ich“, „Mag ich nicht“).

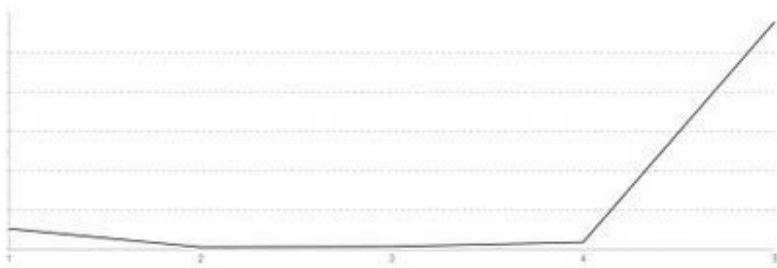


Abbildung 6.3: Bewertungsverhalten auf Youtube (Quelle: Youtube).

Eine weitere Beobachtung, die sich nicht nur auf Youtube beobachten lässt, ist die Vermutung, dass es von den Nutzern im Allgemeinen als eine sehr mühsame Aufgabe empfunden wird, detaillierte Bewertungen abzugeben. Nicht ohne Grund versenden eCommerce-Plattformen wie Amazon immer wieder E-Mails, in denen sie ihre Kunden auffordern, ihre gekauften Artikel zu bewerten. Oft ist dem Kunden der Nutzen

nicht einsichtig genug, um dieser Bitte langfristig auch Folge zu leisten. Werden allerdings nicht ständig neue Bewertungen für neue Gegenstände abgegeben, stagnierte die Menge an Objekten, die das System empfehlen kann. Aus diesem Grund verwenden viele Systeme implizite Nutzungsdaten, wie Click Through Informationen oder Kaufentscheidungen, um zumindest implizit Informationen zu sammeln.

Ein geeignetes Bewertungsschema für das geografische Empfehlungssystem sollte die Erkenntnisse aus den beiden Beobachtungen mit beachten. Zum einen sollten die nötigen Informationen implizit aus dem Nutzungsverhalten der Nutzer erschlossen werden und zum anderen sollte nicht jeder Empfehlungsgegenstand gleich gewichtet werden. Zunächst stellt sich allerdings die Frage, was das Gegenstück einer Bewertung im geografischen Raum ist.

Wie schon im Kapitel zuvor beschrieben, hinterlässt der Nutzer beim Erstellen eines Bildes eine zweidimensionale Ortsangabe, seinen digitalen Fußabdruck. An diesen Positionen führt der Benutzer eine bestimmte Aktion aus. Damit signalisiert er, dass er zu diesem Ort eine besondere Beziehung hat. In Fall von Fotografien hält er seine momentane Sicht auf einem Bild fest. Durch ein einzelnes verortetes Bild fällt dabei jeder Nutzer eine binäre Entscheidung in einem zweidimensionalen Raum. Wie allerdings schon bei dem generischen Ansatz beschrieben, basiert ein kollaborativer Filter darauf, dass verschiedene Nutzer auf gemeinsamen Objekten eine Entscheidung treffen. Würde man die GPS-Position als grundlegendes Objekt festlegen, würden sich dadurch kaum Übereinstimmungen bei den Nutzern feststellen lassen. Zum einen ist die GPS-Position viel zu präzise, zum anderen ist sie im Vergleich zur eigentlichen Position des Nutzers recht ungenau. Um dennoch eine Übereinstimmung bei den Nutzerentscheidungen zu finden, musste ein Raumunterteilungsverfahren eingesetzt werden, dass die gemeinsamen Entscheidungen gruppiert.

Hinterlässt ein Nutzer einen digitalen Fußabdruck innerhalb eines der Cluster, steigt dadurch das Popularitätsmaß dieses Clusters und man kann dies als eine zusätzliche Bewertung dieses Ortes ansehen. Dies ist vergleichbar mit Favoriten auf Onlineplattformen, wie eben auch Youtube. Hinterlässt der Nutzer mehrere Fußabdrücke in einem Cluster, signalisiert er damit ein gesteigertes Interesse an diesem Ort. Bewertet man diese vermehrte Teilnahme an einem Cluster, entsteht dadurch eine nach oben offene Bewertungsskala.

Zur Hilfenahme einer Nearest Neighbor Gruppierung mit einem flexiblen k (kNN) und dem Bewertungsschema nach Popularität haben wir in Kapitel 5 typische Fotografenstandpunkte berechnet. Wir verwenden nun dieselbe Raumzerlegung um festzustellen, an welchen Fotografenstandpunkten ein Nutzer gemeinsam mit anderen Nutzern eine Entscheidung getroffen hat. Diese gemeinsamen Entscheidungen werden im Folgenden für die Berechnung der Ähnlichkeiten wichtig, da Nutzer im Sinne eines kollaborativen Ansatzes sich nur ähnlich sein können, wenn sie Entscheidungen auf denselben Objekten treffen. Tabelle 6.1 zeigt neben dem verwendeten räumlichen Clusterverfahren auch andere Unterteilungsmethoden, die stattdessen verwendet werden können.

Durch die Zerteilung des Raumes entsteht nicht nur eine persönliche Gewichtung von einzelnen Orten durch einen bestimmten Nutzer, auch verschiebt sich der eigentliche Empfehlungsansatz von der Empfehlung von einzelnen Bildern hin zu der Empfehlung von Orten, an denen diese Bilder entstanden sind. Das heißt, das geografische Empfehlungssystem, wie es hier vorgestellt wird, verwendet als Ausgangsmenge keine Menge von Gegenständen, empfiehlt daher auch nicht direkt eine Menge von Bildern, sondern nur eine Zusammenstellung von mehreren Standorten. Das hat auch damit zu tun, dass wir uns in Kapitel 5 auf die Unter-

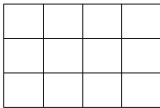
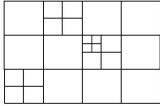
	Das Gitter zerlegt den Raum in einem zuvor definierten Abstand, sowohl in Richtung Längengrad als auch in Richtung Breitengrad. Durch diese Art der Zerlegung läuft man allerdings Gefahr, dass man entscheidende Orte in bis zu vier Teilorte zerlegt und sich die räumlichen Entscheidungen in diese Teilgebiete aufteilen. Dadurch kann es vorkommen, dass dieser Ort unterbewertet wird. Auch die gemeinsamen Entscheidungen der Nutzer würden sich dementsprechend aufteilen, was direkt die Nutzerähnlichkeit verringern würde. Diese Zerlegung von Entscheidungen ist umso wahrscheinlicher je kleiner das Gitter ist. Allerdings sind differenzierte Aussagen auch nur mit einem recht schmalen Gitter möglich.
	Durch Quadtree wird der Raum dynamisch aufgeteilt. Zunächst wird der Raum durch ein recht grobes Gitter zerlegt und dann werden die Gitterflächen schrittweise verfeinert, in denen eine zuvor festgelegte Anzahl von Entscheidungen liegen. Allerdings haben die Zerlegungsflächen, wie auch beim Gitter, ein festes Größenverhältnis, sodass dieselben Probleme auftreten werden.

Tabelle 6.1: Raumzerlegungsverfahren

	Ein Voronoi-Diagramm ergibt sich aus der Festlegung von mehreren Punkten, die den Zentren der Regionen entsprechen, die durch die Zerlegung entstehen. Ein Punkt in einer bestimmten Region ist hierbei immer näher zum entsprechenden Zentrum als zu allen anderen Zentren. Mit Hilfe dieser Zerlegung könnte man den Raum um ausgewiesene Sehenswürdigkeiten zerlegen.
	Die Variante von Nearest Neighbor, die im Heatmap Verfahren verwendet wird, sucht ausgehend von einem Startpunkt nach einem lokalen Maximum, das dann als eigenständige Region gilt. Verschieben sich zwei Regionen übereinander wird die Region mit der höheren Popularität bevorzugt. Durch dieses Verfahren sind wir imstande, den Raum abhängig von den Eingangsdaten zu zerlegen, um somit möglichst aussagekräftige Orte zu finden.

Tabelle 6.2: Raumzerlegungsverfahren

suchung der Standpunkte der Nutzer festgelegt haben, aus den dort erwähnten Gründen.

Nach der Festlegung der Gegenstandsmenge muss man sich nun mit der Frage beschäftigen, wie hoch eine Bewertung durch einen Nutzer für eine bestimmte Region liegt. Im einfachsten Fall wertet man den Aufenthalt in einer Region. Dabei ist es belanglos, wie viele Bilder ein Nutzer in einer Region hinterlassen hat. Diese Art der Bewertung entspricht einer binären Skala, die für den Anwendungsfall durchaus ausreichend ist. Dennoch könnte man sich noch andere Bewertungsskalen überlegen, da der Nutzer durch die unterschiedliche Anzahl von Bildern in den unterschiedlichen Regionen auch eine Rangfolge festlegt, die ähnlich einer Punkteskala von „weniger interessant“ bis „sehr interessant“ interpretiert werden kann. Diese Rangfolge hat dann wiederum dieselbe Auswirkung auf die Nutzerähnlichkeit wie die Vergabe von unterschiedlichen Bewertungen. Dennoch soll uns zunächst eine binäre Skala genügen, da wir uns bei der Berechnung der Nutzerähnlichkeiten auf einen anderen Faktor konzentrieren wollen: Dem persönlichen Gehalt einer Entscheidung im Vergleich zu prototypischen Entscheidungen.

Die Bewertung einer Region kann auch als Umsetzung der Prototypentheorie nach Rosch betrachtet werden. Jeder Benutzer hat eine andere Vorstellung davon, welche Ansichten eines bestimmten geografischen Gebietes typisch sind. Diese Ansichten hat er in Bildern festgehalten. All diese prototypischen Ansichten werden, wie in Kapitel 5 beschrieben, zu einem prototypischen Modell des Ortes vereint. Die Auflistung nach Popularität eines Ortes ergibt eine typische Konzeptualisierung des Ortes. Um nun eine persönliche Empfehlung zu generieren, berechnen wir dieses Konzept im Folgenden nicht mehr auf der gesamten Nutzergemeinschaft, sondern nur noch auf der Gruppe von Nutzern, die sich zuvor am ähnlichsten verhalten haben. In Bezug auf Rosch sind das diejenigen Nutzer, die in der Vergangenheit ähnliche proto-

typische Vorstellungen eines Ortes aufgewiesen haben. Im Folgenden werden mögliche Metriken erarbeitet, die die Ähnlichkeit zwischen den Prototypen von zwei verschiedenen Benutzern bestimmen.

6.3.2 Schritt 2: Ähnlichkeitsberechnung mit Hilfe von räumlichen Entscheidungen

In einem bestimmten Gebiet erstellt ein Nutzer im Normalfall eine ganze Reihe von Bildern und teilt dadurch seine räumlichen Entscheidungen mit. Für jede Region definieren wir diese Informationen als geografisches Multi-Objekt-Tag. Der geografische Fußabdruck dieses Multi-Objekt-Tags besteht dabei aus dem Polygon, das diese Region umschließt. Jedes Objekt in der Menge stellt eine räumliche Entscheidung dar. Alle räumlichen Einzelentscheidungen liegen innerhalb dieses Raumpolygons; dieses verhält sich ähnlich wie das Schlagwort beim Mehr-Objekt-Tagging das Schlagwort, das alle Objekte textlich beschreibt. Die räumlichen Entscheidungen eines Nutzers lassen sich allgemein gesehen als eine Menge

$$A = a_1, a_2, \dots, a_m \quad (6.7)$$

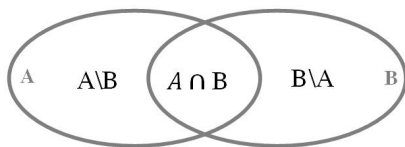
darstellen, wobei a_i eine räumliche Entscheidung eines Benutzers repräsentiert. Es gilt $a = \text{geoTagging}(\text{Bild}, \text{GPSPosition}, \text{Benutzer})$.

Definition 6.1 *Ein geografisches Multi-Objekt-Tag ist ein Tripel aus einer Menge von verorteten Bildern $A = a_1, \dots, a_n$. Nach Definition kann das verortete Bild als eine Relation $a = \text{geoTagging}(\text{Bild}, \text{GPSPosition}, \text{Benutzer})$ dargestellt werden. Insgesamt lässt sich das geografisches Multi-Objekt-Tagging als Relation $\text{geoMultiTagging}(A, \text{Polygon}, \text{Benutzer})$ schreiben.*

Die Liste der räumlichen Entscheidungen kann nun mit den Entscheidungen eines anderen Benutzers verglichen werden, welcher einen anderen Entscheidungsvektor

$$B = b_1, b_2, \dots b_m \quad (6.8)$$

für dasselbe Raumpolygon festlegt. Die Länge der beiden Vektoren muss dabei nicht zwangsläufig gleich lang sein. Vergleicht man die beiden Mengen miteinander, können drei Fälle auftreten, die in der folgenden Abbildung illustriert werden. Es werden zwei Teilmengen unterschieden: Entscheidungen, die von beiden Nutzern gemeinsam getroffen wurden und Entscheidungen, die lediglich von jeweils einem der beiden Nutzer alleine getroffen wurden. Abbildung 6.3.2 gibt einen Überblick über diese Schnittmengen.



Es gibt viele Ähnlichkeitsmaße, die auf geografischen Entscheidungsmengen arbeiten und das wahrscheinlich bekannteste davon ist eine Variante des Tverskymaßes, das von Rodrigues und Egenhofer veröffentlicht wurde. Sie benutzten dabei die Bedingung, dass die Parameter α und β in der Summe 1 ergeben.

$$\text{sim}(A, B) = \frac{|A \cap B|}{|A \cap B| + \alpha |A \setminus B| + \beta |B \setminus A|} \quad (6.9)$$

Dieses Maß erlaubt die Asymmetrie bei Ähnlichkeitsbeziehungen, die man oft bei kognitiven Entscheidungsprozessen beobachten kann. Obwohl diese Asymmetrie eine wichtige Eigenschaft dieses Maßes ist, verwenden wir eine symmetrische Variante des Tverskymaßes, um auf ein anderes Problem genauer einzugehen, das gelöst werden muss, um

die Ähnlichkeit in den vorliegenden Daten zu berechnen. Setzt man $\alpha = \beta = 1$ vereinfacht sich das Tverskymaß auf ein Maß, das man manchmal unter dem Begriff Tanimotomaß findet:

$$\text{sim}(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (6.10)$$

Um die angesprochene Problematik zu illustrieren, nehmen wir zwei hypothetische Extremfälle von Entscheidungen an, die von zwei Besuchern getroffen wurden. In beiden Fällen ist die Menge von Plätzen, die zumindest von A oder B besucht wurden, dieselbe. Die Menge enthält drei der populärsten Orte in Bamberg (Altes Rathaus, Dom und Klein Venedig). Zusätzlich sind drei recht unpopuläre Plätze enthalten, die sehr tief in der Verteilung zu finden sind. Die CPVs sind an Position 22, 23 und 24 nach ihrer Popularität eingeordnet worden.

Obwohl die Menge der Entscheidungen gleich ist, unterscheiden sich die beiden Fälle in der Auswahl der gemeinsamen CPVs. Im ersten Fall haben die beiden Nutzer die drei populärsten Plätze gemeinsam und im zweiten Fall teilen sie die drei unpopulären Entscheidungen miteinander. Nimmt man den einfachen Tanimoto-Ansatz, ergibt sich in beiden Fällen die gleiche Ähnlichkeit von $\frac{3}{6} = 0,5$. Jede Ähnlichkeitsberechnung, die nur die Anzahl der Entscheidungsmengen berücksichtigt bzw. jede Entscheidung gleich gewichtet, wird die gleiche Problematik aufweisen.

Um die berechnete Ähnlichkeit weiter zu differenzieren, sollte folgende Vorüberlegung zum Tragen kommen: Einige Entscheidungen werden häufiger gefällt und haben daher bereits eine breite Zustimmung in der Menge der Benutzer. Es ist also nicht überraschend, wenn zwei Personen diese Entscheidungen gemeinsam haben. Andere Entscheidungen hingegen werden seltener getroffen und spiegeln daher eher den persönlichen Geschmack der jeweiligen Person wieder. Dies ist eine mögliche

Umsetzung der etwas allgemeineren Idee, die im Jahr 1977 von Tversky als diagnostischer Wert einer Vergleichseinheit beschrieben wurde. Wir haben beobachtet, dass die Verteilung der Popularität an verschiedenen räumlichen Positionen einen sehr ausgeprägten Charakter hat – sie ähnelt einer Potenzreihe. Im Folgenden werden wir zwei mögliche Ansätze beschreiben, um den diagnostischen Wert einer Entscheidung zu bestimmen. Beide Ansätze basieren auf dem Popularitätsmaß. Wichtiger als der absolute Wert der Popularität ist allerdings der Anteil an der Gesamtpopularität. Der Wert $rp(c)$ eines CPVs gibt demnach die Popularität relativ zur Gesamtsumme an. Für die populärste Stelle in Bamberg, dem alten Rathaus, ergibt sich zum Beispiel ein Wert von 12,8%. Mehr als jeder zehnte Besucher erstellt demnach ein Bild von dieser Position aus. Umso höher der Wert $rp(c)$ liegt, umso weniger soll diese Entscheidung zur Ähnlichkeit beitragen. Die erste Annahme ist daher, dass der Betrag inverse proportional zu sein muss. Aus dieser Überlegung ergibt sich das Ähnlichkeitsmaß sim_p für die Nutzerähnlichkeit von Nutzer A und B, das in Formel 6.12 beschrieben wird. Ähnliche Überlegungen finden sich auch in der Arbeit von Breese u. a. (1998); diese versuchen den aus der Informationsfilterung verbreiteten Ansatz von „term frequency – inverse document frequency“ (tf-idf) im Bereich der Empfehlungen zu etablieren. Statt Häufigkeiten verwendet das Ähnlichkeitsmaß allerdings die Popularität des Ortes.

$$rp(n) = \frac{p(n)}{\max_{s \in siblings(n)} p(s)} \quad (6.11)$$

$$sim_p(A, B) = \frac{\sum_{c \in A \cap B} rp(c)^{-1}}{\sum_{c \in A \cup B} rp(c)^{-1}} \quad (6.12)$$

Dieses Maß erlaubt es, zwischen den beiden Extrembeispielen zu unterscheiden. Im Beispiel, in dem sich A und B über die populären Plätze

einig sind, ergibt sich eine Ähnlichkeit von $\text{sim}_p(A, B) = 0,15$. Für das zweite Beispiel erhält man eine Ähnlichkeit von $\text{sim}_p(A, B) = 0,84$.

Die Inverse der Popularität steigt sehr rasch an, da der Wert von $rp(c)$ auf Grund der Potenzverteilung sehr schnell gegen null sinkt. Daher ist es sinnvoll die Überbewertung der niedrigen Ränge bei der Ähnlichkeitsbetrachtung etwas abzuschwächen. Als alternativer Ansatz wird der Informationsgehalt der Entscheidung betrachtet, was einen niedrigeren Anstieg des berechneten diagnostischen Wertes verspricht. Der Begriff Informationsgehalt kommt aus der Informationstheorie und wird auch Überraschungswert genannt. Meistens findet man ihn in Bezug auf das Komprimieren von Nachrichten und er gibt an, wie viel Bit mindestens benötigt werden, um eine bestimmte Information zu übertragen. Wir bestimmen den Informationsgehalt eines CPVs c mit Hilfe des negativen Logarithmus der relativen Popularität $-\log_2 rp(c)$. Das alte Rathaus in Bamberg weist einen Informationsgehalt von 2,96 Bit auf, während eine Entscheidung an dem Ort mit Rang 24 informativer ist und 7,15 Bit benötigt. Verwenden wir den Informationsgehalt als Gewichtung, ergibt sich das Ähnlichkeitsmaß sim_i :

$$\text{sim}_i(A, B) = \frac{\sum_{c \in A \cap B} -\log_2 rp(c)}{\sum_{c \in A \cup B} -\log_2 rp(c)} \quad (6.13)$$

Wenden wir dieses Maß auf unsere beiden Beispiele an, können wir wieder beide Fälle unterscheiden. Für den ersten Fall ergibt sich eine Ähnlichkeit $\text{sim}_i(A, B) = 0,38$ und für den zweiten Fall $\text{sim}_i(A, B) = 0,62$.

An diesem Wert lässt sich beobachten, dass $\text{sim}_p(A, B)$ die Unterschiede der beiden Fälle viel deutlicher betont als $\text{sim}_i(A, B)$. Der Unterschied liegt mit dem ersten Maß bei über 500%, während sich die Werte

unter Verwendung des Informationstheoretischen Maßes nur um 61% unterscheiden.

Betrachtet man nun die n -nächsten Nachbarn eines Nutzers spricht man von seiner Nutzer-Gemeinschaft. Es ist zu beobachten, dass die berechneten Nachbarschaften nicht symmetrisch sind (Sarwar u. a., 2000), auch wenn unser Ähnlichkeitsmaß durch die Verwendung des Tanimoto-Ansatzes symmetrisch ist. Das heißt, ein Benutzer A kann einen Nutzer B in seiner Gemeinschaft der n -nächsten Nachbarn vorfinden, umgekehrt ist dies nicht zwangsläufig der Fall: A ist nicht automatisch in der Gemeinschaft von B. Abbildung 6.4 zeigt ein Beispiel, das für zwei Nutzer aus Panoramio berechnet wurde.

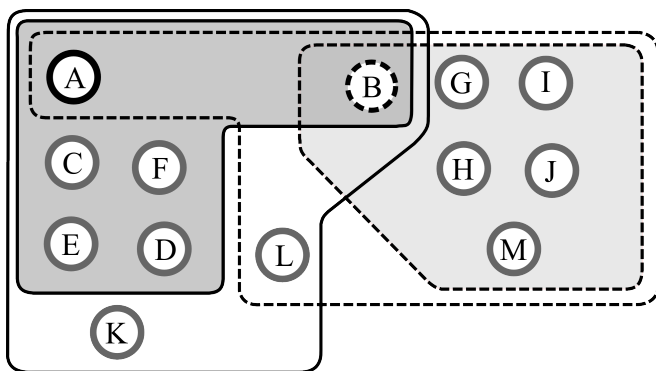


Abbildung 6.4: Beispiel einer unsymmetrischer Nutzergemeinschaften

In diesem Beispiel liegen die ersten fünf nächsten Nachbarn in den gefüllten Flächen, in Grün ausgehend von Nutzer A und in roter Farbe ausgehend von Nutzer B. Die Nutzergemeinschaft, bestehend aus den fünf nächsten Nachbarn für A beinhaltet den Nutzer B. Nutzer A ist allerdings nicht in den nächsten fünf Nachbarn von Nutzer B enthalten. Erst wenn man die nächsten Nachbarn um drei weitere Nutzer erweitert, liegen beide Nutzer in der jeweils anderen Nachbarschaft.

Explizite räumliche Semantik

Wie schon bei der Definition von Mehr-Objekt-Tagging werden Objektgruppen unter einem Konzept vereint. Im geografischen Fall, dem Mehr-Objekt-Geotagging, ist dieses Konzept über bestimmte räumliche Grenzen definiert. Mit deren Hilfe lässt sich feststellen, ob eine räumliche Entscheidung für das jeweilige räumliche Konzept zulässig ist. Bilder, die in Paris verortet sind, können daher niemals eine bildliche Konzeption des räumlichen Konzepts „Bamberg“ sein. Ähnlich wie bei Tags kommen die Stärken einer Folksonomy dann zum Tragen, wenn man mit einem gemeinsamen Vokabular arbeitet. Im sprachlichen Fall ist diese Zuordnung in gemeinsame Konzepte nicht ganz einfach: angenommen der erste Benutzer tagged ein Bild von sich mit dem Wort „München“, der Tag des zweiten Benutzers verwendet allerdings eine andere Sprache und würde sein Bild mit „Munich“ taggen. In so einem Fall ist es nicht mehr trivial festzulegen, ob die beiden dasselbe Konzept meinen. Hinzu kommt, dass Tags in der Regel mit vielen anderen sprachlichen Problemen, wie Schreibfehlern oder zusammengesetzten Wörtern übersät sind, was es noch schwieriger macht, gleiche oder sehr ähnliche Konzepte als solche zu erkennen.

Im räumlichen Fall sind die semantischen Sinneinheiten die geografischen Fußabdrücke. Für eine Menge von Entscheidungen wird ein Polygon als Fußabdruck verwendet, das alle Einzelentscheidungen umschließt. Im Fall von München kann dies zum Beispiel das Polygon sein, das die Stadtgrenze beschreibt und somit das Konzept München geografisch beschreibt. Ein Objekt wird zu diesem geografischen Konzept gezählt, falls es innerhalb dieses Polygons liegt. Dies funktioniert selbstverständlich nur, wenn das zugrundeliegende Konzept eine räumliche Ausdehnung bzw. einen festen Ort hat. Allgemeine Konzepte, wie „Wald“ oder „Flughafen“ lassen sich schlecht in ein konkretes Gebiet fassen. Auch umgangssprachliche Konzeptionen wie „Die Fränkische

Schweiz“ oder „Der Wilde Westen“ haben sehr ungenaue Grenzen. „Die Fränkische Schweiz“ zum Beispiel wird im Allgemeinen festgelegt als Gebiet zwischen den drei Städten Bamberg, Bayreuth und Nürnberg. Eine genaue Umrisslinie kann dabei allerdings nicht bestimmt werden. Diese Arbeit beschränkt sich auf die Verwendung von Toponymen mit scharfen Kanten, die durch einfache Polygone begrenzt sind. Der Umgang mit vagen Gebieten bleibt daher eine offene Fragestellung. Eine mögliche Modellierung dieser vagen Gebiete wird später noch aufgegriffen.

Der erste Punkt, den man bei der Festlegung dieser Grenzen beachten sollte, ist, dass die Popularität abhängig vom betrachteten Raum ist. Dies gilt insbesondere auch für die relativen Popularitäten, die in den Ähnlichkeitsmaßen in den Gleichungen 6.13 und 6.12 zu finden sind. Die Popularitätsgewichtung legt für ein bestimmtes Gebiet den semantischen Kern fest. Erweitern wir allerdings das betrachtete Gebiet, vergrößert sich mit hoher Wahrscheinlichkeit auch die Gesamtsumme und dadurch sinkt die Signifikanz der einzelnen Entscheidungen. Dabei können aber auch sehr populäre Orte hinzukommen, die zuvor nicht in der Menge der CPVs enthalten waren und diese können Entscheidungen aus dem semantischen Kern verdrängen. Hinzu kommt, dass es schwierig ist, die Unterschiede von zwei Menschen in einer einzigen Stadt zu bestimmen, wenn man dabei Entscheidungen im ganzen Land in Betracht zieht. Nimmt man zum Beispiel an, dass beide Benutzer 500 räumliche Entscheidungen, verteilt über ganz Deutschland, getroffen haben und 10 davon in einer Stadt liegen. Wir nehmen weiterhin an, dass die zehn Entscheidungen in dieser Stadt deckungsgleich sind, dann würde sich eine Ähnlichkeit nach Tanimoto von $\frac{10}{980} \approx 0,01$ ergeben. Betrachtet man allerdings nur das Stadtgebiet für sich alleine, würde man eine maximale Ähnlichkeit feststellen und alle anderen Regionen würden als nicht entscheidbar zählen. Eine sinnige Bedingung für die Auftei-

lung in einzelne Gebiete ist, dass dadurch keine Orte getrennt werden, an denen Benutzer im Normalfall zusammengehörige Entscheidungen treffen. Vor allem für einen touristischen Anwendungsfall macht es daher Sinn, die Stadt (vorwiegend im Hinblick auf die wachsende Zahl der Städtetouristen, siehe Landgrebe (2005)) als grundlegendes Raumobjekt zu verwenden. Unterteilungen dieser Art existieren bereits als fertige Produkte, zum Beispiel dem Nomenclature of Territorial Units for Statistics (NUTS), welche als Hintergrundwissen für das vorgestellte Verfahren verwendet werden kann.

Die zweite Beobachtung ist die bessere Modellierung des Nutzerverhaltens durch eine sinnvolle Aufteilung des Raumes. Dies ist vor allem dann interessant, wenn man unterschiedliche Ebenen von räumlichen Unterteilungen unterscheidet. NUTS stellt zum Beispiel bereits eine hierarchische Partitionierung des Raumes bereit, wobei jede Ebene eine Tessellierung darstellt. Eine Tessellierung bedeckt den kompletten Raum, ohne dass sich die einzelnen Teilgebiete überlappen. Es liegt eine komplette Partitionierung des Raumes vor. Der Vorteil eine Tessellierung zu verwenden ist, dass alle möglichen räumlichen Entscheidungen der Nutzer für die Ähnlichkeitsberechnungen auch berücksichtigt werden. Bei einer Tessellierung wird zu dem jede Entscheidung auch genau nur ein einziges Mal für einen Ähnlichkeitsvergleich verwendet. Neben der Tessellierung sind auch andere Raumunterteilungen vorstellbar. Man würde allerdings auf eine oder mehrere dieser Eigenschaften verzichten müssen. Bei einer Patchworkunterteilung könnten zum Beispiel einige Entscheidungen nicht berücksichtigt werden, da sie außerhalb eines gültigen Gebietes lägen, was zu einer Fehleinschätzung der Ähnlichkeit führen kann. Diese Raumunterteilung macht zum Beispiel Sinn, wenn sich die Ähnlichkeiten der Benutzer nur auf spezielle Gebiete stützen soll. Ein Anwendungsfall wäre die Analyse von Städtetouristen, die sich schon im Voraus auf die Stadtgebiete beschränken

bzw. Entscheidungen außerhalb des Stadtgebiets unerheblich für den Vergleich ihres räumlichen Verhaltens sind.

Ließe man Überlappungen zu, würde man die Entscheidungen, die dort getroffen worden sind, überbewerten. In diesem Fall spricht man von einer Coverage. Dieser Fall tritt dann auf, wenn man sich mit umgangssprachlichen Gebieten beschäftigt, die keine festen Grenzen haben. Ein Beispiel ist das vage Gebiet der Fränkischen Schweiz, das auch die Stadtbereiche von Bamberg, Bayreuth und Nürnberg mit enthalten könnte.

Abschließend kann man sagen, dass eine Tessellierung wohl die sinnvollste Raumzerlegung darstellt, obwohl es für das Verfahren letztendlich unerheblich ist, welche verwendet wird. Abbildung 6.5 gibt einen Überblick über drei mögliche Raumzerteilungen (Tessellierung, Patchwork, Coverage) zusätzlich zu den daraus resultierenden Entscheidungsmengen, die bei der jeweiligen Raumzerlegung beachtet werden (graue Flächen). Mögliche Raumzerlegungsmuster (a)-(c) und die zugehörigen Entscheidungsmengen (d)-(f) (graue Fläche).

Um die Möglichkeit zu haben auf verschiedenen Detailgraden zu argumentieren, wird im Folgenden von einer hierarchischen räumlichen Zerlegung ausgegangen. Dadurch können räumliche Entscheidungen besser interpretiert werden. Entscheidungen auf unterster Ebene können so zu einer räumlichen Entscheidung auf einem höheren Level generalisiert werden. Dabei wird zum Beispiel aus einer Entscheidung für eine bestimmte Position, eine Entscheidung für eine bestimmte Stadt. Ein Bild vom „Alten Rathaus“ in Bamberg ist somit auch eine Entscheidung für die Stadt Bamberg und die Region Franken und so weiter.

Die hierarchische Zerlegung kann als Graph modelliert. Ein Graph besteht aus dem Tupel zweier Mengen, wobei $N = n_1, \dots, n_m$ die Menge der Knoten des Graphen ist und $E = e_1, \dots, e_n$ die Menge aller Kanten im Graphen G . Jeder Knoten im Graphen stellt dabei eine Re-

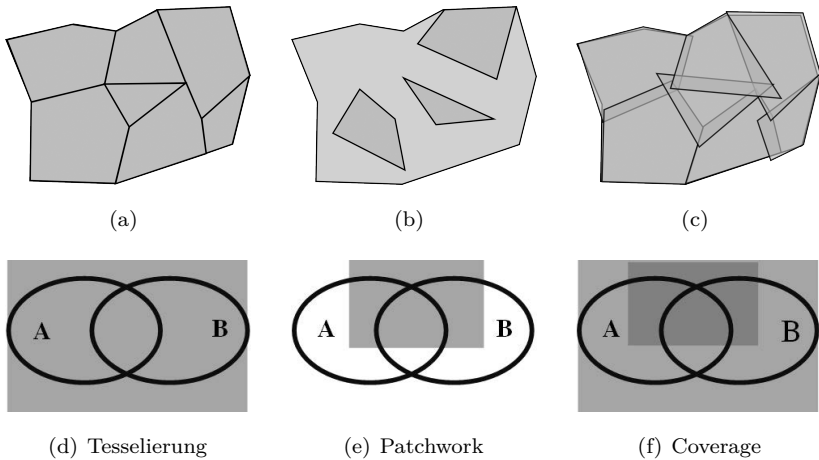


Abbildung 6.5: Raumzerlegungen und mögliche Entscheidungsmengen

gion dar, das beinhaltet ein Konzept und einen geografischen Fußabdruck. Im geografischen Fall lässt sich das Konzept durch den Namen der Region beschreiben, die geografisch durch ein Polygon als Fußabdruck beschrieben wird. Eine Kante im Graphen repräsentiert eine Teil-von-Beziehung zwischen zwei Regionen, im Fall einer Tessellierung oder einer Patchwork-Zerlegung würde sogar eine räumliche enthalten-in-Beziehung vorliegen. Das heißt die Beziehungen zwischen den Knoten sind sogar berechenbar. Sinnvoll für den touristischen Anwendungsfall ist eine Zerlegung in verschiedene Stadtbereiche auf der untersten Ebene. Im hierarchischen Fall hat der Graph eine Richtung und beginnt an einem Wurzelknoten, der die komplette Fläche der Erde umschließt. Es folgen verschiedene Ebenen, die eine bestimmte Granularität in der Zerlegung gemeinsam haben sollten. Knoten aus einem der oberen Ebenen können nur mit Knoten aus einer tieferen Ebene durch eine Kante miteinander verbunden sein. Diese Knoten gehen eine Vater-Kind-Beziehung ein. In dieser Beziehungsmetapher definieren

sich Geschwisterknoten, als Knoten mit einem gleichen Elternknoten. Knoten, die keine Kinder haben, werden Blattknoten genannt.

Um die unterschiedlichen räumlichen Entscheidungen zu repräsentieren, muss eine weitere Ebene in den Graphen eingefügt werden, die den räumlichen Clustern in den jeweiligen Gebieten entsprechen. Diese Arbeit verwendet dafür die Regionen, die der Heatmapper für jede dieser Gebiete berechnet hat, den CPVs. Ein CPV wird ebenfalls als Knoten im Graphen repräsentiert und der geografische Fußabdruck entspricht der endgültigen Position des kreisförmigen Suchkreises, der für den Clusteralgorithmus verwendet wurde. Jeder CPV-Knoten beinhaltet alle Bilder, die in dieser Region enthalten ist. Jedes Bild wiederum könnte man ebenfalls als eigenständigen Knoten modellieren, was allerdings keinen Mehrwert für die verwendeten Maße, wie Popularität, bringen würde. Es reicht daher aus, entscheiden zu können, ob ein verortetes Bild in einem CPV-Knoten enthalten ist oder nicht. Der Urheber dieses Bildes gilt als Teilnehmer an dem Knoten. Man testet dazu, ob das Bild innerhalb des entsprechenden Fußabdrucks liegt. Für Knoten in den höheren Ebenen gilt ein Benutzer als Teilnehmer, wenn er mindestens an einem seiner Kinder teilgenommen hat. Ein Nutzer, der in Bamberg Bilder eingestellt hat, tut dies auch automatisch für Bayern und wird bei der Berechnung für Bayern mit beachtet.

Alternativ kann an Stelle von einem der beiden räumlichen Clusterverfahren, die in Abschnitt 5.4.1 definiert wurden, auch ein anderes Gruppierungsverfahren verwendet werden. Es sollte allerdings in der Lage sein, die Einzelbilder zu gleichbedeutenden räumlichen Entscheidungen zu gruppieren. Die daraus resultierenden Gruppen enthalten Bilder, die austauschbar sind. In unserem Fall sind es Bilder, die einen ausreichend ähnlichen Standort des Fotografen haben. Auch von Hand erzeugte Regionen könnten in Frage kommen, in die Bilder zugeordnet werden können. Abgesehen von der Analyse der POVs sind Verfahren

wie Photosync (Snaveley u. a., 2008) in der Lage auch Bildgruppen zu liefern, die das gleiche Objekt fotografiert haben, so lange es sich um Gebäude oder ähnliches handelt. Dieses Bildanalyseverfahren arbeitet mit Hilfe der Erkennung von SIFT-Punkten in den Bildern, um diese aufeinander abzubilden. Dabei verwendet man fotogrammetrische Verfahren, um zum einen das Objekt selbst zu rekonstruieren, aber auch, um den Standort der Kamera zu präzisieren. Auch Verfahren, die die Semantik der angefügten Tags untersuchen (Rattenbury und Naaman, 2009), versuchen, über den POV hinaus die abgebildeten Objekte zu analysieren. Diese Verfahren könnten in Zukunft rein geografische Clustern ergänzen, sobald ihre Ergebnisse durch die steigende Anzahl von verorteten Bilder ausreichend sind. Damit könnte zumindest innerhalb eines CPVs die Hauptansicht berechnet werden und die Auswahl für einen guten Stellvertreter für das gesamte Cluster verbessert werden. Egal welches Verfahren man einsetzt, am Ende wird eines der Bilder aus dem Cluster als Repräsentant für alle Bilder des Clusters ausgewählt. Wird ein Cluster durch das Empfehlungsverfahren ausgewählt, wird nur dieses eine Bild als Stellvertreter für alle anderen Aufnahmen empfohlen.

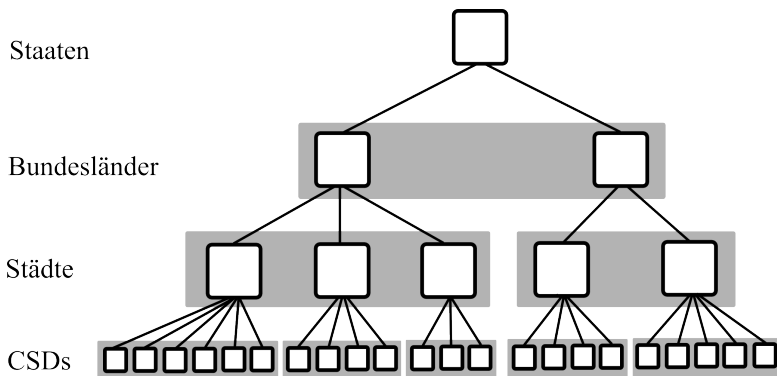


Abbildung 6.6: Partonomy

Für jeden Knoten im Graphen werden nun verschiedene Gewichte berechnet, die für die Berechnung der Ähnlichkeit zwischen zwei Benutzern benötigen werden. Zum einen ist das die Popularität des Gebietes und zum anderen, davon abgeleitet, der diagnostische Wert eines Knoten. Geht man von einer Tesselation aus, kann die Popularität rekursiv bestimmt werden. Dazu berechnet man zunächst die Popularität der Blattknoten (CPVs). Die Popularität der anderen Ebenen ergibt sich aus der Summe der Popularität der Kinder. Ansonsten muss die Popularität für jeden Knoten neu berechnet. Die Funktion $p(n)$ entspricht dem Popularitätsmaß aus Gleichung 5.1 angewendet auf die Menge aller Entscheidungen, die in dem jeweiligen Knoten enthalten sind.

$$popularity(n) = \sum_{c \in children(n)} popularity(c) \quad (6.14)$$

Das zweite Maß folgt den vorangegangenen Überlegungen über den diagnostischen Wert einer Entscheidung. Für einen bestimmten Knoten wird demnach der relative Informationsgehalt im Vergleich zu seinen Schwesterknoten (Knoten, die denselben Vater haben) berechnet. Zunächst berechnen wir für jeden Knoten den Informationsgehalt mit Hilfe der folgenden Formel:

$$information(n) = -\log_2 rp(n) \quad (6.15)$$

Die Gleichung ergibt das Informationsmaß. Übertragen gesehen gibt das Informationsmaß eines Knotens an, wie viel Information über den Benutzer selbst in der Entscheidung steckt, die er durch die Teilnahme an diesem Knoten vermittelt hat. Informationen meint in diesem Zusammenhang das Maß der persönlichen Entscheidung. Im vorliegenden Fall wird die relative Popularität in Bezug auf die Schwesternknoten als Auftretenswahrscheinlichkeit angenommen. Beim populärsten Knoten

wird davon ausgegangen, dass er mit der höchsten Wahrscheinlichkeit von 1 vorkommt, während alle anderen Ränge immer unwahrscheinlicher in einer Menge von Entscheidungen auftreten werden bzw. ein Nutzer sich an diesen Knoten mit seinen Bildern beteiligt. Es gilt also $g(n) = rp(n)$.

Das letztendliche Gewicht eines Knotens ergibt sich durch den Informationswert relative zum maximalen Informationsmaß, das ein Geschwisterknoten aufzeigen kann. Dieses Maß skaliert den Informationsgehalt eines Knotens lediglich zwischen 0 und 1.

$$w(n) = \frac{\text{information}(n)}{\max_{s \in \text{siblings}(n)} \text{information}(c)} \quad (6.16)$$

Auf Grundlage dieses Baumes lassen sich nun verschiedene Ähnlichkeitsmaße bestimmen, die die unterschiedliche Belegung von Nutzerentscheidungen innerhalb des Baumes vergleichen. Jedes Maß kann auf unterschiedlichen Ebenen des Baumes angesetzt werden, wobei man auf der untersten Ebene die differenziertesten Aussagen über Nutzerähnlichkeiten treffen kann.

$$\widehat{r_{u,i}} = \overline{R_u} + \frac{\sum_{u' \in N_u} (r_{u',i} - R_{u'}) \cdot \text{sim}(u, u')}{\sum_{u' \in N_u} |\text{sim}(u, u')|} \quad (6.17)$$

Räumliche Ähnlichkeitsmaße

Durch die zusätzlichen Ebenen lassen sich auch Aussagen über die Ähnlichkeiten der Benutzer bestimmen, auch wenn sie auf der untersten Ebene keine gemeinsamen Entscheidungen aufzeigen. Ein Beispiel für einen solchen Fall sind zwei Nutzer, die zwar dieselbe Stadt, aber nicht dieselben Plätze besucht haben oder dasselbe Land, aber nicht dieselben Städte. Man kann grundsätzlich sagen, dass zwei Nutzer sich am ähnlichsten sein sollen, wenn sie dasselbe europäische Land, in diesem Land dieselbe Region, in der Region dieselben Städte und in die-

sen Städten dieselbe räumliche Entscheidungen getroffen haben. Abbildung 6.7 gibt eine Übersicht über die verschiedenen Ansätze, um Entscheidungen auf verschiedenen Ebenen mit einfließen zu lassen.

- *Single-Layer feature similarity (SFS)*: Das erste Verfahren, um die Nutzerähnlichkeiten zu bestimmen, entspricht dem Standardansatz von Resnick, der einzelne Zeilen der Nutzer-Item-Matrix miteinander vergleicht. Allerdings werden anstatt der Gegenstände die räumlichen Entscheidungen betrachtet. Jede mögliche räumliche Entscheidung auf einer Ebene (bzw. jede Teilnahme an einem Knoten einer bestimmten Ebene) wird durch die Einträge in einem Vektor repräsentiert. Die Kardinalität dieses Vektors entspricht demnach der Anzahl aller möglichen Entscheidungen. Ein Eintrag in diesem Vektor entspricht der Bewertung dieses Ortes, die der Benutzer dort hinterlassen hat. Eine einfache Heuristik ist dabei die Anzahl der Bilder als Bewertung zu verwenden.

$$\text{sim}_{SFS}(u_a, u_b) = \frac{\vec{I}_{u_a} \cdot \vec{I}_{u_b}}{|\vec{I}_{u_a}| \cdot |\vec{I}_{u_b}|} \quad (6.18)$$

Die unterschiedlichen Nutzervektoren können nun mit Hilfe von Korrelationsfaktoren verglichen werden. In den meisten kollaborativen Filtern wird der Pearson-Korrelationsfaktor oder die Cosinusähnlichkeit verwendet. In den eigenen Testläufen wurde für dieses Verfahren die Cosinusähnlichkeit verwendet, da in dem Fall die Anzahl der an einer räumlichen Einzelentscheidung beteiligten Bilder eine nicht so große Rolle spielen soll, wie es beim Pearson-Korrelationsfaktor der Fall sein würde. Bei Cosinus wird lediglich der Winkel zwischen den Vektoren berechnet, wobei deren Länge keine Rolle spielt. In diesem Maß werden keine räumlichen Hintergrundinformationen mit verwendet, was die Hypothese zulässt, dass dieses Maß die ungenausten Ergebnisse liefert.

- *Two Layer feature similarity (TFS)*: Dieses Maß folgt der Idee, die Ähnlichkeit in unterschiedlichen Region einzeln zu berechnen und daraus für eine höhere Ebene eine Gesamtähnlichkeit zu entwickeln. Dazu wird wie im ersten Fall die Ähnlichkeit der Benutzer durch die Cosinusähnlichkeit berechnet. Allerdings für die Knoten eines gemeinsamen Vaterknotens einzeln, um danach die gewichtete Summe gemäß Gleichung 6.19 zu berechnen. Die Ähnlichkeit der beiden Benutzer skaliert dabei den Wert eines Knotens $w(n)$ und die Summe wird dadurch auf einer Skala von 0 bis 1 skaliert.

$$sim_{TFS}(u_a, u_b) = \frac{\sum_{n \in cities} sim_{SFS}(u_a, u_b, n) \cdot w(n)}{\sum_{n \in cities} w(n)} \quad (6.19)$$

Man kann diesen Ansatz von einer beliebigen Ebene aus starten und von jedem Knoten aus die Ähnlichkeit in den Kindern berechnen. In der Evaluation werden wir dieses Maß verwenden, um die Ähnlichkeit im Hinblick auf die Ebene der Städte zu berechnen.

- *Two-Layer information similarity (TIS)*: Dieses Maß verfolgt grundsätzlich denselben Ansatz wie das Two-layer feature Maß. Allerdings wird in diesem Fall die Ähnlichkeit auf der unteren Ebene anders berechnet. Bei diesem Maß sollen alle Vorüberlegungen über den diagnostischen Wert einer geografisch räumlichen Entscheidung zum Tragen kommen. Zunächst wird die Ähnlichkeit der Entscheidungen auf den Geschwisterknoten, ausgehend von den Knoten einer bestimmten Ebene, berechnet. Will man die Ähnlichkeit auf Stadtebene ausrechnen, so werden jeweils die Entscheidungen der einzelnen Benutzer in den CPV-Knoten in jeder Stadt miteinander verglichen und das Maß als Ähnlichkeit für diese Stadt berechnet. Die Berechnung verwendet den Informationsgehalt der Knoten um die Ähnlichkeit nach dem Tanimotomaß zu berechnen. Die Gesamtähnlichkeit der Benutzer auf allen gemeinsamen Knoten berechnet sich dann wiederum aus

der folgenden Gleichung, wobei dieses Mal sim_i verwendet wird anstatt sim_{TFS} .

$$sim_i(u_a, u_b, n) = \frac{\sum_{c \in A_n \cap B_n} information(c)}{\sum_{c \in A_n \cup B_n} information(c)} \quad (6.20)$$

$$sim_{TIS}(u_a, u_b) = \frac{\sum_{n \in cities} sim_i(u_a, u_b, n) \cdot w(n)}{\sum_{n \in cities} w(n)} \quad (6.21)$$

Wie auch schon beim Two-layer-feature-similarity-Ansatzes wird im Folgenden, dieses Maß auf die Ebene der Städte angewendet. Andere Einstiegsebenen sind aber ebenso möglich, wobei das Maß dann an Aussagekraft verliert, da örtliche Unterschiede verloren gehen.

- **Geographic coverage similarity (GCS):** Bei den bisher vorgestellten Maßen wurden jeweils nur eine oder zwei Ebenen für die Berechnung der Ähnlichkeit betrachtet. Dieser Ansatz deckt alle Ebenen des Zerlegungsbaumes ab und ist daher immer in der Lage eine Ähnlichkeit zwischen zwei Benutzern festzustellen. Zumindest, wenn man davon ausgeht, dass alle Bilder im Wurzelknoten, also der Erde, enthalten sind, sind die Teilmengen, die diejenigen Knoten enthalten, an denen entweder A oder B teilnehmen. Die Ähnlichkeit bestimmt sich aus der Summe der Knotengewichte, ausgehend von Knoten, an denen beide Nutzer teilgenommen haben, geteilt durch die Summe aller Knotengewichte, ausgehend von Knoten, an denen mindestens ein Nutzer teilgenommen hat.

$$sim_{GCS}(u_a, u_b) = \frac{\sum_{n \in N_{u_a} \cap N_{u_b}} w(n)}{\sum_{n \in N_{u_a} \cup N_{u_b}} w(n)} \quad (6.22)$$

Jedes dieser Maße wertet die einzelnen Entscheidungen mit einem anderen Gewicht. Welches Maß am besten geeignet ist, relevante Nutzergruppen zu finden, die sich gegenseitig gute Empfehlungen aussprechen können, wird die Evaluation in Kapitel 8 zeigen. TIS und GCS sind in

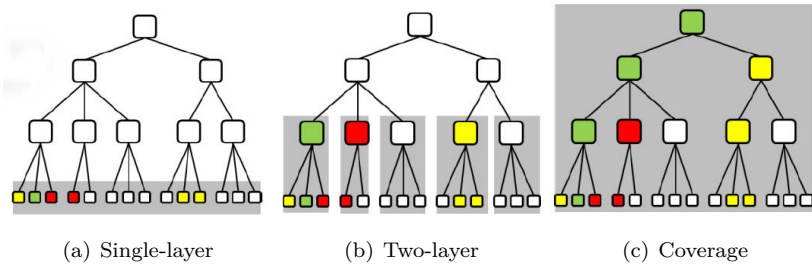


Abbildung 6.7: Ähnlichkeitsmaße und die unterschiedlichen Ebenen, die sie ansprechen. Knoten mit grüner Farbe sind geografische Entscheidungen, die beide Nutzer gemeinsam getroffen haben. Rote und gelbe Knoten zeigen Entscheidungen, die nur von einem einzigen Nutzer gewählt wurden.

der Lage die Annahmen über ein gutes geografisches Ähnlichkeitsmaß aus Abschnitt 6.3.2 zu erfüllen und versprechen somit die besten personalisierten ortsbezogenen Empfehlungen.

6.3.3 Schritt 3: Nachbarschaftswahl mit Hilfe von räumlichen Entscheidungen

Wie schon in Abschnitt 6.17 erwähnt, lässt sich die Auswahl der geeigneten Nutzer über das verwendete Ähnlichkeitsmaß bestimmen. Im geografischen Fall können allerdings von vornherein einige Benutzer ausgeschlossen werden, die für die Berechnung einer Empfehlung nicht relevant sind. In den meisten Fällen sind das Nutzer, die trivialerweise keine Ähnlichkeit zu dem momentanen Nutzer aufweisen können. Man kann daher folgende Filterungsregel für die Suche nach geeigneten Kandidaten festlegen:

- Abhängig vom aktiven Nutzer identifiziert man die Gebiete, in denen er bereits eigene Eindrücke hinterlassen hat. Ein Kandidat muss ebenfalls mindestens in einem dieser Gebiete räumliche Entscheidungen getroffen haben, ansonsten ist seine Ähnlichkeit im Fall von SFS, TFS und TIS gleich 0. Dieses Ausschlussverfahren gilt nicht für das GCS, da dieses Maß immer in der Lage ist, eine Ähnlichkeit festzustellen.
- Zusätzlich muss ein möglicher Kandidat in dem gewählten Zielgebiet bereits räumliche Entscheidungen getroffen haben. Nur ein Kandidat, der auch im Zielgebiet beteiligt ist, kann seine Erfahrungen dort auch einbringen.

Nachdem die Ähnlichkeiten der übrigen Kandidaten berechnet wurden, kann nun eines der Standardauswahlverfahren verwendet werden, um die nächsten Nachbarn zu wählen. Für die eigene Implementation wurde vorerst beschlossen, alle Nutzer zu verwenden, die eine Ähnlichkeit aufweisen. Weist ein Benutzer eine sehr geringe Ähnlichkeit auf, wird sein Anteil am Ergebnis selbst ebenfalls sehr gering sein, so dass diese Festlegung das Ergebnis nicht negativ beeinflussen wird.

6.3.4 Schritt 4: Aggregation der räumlichen Entscheidungen

Ziel ist es nun, mit Hilfe der nächsten Nachbarn die Plätze der gewünschten Region neu zu bewerten. Die Bewertung soll nicht mehr an der Popularität aller Nutzer gemessen werden. Die Definition eines Platzes entspricht weiterhin den berechneten CPVs aus Kapitel 5, allerdings wird ihre Popularität auf Null gesetzt und dann mit Hilfe der folgenden Funktion für einen Nutzer u und einen Knoten n neu berechnet.

$$w_{personal}(u, n) = \sum_{u_s \in U_{sim}} sim(u, u_s) \cdot \gamma(u_s, n) \quad (6.23)$$

Die binäre Funktion $\gamma(u, n)$ gibt an, ob ein Benutzer u eine räumliche Entscheidung am Ort n getroffen hat. Hat der Benutzer mindestens ein Bild an diesem Ort zur Verfügung gestellt, nimmt die Funktion den Wert 1 an, ansonsten ist ihr Wert 0. An Orten, an denen keiner der nächsten Nachbarn eine Entscheidung getroffen hat, ist keine persönliche Bewertung möglich und es gilt $w_{personal} = 0$.

Im Gegensatz zum allgemeinen Ansatz nach Resnick wird dieser Wert nicht mit der Summe aller Ähnlichkeiten skaliert. Dieser Schritt ist für unseren Anwendungsfall nicht wichtig, da es nicht darum geht eine exakte Bewertung vorherzusagen, sondern lediglich eine Ordnung auf der Menge der Plätze festzulegen.

Abbildung 6.8 und Abbildung 6.9 zeigen die Veränderung der persönlichen Gewichtung im Verlauf der Popularitätskurve. Die Cluster beider Diagramme sind nach Popularität $p(n)$ geordnet. Allerdings wurde im zweiten Diagramm eine persönliche Gewichtung berechnet. Es ist gut zu sehen, wie einige Orte im langen Schwanz der Verteilung viel höher gewichtet werden, als dies noch durch die typische Bewertung der Fall war. Ordnet man die Orte nun nach ihrem entsprechenden Popularitätsmaß neu, werden diese Orte bei einer Top- n Empfehlung viel wahrscheinlicher in die Ergebnismenge aufgenommen, als dies noch durch die typische Popularitätsverteilung der Fall war. Der persönliche Popularitätswert einiger Orte fällt mit hoher Wahrscheinlichkeit auf 0 ab und diese Orte werden daher ganz aus der möglichen Empfehlungsliste entfernt.

Das Auftreten von Popularitätsspitzen im langen Schwanz der Verteilung ist vor allem interessant, weil sich dort die Objekte befinden, die nur für wenige Benutzer eine Bedeutung haben. Chris Anderson beschrieb zum ersten Mal ausführlich dieses Phänomen, dass Kaufentscheidungen zwar eine große Übereinkunft bei den populärsten Objekten aufweisen, den sogenannten Bestsellern, es aber gleichzeitig eine

unglaubliche Vielzahl von unterschiedlichen Objekten gibt, die ganz wenig gekauft wurden. Dennoch ist die Anzahl der Gegenstände, die viel verkauft werden im Vergleich zu den wenigen Bestsellern um einiges höher, so dass sie multipliziert mit Anzahl der verkauften Exemplare mehr Verkaufsvolumen ergeben. Die Problematik bei dem Verkauf von Gegenständen aus dem langen Schwanz ist allerdings, dass sie nicht jedem angeboten werden können. Vor allem bei digitalen Verkaufswegen, bei denen keinerlei Lager- oder Versandkosten anfallen, bietet sich hier ein interessantes Geschäftsmodell, so wie auch schon der Untertitel des Buches es ausdrückt: „sell more of less“.

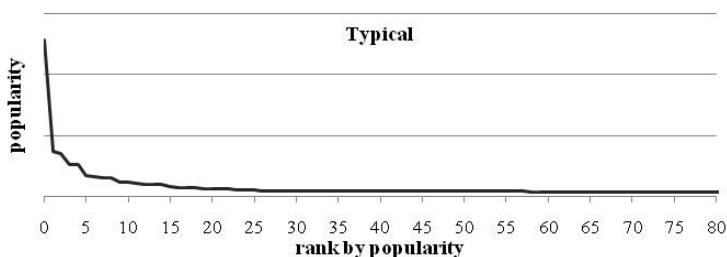


Abbildung 6.8: Heatmap sortiert nach Popularität

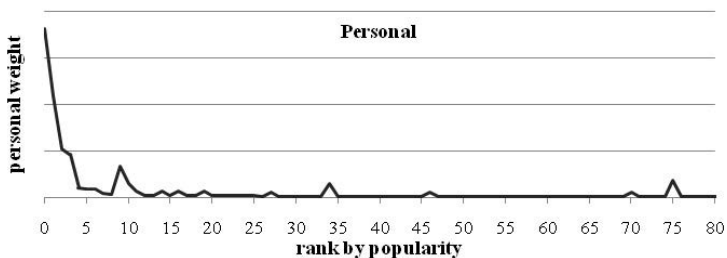


Abbildung 6.9: Gleiche Heatmap allerdings mit persönlicher Gewichtung

Die Plätze mit einer Bewertung von 0 entsprechen den Werten, die in herkömmlichen kollaborativen Empfehlungen durch das „new-item“-Problem entstehen können. Das sind unter anderem Orte, die nur ganz wenige Menschen überhaupt besucht haben. Eine Möglichkeit, sie dennoch in die Ergebnismenge mit aufzunehmen, ist die zusätzliche Gewichtung eines Knotens mit dessen Popularität. Dadurch wird jeder Ort über das Nullniveau gehoben (nach Definition von Popularität hat jeder Platz mindestens Popularität 1) und hat somit eine theoretische Chance, empfohlen zu werden.

7 Tripod GeoRecommender

7.1 Projekt Tripod

Teile der vorliegenden Doktorarbeit entstanden im Rahmen eines europäischen Projektes. Das durch die EU geförderte Forschungsprojekt Tripod hatte sich zum Ziel gesetzt, Verfahren zu entwickeln, mit denen sich Metainformationen über Bilder automatisch erzeugen lassen und Bilder einfacher kategorisieren und organisieren lassen. Von 2007 bis 2010 entstand in der Zusammenarbeit von 10 universitären und industriellen Partnern aus sieben europäischen Ländern mehrere Verfahren, um sich diesem Ziel zu nähern. Grundlage war, wie auch bei meiner Arbeit, Ortsinformationen, die man zusätzlich zum eigentlichen Bildmaterial auswerten konnte. Insbesondere die von heutigen Kameras erfassten GPS-Daten standen im letzten Projektabschnitt im Fokus der Forschungen. Durch die GPS-Informationen sind die Teilsysteme von Tripod in der Lage die Bilder mit Hintergrundinformation aus anderen georeferenzierten Quellen zu verbinden.

Hauptziel des Projekts war es, für den Menschen lesbare Bildtitel und Beschreibungen zu generieren. Zusätzlich sollten sinnvolle Schlüsselwörter erzeugt bzw. vorhandene ergänzt werden. Die Problemlage aus Anwendersicht lässt sich gut an dem Fratelli-Alinari-Archiv in Florenz, einem der Projektpartner, illustrieren. Das Archiv verwaltet einen großen Bestand von historischen Fotografien aus der Toskana und Italien, der fortlaufend durch aktuelle Pressefotografien ergänzt wird. Täglich müssen großen Menge an Fotomaterial erschlossen werden. Ohne

Unterstützung durch automatische Verfahren der Betitelung werden die Bildarchivare diese Aufgabe bald nicht mehr bewältigen können. Das Projekt hatte nicht das Ziel, eine Universallösung für mögliche Motive anzubieten, sondern beschränkte sich ausschließlich auf Bilder mit geografischem Bezug, das heißt hauptsächlich auf Bilder von Gebäuden oder Landschaften. Bildbasierte Verfahren wurden dabei bewusst vernachlässigt, um sich mehr auf einen reinen geografischen Ansatz zu konzentrieren. Dennoch konnte man auf ein paar einfache aber auch zuverlässige Verfahren nicht verzichten. Sie wurden hauptsächlich dazu verwendet, Bilder auszuschließen, die nicht dieser Einschränkung genügen. Mit Hilfe von Gesichtserkennungsverfahren konnten zum Beispiel Bilder ausgeschlossen werden, die hauptsächlich Menschen zeigen.

Die Verfahren wurden dabei in ein Rahmenwerk umgesetzt, das sich hauptsächlich aus Datenquellen, wie topographische Karten, Ortsnamenverzeichnisse, Landnutzungsdaten und anderen Internetquellen, bedient. Für das Projekt selbst wurden aber auch eigene Datenquellen wie beispielsweise eine passenden Konzeptontologie für derartige Bilder aufgebaut. Die einzelnen Systeme können dabei über Webservice-Schnittstellen miteinander kombiniert werden.

Die Bilder selbst mussten dabei nicht zwangsläufig über eine GPS-Position verfügen, ansonsten müsste man auch einen Großteil der bestehenden Archive als Bildquellen ausschließen. In der ersten Projektphase wurde zunächst nach Möglichkeiten geforscht, den Aufnahmeort auf bestehenden Bildern mit Hilfe ihres Bildtitels zu rekonstruieren. Diese Anforderung lieferten vor allem die digitalen Datenbestände des Fratelli-Alinari-Archivs, da dort viele historische Bildbestände in digitaler Form vorliegen, die zwar einen Bildtitel aufweisen, aber über keine GPS-Position verfügen. Erst in den neuen professionellen Fotokampagnen von Alinari werden die Bilder direkt mit GPS-Sensoren verortet. Bereits zu Beginn des Projektes musste man daher zwei unterschiedliche

Qualitätsstufen bei den Ortsinformationen eines Bildes unterscheiden, für die zum Teil eigene Lösungen angeboten werden mussten.

Szenario 1: Nur Bildtitel Ausgehend von einem Bild mit einem Bildtitel werden in diesem Satz nach Toponymen, wie Städtenamen oder Adressen, gesucht. Dabei kann es allerdings zu Doppeldeutigkeiten kommen, zum Beispiel findet man die meisten europäischen Großstädte unter gleichem Namen auch auf dem amerikanischen Kontinent. Aus diesem Grund war schon dieser erste Schritt keine triviale Aufgabe und es wurden verschiedene Verfahren verwendet, um diese Doppeldeutigkeiten aufzulösen und für jedes mögliche Toponym mit anzugeben, wie wahrscheinlich es ist, dass genau dieses Toponym gemeint ist. Taucht zum Beispiel das Toponym „Paris“ zusammen mit „Illinois“ eingeleitet mit dem Bindewort „in“ im Bildtitel auf, kann man davon ausgehen, dass nicht das ansonsten naheliegende Paris in Frankreich gemeint ist. Wird Paris allerdings nur durch das zusätzliche Toponym „USA“ eingegrenzt muss man sich zwischen sieben weiteren Städten mit dem Namen „Paris“ entscheiden. Am Ende legt sich das System auf ein Toponym fest und berechnet daraus ein vorläufiges Raumpolygon.

In einem zweiten Schritt wird im Bildtitel weiter nach räumlichen Beziehungswörtern gesucht, wie zum Beispiel „südlich von“ oder „in der Nähe von“. Je nachdem, welche räumliche Beziehung identifiziert wurde, wird das vorläufige Raumpolygon mit Hilfe von festgelegten Regeln weiter verfeinert, erweitert bzw. verschoben um ein endgültiges Polygon zu erzeugen. In diesem eingegrenzten Gebiet werden nun, in den vorhandenen Datenquellen, möglichen Informationen gesucht, die auf das Bild zutreffen könnten.

Wichtigster Schritt ist dabei die Verwendung der Konzeptontologie, die für Tripod implementiert wurde. Die Auswahl von Konzepten erfolgt dabei auf drei verschiedene Arten, mit Hilfe des Raumpolygons, dem Typus der Szene und des identifizierten Subjektes im Titel. Durch jedes

dieser Verfahren wird eine Liste von Elementen und Qualitäten berechnet. Tripod verwendet für die Konzeptontologie außerdem eine Konzepthierarchie, daher können mehrere Konzepte auch zu einem übergeordneten Konzept zusammengefasst werden. Zusätzlich lassen sich die Konzepte in andere Sprachen übersetzen, wenn die Ontologie in dieser Sprache vorliegt und die einzelnen Konzepte mit den entsprechenden Übersetzungen verlinkt wurden.

In einem letzten Schritt werden die bisherigen Schlüsselwörter durch Schlüsselwörter aus kollaborativen geografischen Bilddatenquellen ergänzt. Beispiele für derartige Bilddatenquellen findet man in Kapitel 3, wobei für diesen Schritt die Daten aus Panoramio verwendet wurden. Das Tripod-System gewichtet nun alle Schlüsselwörter nach ihrer Relevanz für den konkreten geografischen Kontext des Bildes. Die obersten n Begriffe werden dann in die Ergebnismenge übernommen.

Szenario 2: Bild mit Lagesensoren Sind bereits räumliche Sensordaten vorhanden, kann das Raumpolygon aus dem ersten Schritt viel detaillierter und präziser erzeugt werden. Je nachdem welche Lagesensoren zur Verfügung stehen, werden daraus der entsprechende Sichtbereich der Aufnahme berechnet. Bei der Rekonstruktion des Sichtbereichs wird davon ausgegangen, dass sich der Fotograf zum Zeitpunkt der Aufnahme auf dem Boden befunden hat.

Stehen nur zweidimensionale Hintergrunddaten zur Verfügung, lässt sich daraus auch nur ein zwei dimensionaler Sichtbarkeitsbereich erzeugen. Liegt nur eine GPS-Koordinate vor, verwendet das System einen kreisrunden Suchpuffer. Kann man zusätzlich noch auf eine Richtungsinformation zurückgreifen, wird dieser Kreis durch einen Kreiskegel ersetzt. Der entsprechende Sichtwinkel ergibt sich aus den Einstellungen der Optik und den Abmessungen des in der Kamera verbauten Aufnahmechips. Liegen zusätzlich zur Position der Kamera auch Richtungsdaten vor, kann ein dreidimensionaler Sichtkegel der Kamera be-

rechnet werden. Der Radius des Sichtkreises(-kegels) hängt davon ab, ob das Bild im Stadtbereich oder auf dem Land aufgenommen wurde, da man davon ausgehen kann, dass die Besiedlungsdichte in Städten dafür sorgt, dass die Sicht meistens schon nach wenigen Metern von einem Gebäude unterbrochen wird.

Die Sichtbarkeit innerhalb dieses zweidimensionalen Bereiches kann mit Hilfe von Geländemodellen weiter eingeschränkt werden. Diese Methode beruht darauf, dass der Schattenbereich hinter einer Anhöhe nicht mehr vom Betrachter gesehen werden kann. Diese Anwendungsweise macht allerdings nur für ein ländliches Gebiet Sinn, da in einer Stadt die Sicht viel wahrscheinlicher durch ein Gebäude nach einigen Metern unterbrochen wird, als durch weit entfernte Berge. Innerhalb von Städten kann zusätzlich zu einem Geländemodell auch auf 3D-Modelle einer Stadt zurückgegriffen werden. Als erste Annäherung an ein 3D-Modell einer Stadt wurden auch ein Verfahren entwickelt, das ein einfaches Stadtmodell generiert in dem es Liegenschaftsdaten mit einem detaillierten Höhenmodell extrudiert. Tabelle 7.1 gibt eine Übersicht über die möglichen Ausgangsinformationen und über die daraus berechneten Sichtbarkeitsmodelle.

	GPS	GPS + Kompass	GPS + Kompass + Neigungswinkel
Stadt	Kreis (kleiner Radius)	Kreis Sektor (kleiner Radius)	Kreis Sektor (kleiner Radius)
Stadt mit extrudiertem Stadtmodell	Kreis (kleiner Radius)	Kegelstumpf	Kegelstumpf
Stadt mit texturiertem 3D Stadtmodell	Kreis (kleiner Radius)	Kegelstumpf mit korrigierter Kameraposition und Richtungswinkel	Kegelstumpf mit korrigierter Kameraposition und Richtungswinkel
Land	Kreis (großer Radius)	Kreis Sektor (großer Radius)	Kreis Sektor (großer Radius)
Land mit Geländemodell	Kreis mit Sichtbarkeitsberechnung	Kreis Sektor mit Sichtbarkeitsberechnung	Kreis Sektor mit Sichtbarkeitsberechnung

Tabelle 7.1: Verfügbare Sensorendaten und die entsprechenden Sichtbarkeitsmodelle

Lag ein 3D Model der Gegend vor, kann ein Abgleich der Bildaufnahme mit der dreidimensionalen Szene vorgenommen werden. Dieser Abgleich war durch die fehlerbehafteten Sensordaten (vor allem GPS und Kompass waren je nach Gerät recht ungenau) auch deshalb nötig, da die gerenderte Szene nur in den wenigsten Fällen der Ansicht auf dem Bild selbst entsprach. Um diesen Abgleich zu erreichen wurden Szenen mit einer leichten Abweichung des Kamerastandpunktes und des Richtungswinkels berechnet und die berechneten Bilder mit der Bildaufnahme verglichen. Die einzelnen Gebäude wurden letzten Endes über Farbcodes im gerenderten Bild bestimmt, das heißt jedes Gebäude hat eine eindeutige Farbe, die im gerenderten Bild dann einen bestimmten prozentualen Anteil aufzeigen.

Um die identifizierten Objekten nach Wichtigkeit zu sortieren, berechnet das Tripod-System die sogenannte visuelle Salienz der einzelnen Bildteile. Dabei arbeitet Tripod hauptsächlich auf der Grundlage von bildgestalterischen Paradigmen. Das wichtigste Maß für eine gelungene Bildkomposition ist dabei der goldene Schnitt. Der goldene Schnitt zerlegt eine Strecke ungefähr in drei gleiche Teile. Wird eine Fläche durch ein Objekt in diesem Verhältnis zerteilt, wirkt sie interessant, aber auch nicht unruhig. Bei den Kameraherstellern hat es sich daher etabliert, dass die Kamera dem Nutzer zwei horizontale und zwei vertikale Hilfslinien als Orientierung anzeigt, um so eine möglichst interessante Bildaufteilung zu erhalten. Durch diese Hilfslinien wird die Fläche in neun gleich große Flächen aufgeteilt. Diese Hilfslinien schneiden sich in vier Punkten, an denen meistens das Hauptobjekt im Bild platziert wird. Aus diesen bildgestalterischen Mitteln werden von Tripod Objekte entlang der Linien oder in der Nähe der vier Schnittpunkten höher bewertet, als zum Beispiel Objekte, die direkt im Zentrum liegen. Abbildung 7.1 illustriert diese Arbeitsschritte. An diesem Beispiel lässt sich verdeutlichen, dass das rote Gebäude am wahrscheinlichsten das

Motiv des Bildes darstellt, da es sich in zwei Schnittpunkten der vier photographischen Hilfslinien befindet und zudem die größte Fläche im Bild einnimmt.

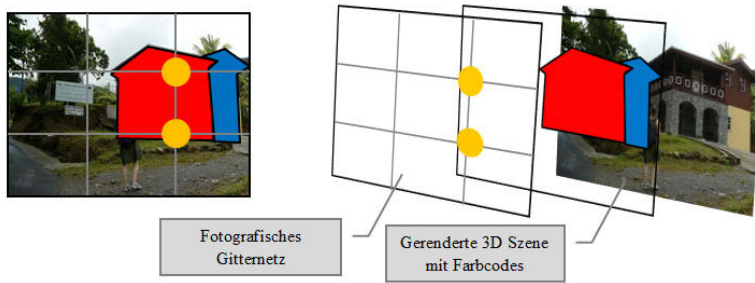


Abbildung 7.1: Identifizierung des Hauptobjekte im Bild mit Hilfe von Farbcodes und photographischen Hilfslinien

Wie schon im ersten Szenario ist das Ergebnis eine meistens recht lange Kandidatenlisten aus Konzepten und Schlagwörtern, von denen die meisten jedoch irrelevant für den entsprechenden Fotografenstandpunkt sind. Daher wird ausgehend von den identifizierten Toponymen und dem geografischen Ort in verschiedenen Webquellen nach einer Bestätigung für diese Schlagwörter gesucht. Zum einen wird die relative Häufigkeit von möglichen Begriffspaaren, aus Toponym + Kandidat berechnet, ähnlich der Kookkurrenzanalyse aus Kapitel 5. Zum anderen werden über Flickr verortete Bilder aus dem umliegenden Gebiet geladen, um anschließend zu überprüfen, ob die einzelnen Kandidaten ebenfalls in den Metadaten dieser Bilder zu finden sind. Um schließlich einen menschenlesbaren Titel zu erzeugen, werden die Informationseinheiten, wie Städtenamen oder berühmte Gebäude im Sichtbereich, in ein umgangssprachliches Template eingebettet. Liegt dem System nur eine einfache Position als räumliche Information vor, versucht das System zunächst die naheliegenden bekannten Toponyme zu benennen

und es können Bildtitel, wie „Aufgenommen in der Nähe von Innerthal im Schweizer Kanton Schwyz“ erstellt werden. Enthalten die Metainformationen sogar Richtungsangaben kann das wahrscheinlichste Hauptmotiv identifiziert werden und ein Titel wie „Zu sehen ist der Wägitalersee, ein Stausee“ generiert werden. Abbildung 7.2 zeigt ein beispielhaftes Ergebnis für dieses Szenario. In diesem Fall wurde ein englischer Titel erzeugt, der ins Deutsche übersetzt „Wägitalersee fotografiert am frühen Nachmittag in der Nähe der Wägitalerstrasse im Innerthal (Region)“ lautet. Interessant ist vor allem die Verwendung von verschiedensprachigen Datenquellen, in dem Fall das englischsprachige Corine und das deutschsprachige Swisstopo Vec 200. Alleine durch das Mapping in die eigene Konzeptontologie ist es möglich, beide Datensätze konsistent und ohne Doppeldeutigkeiten zu verwenden. Die Konzeptontologie ist somit auch ein Vermittler zwischen den einzelnen Datensätzen.

Technisch gesehen wurde jeder Teilprozess in ein eigenes Modul ausgelagert und als Webservice umgesetzt. Dieses grundsätzliche Vorgehen ermöglichte es, verschiedene Arbeitsabläufe aufzubauen und verschiedene Teilsysteme, wie die Konzept-Ontologie, für mehrere Arbeitsabläufe mehrfach zu verwenden. Diese Webservices können mit Hilfe von Orchestrierungssprachen, wie BPEL (Business Process Execution Language) miteinander verbunden werden. Im Projekt selbst wurde eine eigene Orchestrierungs-Engine in PHP umgesetzt. Ausgehend von den zwei Szenarien, mit unterschiedlichen Anforderungen an die Metadaten des Bildes, wurden zwei konkrete Prozessketten aufgebaut. Diese durchlaufen nach einem festen Schema bestimmte Module, um den erwünschten natürlingsprachigen Bildtitel, Schlagwörter und Beschreibungen zu generieren.

Zusätzlich zum innovativen Ansatz Metadaten für ein Bild zu erzeugen wurde nach neuen Möglichkeiten geforscht, auf diese Daten zuzugreifen.



Wägitalersee photographed in the early afternoon near Wägitalerstrasse in the Innerthal (Region)

Keywords: Agriculture, Moor, Lake, Forest

Datasets: Corine, Swisstopo Vec 200

Spotted feature types:

Corine: 0_wcsNodata, 23_231_pastures, 51_inlandWaters, 32_shrubAndorHerbaceousVegetationAssociations, 31_forests
Vec200: Wald, Fluss, Stausee, Building, Seeufer, Nebenströ, Uebrig

Concepts:

NaturalGrassland, Rape, PermanentCrops, Moor, Streets, Orchards, Pond, Rivers, AgriculturalLand, Potato, Straw, Heath, Dairy, Byways, Woodland, Nursery, Lakes, Tributary, Shorelines, Reservoir, Plantation, ArableLand, Pasture, Cereals, Meadow, Stream, MoorsAndHeathland, Forest, Tarn, Grass



Abbildung 7.2: Ergebnis des Tripod Workflows und die verwendeten Datenquellen

Zum einen über eine Suche und zum anderen über ein Empfehlungssystem. Beide Ansätze nutzen dabei die neuen geografischen Informationen der verorteten Bildern. Die eingeführte geografische Suche beruht dabei auf räumliche Diversität. Das Suchsystem versucht demnach, die Ergebnismenge möglichst gleichmäßig über den geografischen Raum zu streuen. Diese Streuung ergibt vergleichbare Ergebnisse, wie sie in Kapitel 2 für bildliche Illustrationen einer Karte auf Googlemaps zu sehen waren, wobei die Ergebnismenge durch bestimmte Schlüsselwörter eingegrenzt werden kann. Neben der Suche wurden auch die Verfahren aus den Kapiteln 5 und Kapitel 6 für das Projekt Tripod als eigenständiges Empfehlungssystem umgesetzt und in die Benutzeroberfläche des Tripod-Prototypen integriert.

Der Empfehlungsansatz aus Kapitel 6 baut dabei auf das Modul für prototypische Bildgruppen auf, dennoch sind beide System getrennt voneinander über zwei öffentliche Schnittstellen von außen zugänglich, so dass man von zwei Systemen reden kann. Das erste System ist der Heatmapper, der typische Standpunkte für ein räumliches Gebiet berechnet, die Fotografen zuvor in der Stadt eingenommen haben. Das zweite System ist der GeoRecommender und verwendet die berechnete Gewichtung des Heatmappers, um typische Fotografien in der eigenen Bilddatenbank zu finden oder die Standpunkte über einen kollaborativen Filter neu zu gewichten. Abbildung 7.3 gibt eine Übersicht über die einzelnen Schichten des kompletten Systems. Auf der untersten Ebene können verschiedene Datenquellen, die entweder über ein öffentliches Interface wie REST oder SOAP verfügen, an das System gebunden werden. Die Daten werden über eine Übersetzungsschicht in eine gemeinsame interne Repräsentation überführt. Im Fall des geografischen Bildempfehlungssystems kommuniziert der Heatmapper über die Übersetzungsschicht mit verschiedenen öffentlichen Bildgalerien, wie zum Beispiel Flickr oder Panoramio. An dieser Stelle werden aber auch die

Bilddaten derjenigen Galerie eingeladen, für die später eine Empfehlung ausgesprochen wird. Um die Abfrage von häufig genutzten Daten zu beschleunigen, wird in dieser Schicht auch ein Puffer verwendet, der überflüssige Anfragen an externe Dienste verhindern soll. Die externen Anfragen führen leider ansonsten zu Verzögerungen, da die Kommunikation in den meisten Fällen über das Internet erfolgt und das die Roundtripzeiten von Anfrage und Antwort um einiges verlängert. Nach dieser Übersetzungsschicht folgt die eigentliche Anwendungsschicht, in der die Daten von dem Heatmapper System zu räumlichen Entscheidungen vorverarbeitet werden. Diese räumlichen Entscheidungen werden dann entweder vom Heatmapper selbst oder dem Empfehlungssystem gewichtet. Aus diesen gewichteten räumlichen Entscheidungen werden dann die Ergebnislisten erzeugt. Im Fall des Heatmappers ist das Ergebnis eine Heatmap des gewählten Bereiches und im Fall des Empfehlungssystems eine Liste mit empfohlenen Bildern, die die entsprechenden räumlichen Entscheidungen illustrieren. Beide Dienste sind unabhängig voneinander über eine SOAP-Schnittstelle erreichbar um von verschiedenen Clientanwendungen angesprochen zu werden. Auf die einzelnen Anwendungssysteme, dem Heatmapper und dem GeoRecommender, und zwei beispielhafte Clientanwendungen, die diese Systeme nutzen, wird im Folgenden noch einmal genauer eingegangen.

7.2 Heatmapper Prototyp

Ein erster Prototyp des Systems wurde als öffentliche Webapplikation erstellt, die unter <http://www.kinf.wiai.uni-bamberg.de/tp5/heatmap.php> zugänglich ist. Dies ist eine PHP Webanwendung, mit deren Hilfe die beschriebenen Arbeitsschritte und Maße getestet wurden. Dieser Schritt war sinnvoll, da bereits die angesprochenen Bildgalerien über eine Webschnittstelle verfügten und

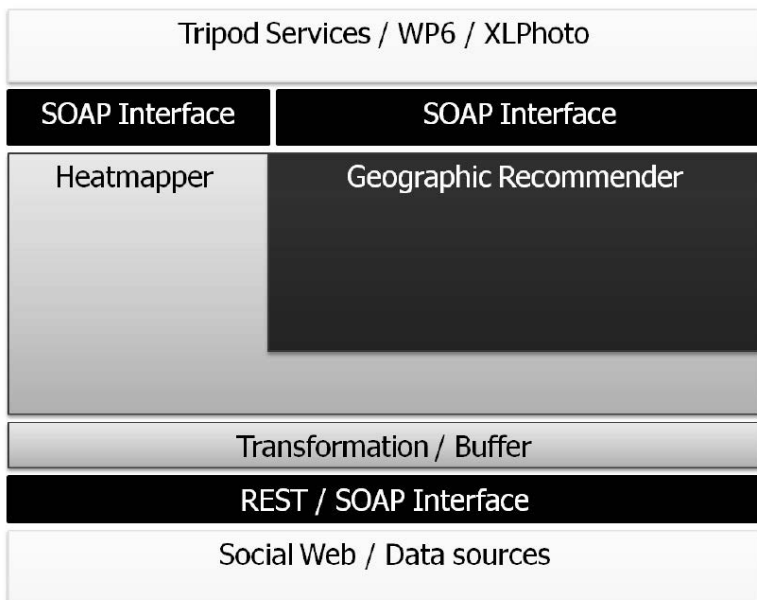


Abbildung 7.3: Übersicht der Einbettung von Heatmapper und GeoRecommender

als Austauschformat JSON (JavaScript Object Notation) verwendet wird. JSON ist ein recht einfaches Textformat, das auf einer Unter-
menge der JavaScript-Sprache basiert, was es vereinfacht es mit einer
Scriptsprache, wie PHP oder JavaScript einzulesen und zu parsen. Die
Verarbeitung der Daten wird dann direkt von der PHP-Anwendung
übernommen. Die Ergebnisse dieser Anwendung werden über den
Browser ausgegeben. Weiterer Vorteil eines Webinterfaces ist, dass
das System von überall einsichtig ist, wo auch Internet zur Verfügung
steht. Die hohe Verfügbarkeit erleichtert zudem die Kommunikation
mit anderen Menschen über so ein System. Allerdings war eine Vorgabe
des Projekt Tripods die Verfügbarkeit der einzelnen Systeme über
eine Webservice-Schnittstelle. Hinzu kam, dass eine Umsetzung der
Businesslogik in einer höheren Programmiersprache einen deutlichen
Geschwindigkeitsgewinn bedeutet. So wurde die Businesslogik in einer
Javaanwendung gekapselt und über einen Webservice für das We-
binterface zugänglich gemacht. Dieser Webservice konnte dann sowohl
für Tripod aber auch für weitere Anwendungen, wie zum Beispiel
von einer Facebook-Applikation genutzt werden. Zusätzlich zeigen
die Screenshots in Kapitel 6 die optimierte Darstellung für mobile
Endgeräte. Webapplikation in einem mobilen Umfeld haben dabei
einen besonderen Reiz, da sich mittlerweile über Browser Plug-Ins
wie Google Gears oder den neuen Möglichkeiten von HTML5 möglich
ist, die momentane Position des Endgerätes abzufragen. In diesem
Umfeld lassen sich daher die 10 populärsten Orte um den momentanen
Standpunkt des Gerätes berechnen, ganz im Sinne eines ortsbezogenen
Dienstes. In den beiden anderen Webanwendungen kann die Region,
die von Interesse ist, aber auch frei gewählt werden.

Das Besondere an der Umsetzung des Heatmappers ist es, dass die
Daten selbst nicht zentral gehalten, sondern aus verschiedenen exter-
ne Datenquellen geladen werden. Um auf die Anbindung dieser Da-

tenquellen so flexible wie möglich zu gestalten, wurde eine Mediation Architektur verwendet. Eine zentrale Klasse, der Mediator, ist für die Abfrage der externen Datenquellen verantwortlich. Dabei werden alle externen Schnittstellen mit Hilfe eines einheitlichen Interfaces gekapselt und beim Mediator angemeldet. Diese gekapselte Schnittstelle nennt man Translator, da er die externe Schnittstelle in eine interne Repräsentation übersetzt.

Wird nun eine Anfrage an den Mediator gesendet, kann dieser die Anfrage an alle aktiven externen Datenquellen verteilen und die einzelnen Ergebnisse bündeln. Soll in Zukunft eine neue Datenquelle eingefügt werden, muss lediglich ein neuer Translator geschrieben und beim Mediator angemeldet werden. Implementiert wurden Translatoren für Flickr, Panoramio, Locr, Alinari und dem Prototypen der geografischen Suchmaschine, die für das Projekt Tripod entwickelt wurde. Auch für die Evaluation wurde ein Translator für die Daten aus einem Experiment erstellt. Abbildung 7.4 gibt eine Übersicht über die verwendete Mediation-Architektur.

7.2.1 Webservice

Der Webservice bietet vier Methoden über eine SOAP-Schnittstelle an. Über die Methoden *getHeatmapByRegion* und *getHeatmapByPosition* kann das System nach einer vollständigen Heatmap gefragt werden. Die Methode zur Berechnung einer Heatmap kann mit Hilfe von *getHeatmapByRegion* auf eine bestimmte Region angewendet werden, die als WKT (Well Known Text, zum Beispiel „POLYGON((10 49, 10 50, 11 50, 11 49, 10 49))“) angegeben werden kann. Die zweite Methode *getHeatmapByPosition* berechnet die Heatmap in einem Gebiet um einen angegebenen Punkt. Dazu wird ein Puffer um den angegebenen Punkt berechnet und mit dessen Hilfe die Heatmap berechnet.

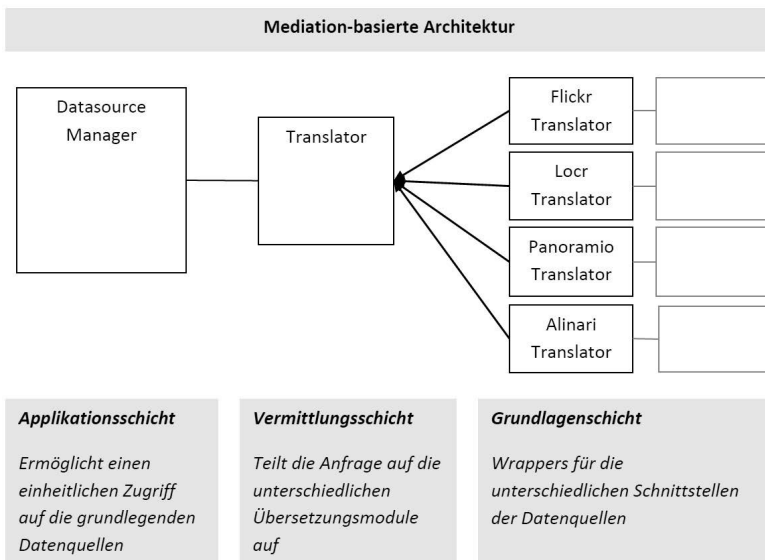


Abbildung 7.4: mediationbasierte Architektur

Alternativ zu der Anfrage nach einer kompletten Heatmap stellt der Webservice auch die Methoden *getPopularityByRegion* und *getPopularityByPosition* zur Berechnung der Popularität zur Verfügung. Bei der Anfrage nach einer Region werden zunächst die Cluster gebildet und dann die berechnete Popularität von dem Cluster in der Region aufsummiert. Bei der Anfrage für einen bestimmten Punkt wird zunächst eine Heatmap für die nächste Umgebung berechnet und dann die Popularität des am nächsten gelegenen CPVs zurückgegeben. Als Abstandsmaß wird der euklidische Abstand zum Schwerpunkt des Clusters verwendet. Die Vorteil der beiden Methoden liegt an dem geringeren Payload der Antwort-Nachricht, da hier nur ein einzelner Zahlenwert übertragen werden muss und nicht wie bei den ersten beiden Methoden eine recht umfangreiches XML Dokument (je nach Größe und Lage der Region).

7.3 GeoRecommender Prototyp

Das geografische Empfehlungssystem baut auf die Ergebnisse des Heatmappers auf und erweitert ihn um die Möglichkeit, Empfehlungen für ein frei wählbares Gebiet zu erzeugen. Die Anforderungen an das Empfehlungssystem für das Projekt Tripod gingen allerdings über den Einsatzbereich eines geografischen Empfehlungssystems hinaus. Daher war es nötig ein hybrides System zu implementieren, das es erlaubt, verschiedene Empfehlungsansätze miteinander zu verbinden. Grundsätzlich sollte das System dennoch einen benutzerbasierten Ansatz verfolgen, sodass zunächst eine Möglichkeit geschaffen werden musste, das Empfehlungssystem auf Verhalten des Nutzers auf Clientseite aufmerksam zu machen. Dazu gibt es generell zwei Umsetzungsmöglichkeiten, einen Push- oder einen Pull-Service.

Um das Empfehlungssystem als einen Pull-Service zu implementieren, muss der Client jeden Schritt des Nutzers protokollieren. Das Emp-

fehlungssystem würde dann regelmäßig beim Client nachfragen, welches Nutzerverhalten sich verändert hat und dann dementsprechend Empfehlungen verändern. Dies setzt einen hohen Integrationsaufwand voraus, da auch Schnittstellen für das Empfehlungssystem zum Client geschaffen werden müssen, damit er auf die Nutzerdaten auf Clientseite zugreifen kann.

Bei einem Push-Service bietet das Empfehlungssystem selbst eine Schnittstelle an, über die das Nutzerverhalten gemeldet wird. Das heißt, sowohl Initiative als auch Datenkontrolle liegen beim Clientsystem. Dieses bestimmt selbst, welches Nutzerverhalten es meldet. Um das Empfehlungssystem in eine bestehende Anwendung zu integrieren, müssen daher in die eigene Anwenderlogik zusätzlich die Schnittstellenaufrufe an das Empfehlungssystem implementiert werden.

Innerhalb von Tripod wurde das Empfehlungssystem als Push-Service umgesetzt. Die Methode, um eine Nutzeraktion beim Empfehlungssystem zu melden, lautet *setRecommendation(Bild, Nutzer, Aktion, Fußabdruck)*. Die Parameter für diese Methode sind die Kennung des Bildes, der Urheber der Aktion, die Aktion, welche auf dem Bild ausgeführt wurde, und ,falls vorhanden, der geografische Fußabdruck des Bildes.

Nachdem der Nutzer seine ersten Präferenzen mitgeteilt hat, kann dieser eine Empfehlungsanfrage stellen. Die Abfrage erfolgt über die Methode *getRecommendation(Nutzer, Fußabdruck)*, die mehrere Anfragemuster ermöglicht, da beide Parameter optional sind. Auf jedes der vier möglichen Anfragemuster reagiert das System mit einer passenden Empfehlungsmethode. Abbildung 7.5 gibt einen Überblick über den inneren Aufbau des hybriden Empfehlungssystems. Generell kann das Empfehlungssystem in zwei logische Teile unterteilt werden:

- Der linke Ast besteht aus Empfehlungssystemen, die eine geografische Empfehlung berechnen. Diese Teilsysteme ergeben dabei

eine Empfehlungskette, in der sich die beiden Teilsysteme befinden, die in dieser Arbeit beschrieben sind: der Heatmapper, für die Empfehlung von typischen Bildern, und das geografische Empfehlungssystem, für die Berechnung einer personalisierten Empfehlung.

- Im rechten Ast des Systems befinden sich Empfehlungssysteme, die ohne geografische Anforderungen auskommen. Zunächst wird eine versucht eine Empfehlung mit Hilfe eines nutzerbasierten kollaborativen Filters zu berechnen. Eine Vorgabe an diesen Filter war, dass der Nutzer nicht selbst eine Bewertung der einzelnen Bilder abgeben sollte, was für eine direkte Umsetzung eines kollaborativen Filters nach dem Vorbild von Resnick nötig gewesen wäre. Stattdessen werden die Bewertungen indirekt aus dem Nutzerverhalten berechnet. Dazu werden die einzelnen Aktionen des Benutzers vom System auf eine Bewertung auf einer Intervallskala abgebildet. Diese Abbildung wird durch das Modul TranslationRecommender durchgeführt und dann ein Kollaborativer Filter auf die abgebildeten Bewertungen angewendet. Softwaretechnisch betrachtet kapselt der TranslationRecommender einen kollaborativen Filter. Die möglichen Nutzeraktionen und die Abbildung auf einen konkreten Zahlenwert sind in Tabelle 7.2 zu sehen.

Jede gemeldete Nutzeraktion wird vom Empfehlungssystem in einer Datenbank gespeichert. Um Nutzeraktionen möglichst konsistent zu halten müssen einige von ihnen ebenfalls gemeldet werden, auch wenn sie für die eigentliche Bewertung nicht benötigt werden, zum Beispiel um die Positionsangaben eines Bildes zu aktualisieren. Wird ein Bild komplett gelöscht, müssen auch die Datensätze des zugehörigen Nutzerverhaltens aus den Datensätzen des Empfehlungssystems entfernt werden.

An den Bezeichnungen von einigen Aktionen lässt sich der zugrunde liegende Anwendungsfall erkennen. Ein Beispiel ist der Begriff der Lightbox. Eine Lightbox ist das digitale Gegenstück zum Lichttisch der analogen Fotografie und dient zur Sichtung von kleineren Bildgruppen. Verwendet wird eine Lightbox hauptsächlich, um Alternativen eines Motivs oder zu einem bestimmten Thema vor einem Kaufvorgang gegenüberstellen zu können. Technisch gesehen ist sie vergleichbar mit einem persönlichen Ordner, in den Bilder aus der Gesamtsammlung gesondert gruppiert werden können. Ähnlich wie bei der Voransicht, muss der Nutzer ein gewisses Interesse an dem Bild haben, um es in so eine private Sammlung einzuordnen. Eine spezielle Form der Lightbox stellt dabei der Einkaufskorb dar, der ebenfalls eine private Sammlung von Bildern ist. Allerdings drückt der Nutzer durch das Einfügen eines Bildes in diese Sammlung schon eine gewisse Kaufabsicht aus und das Interesse an diesen Bildern ist daher höher zu werten als das Interesse an den einzelnen Bildern in einer Lightbox.

Der zweite Empfehlungsansatz, der für Tripod den Fall einer nicht-geografischen Empfehlung umgesetzt wurde, entspricht den aktuellen Bildern, für die sich die Mehrzahl der Nutzer interessierte. Diese Empfehlung hat die Einschränkung, dass der Nutzer selbst noch keine Aktion auf einem der Bilder ausgeführt hat. Dieser Ansatz bildet die Minimalempfehlung, falls keine andere Empfehlung erfolgreich war.

Tabelle 7.3 gibt noch einmal eine Übersicht über die vier möglichen Anfragemuster für das Empfehlungssystem. Die Einträge sind dabei folgendermaßen zu interpretieren: Ein Bindestrich bedeutet, dass der entsprechende Parameter nicht gesetzt ist, und ein Stern, dass ein gültiger Wert übergeben wurde. Für eine persönliche und geografische Empfehlung muss demnach sowohl eine eindeutige Identifikation des Nutzers übergeben werden, wie auch ein gültiger Fußabdruck im WKT-Format.

Aktion	Id	Wert	Beschreibung
FOCUS	101	1	Bild wurde in Thumbnailgröße angezeigt
PREVIEW	102	2	Bild wurde in einer Voransicht geöffnet
DOWNLOAD	103	3	Bild wurde heruntergeladen
ADDLIGHTBOX	201	2	Bild wurde zu einer Lightbox hinzugefügt
REMOVELIGHTBOX	202	-1	Bild wurde aus der Lightbox entfernt
ADDCHECKOUT	203	2	Bild wurde zum Warenkorb hinzugefügt
REMOVECHECKOUT	204	-1	Bild wurde aus dem Warenkorb entfernt
GEOCODE	301	-	Bild wurde georeferenziert
REMOVE	401	-	Bild wurde aus der Bildgalerie entfernt

Tabelle 7.2: beispielhafte Gewichtung der Interaktionen mit einem Bild.

U	FP	Empfehlungsmodul
-	-	<i>RandomRecommender</i>
*	-	<i>CollaborativeRecommender</i> → <i>RandomRecommender</i>
-	*	<i>HeatmapRecommender</i>
*	*	<i>PersonalHeatmapRecommender</i> → <i>HeatmapRecommender</i>

Tabelle 7.3: Anfragemuster und verwendbare Empfehlungsmodule. Ein Pfeil in der Tabelle verdeutlicht eine Fallbackstrategie. U steht für Nutzer (User) und FB für den geographischen Fußabdruck (Footprint)

Abbildung 7.5 zeigt den internen Aufbau des Empfehlungssystems als Klassendiagramm, wobei die grau unterlegten Klassen die vier Empfehlungsansätze bereitstellen. Die Bearbeitungskette beginnt mit dem GeoSwitchingRecommender, der darüber entscheidet, ob die Anfrage eine geografische Anforderung enthält und schickt die Anfrage dann entweder einem GeoRecommender oder einem normalen Recommender Modul. Beide Teilläste enthalten einen ChainRecommender, der das Pattern einer „Zuständigkeitskette“ verfolgt. In diese Kette können beliebig viele Empfehlungsansätze eingebunden werden sowie verschiedene Strategien gewählt werden, wie diese Kette durchlaufen werden soll. In der momentanen Umsetzung wird ein Empfehlungsmodul nach dem anderen ausgeführt. Wird dabei eine Empfehlung mit mindestens einem Bild erzeugt, wird die Kette unterbrochen und das Ergebnis als Empfehlung zurückgegeben. Hier sind andere Strategien, wie sie auch in der Übersicht in Kapitel 4 unter den hybriden Systemen beschrieben werden, denkbar.

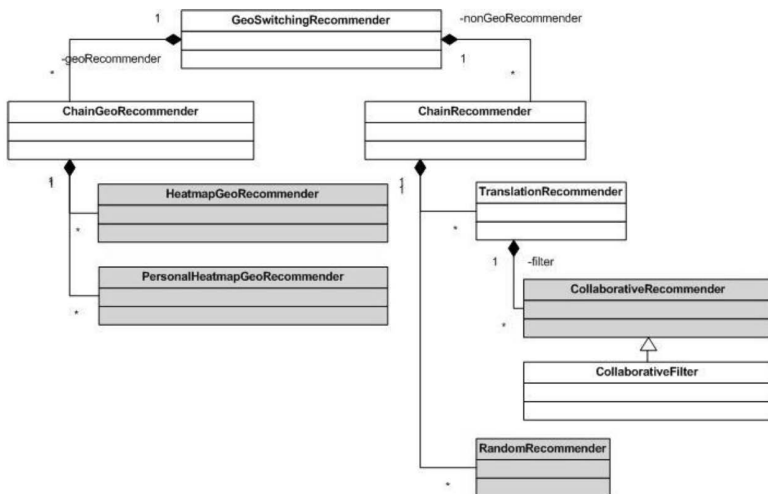


Abbildung 7.5: interne Teilsysteme des GeoRecommender Prototypen

Im Folgenden wird die Einbettung der beiden Empfehlungssysteme in zwei verschiedene webbasierte Plattformen beschrieben. Für das Projekt Tripod hat man sich dabei entschlossen ein bestehendes Produkt zu erweitern und für eine zweite Integration wurde die Facebook-Plattform verwendet, um eine eigenständige Applikation für die Nutzergemeinschaft des sozialen Netzwerkdienstes anzubieten. Die beiden Systeme verdeutlichen die Flexibilität des Systems, da sie auf unterschiedlichen Bildquellen und Eingabe-Metaphern beruhen. Der Tripod-Prototyp zum Beispiel verwendet Bilder aus dem eigenen Archiv und verwendet das Browsing-/Downloadverhalten des Nutzers auf der restlichen Seite. Eine räumliche Entscheidung wird demnach dann getroffen, wenn der Nutzer ein verortetes Bild herunterlädt oder es in einer Voransicht betrachtet.

Wichtig für eine Einbettung in ein bestehendes System ist die Anforderung, dass der Nutzer sich dem System gegenüber identifizieren kann. Dies ist für den Tripod-Prototypen der Fall, da das System eine Nutzererkennung für die eventuellen Einkäufe braucht und um spezielle Teile der Galerie nur für bestimmte Kunden zugänglich zu machen. Für ein soziales Netzwerk wie Facebook, ist eine eindeutige Nutzeridentifikation unumgänglich.

7.3.1 Einbettung in eine Bildergalerie am Beispiel des Tripod-Prototypen

Abbildung 7.6 zeigt die Integration des geografischen Empfehlungssystems im Zusammenspiel mit den anderen Modulen des Projekts. Dabei spielen drei Module eine wesentliche Rolle:

- **Bild-Managementsystem:** Ist für die Darstellung der Bilddatenbank für den Nutzer zuständig. Über dieses Modul hat der Nutzer Zugriff auf die Bildbestände. Dieses Modul hat zusätzlich die

Kontrolle über den Datenaustausch zwischen der Bilddatenbank, der Suchmaschine und dem Empfehlungssystem.

- Geografic Search Engine: Ermöglicht den geografischen Zugriff auf die Bilddatenbank des Tripod Prototypen.
- Geografic Recommender: Berechnet eine Empfehlung für den momentanen Nutzer und/oder den angegebenen Raum.

Als Bild-Managementsystem für das Projekt-Tripod wurde ein bestehendes Produkt der Firma Centrica ausgewählt und erweitert: XLPhoto. XLPhoto ist eine Serverlösung für Digital-Asset-Management speziell für professionelle Bilderarchive. Das System bietet für den Endnutzer eine Weboberfläche, die für die neuen Funktionalitäten der Tripodmodule angepasst wurde. Im Folgenden wird daher von dem Tripod Prototypen die Rede sein, wenn die angepasste Version von XLPhoto gemeint ist.

Der Tripod-Prototyp beinhaltet intern eine Bilddatenbank aller Bilder, die sowohl für die Suche als auch die Empfehlung zugänglich sein sollen. Ebenso, wie schon das Empfehlungssystem, wurde die Suche nur lose als Push-Service angebunden. Auch die Suche hat keinen direkten Zugriff auf die Daten des Tripod Prototypen. Um ein Bild daher für die Suche sichtbar zu machen, muss es zunächst bei der Suchmaschine angemeldet werden. Erst dann kann die Suchmaschine die Metadaten entsprechend indexieren. Ändern sich die Metadaten oder wird das Bild gelöscht, muss dies ebenfalls erst wieder an die Suchmaschine gemeldet werden. Somit hat der Tripod Prototyp die volle Kontrolle über die Indexierung der Bilder.

Auch für das Empfehlungssystem wird jede Aktion, die auf einem Bild ausgeführt wird, dem geografischen Empfehlungssystem über seine SOAP-Schnittstelle gemeldet. Im Falle des Tripod Prototypen wird eine Untermenge der Aktionen aus Tabelle 7.2 verwendet.

Wird nun eine geografische Empfehlung von dem System verlangt, wird über eine geografisch eingeschränkte Suche mit Hilfe der geografische Suchmaschine auf die Bilder des Tripod Prototypen zugegriffen. Das bedeutet eine Empfehlung ist immer dann möglich, wenn die Ausgangsmenge sich ebenfalls im geografischen Suchindex der Suchmaschine befindet. Alternative hätten man auch eine Empfehlung der fotografischen Standpunkte zunächst mit Hilfe von anderen Datensätzen berechnen und dann einen Abgleich mit den Standpunkten der Bilder aus der zu filternden Datenbank über die Ortsangaben auf den Bilder machen können. Dadurch wären allerdings fotografische Positionen bzw. räumliche Entscheidungen, die nur in einem der beiden Datensätzen vorkommen, vernachlässigt worden. Zudem wäre das Ergebnis alleine durch die Datenquellen bestimmt, die durch die eigentlichen Gewichtungsschritte verwendet wurden. Diese Eigenschaft kann aber durchaus sinnvoll sein, wenn man zum Beispiel eine Datenbank mit professionellen Aufnahmen durch die Bilder von Nutzern filtern will, ohne das Empfehlungsergebnis durch die professionellen Bilder zu verändern. Mit anderen Worten müssen dadurch die räumlichen Entscheidungen nicht aus der gleichen Quelle kommen, wie die bildliche Repräsentation der entsprechenden Orte.

Abbildung 7.7 zeigt zwei Screenshots des Tripod Prototypen. Abbildung 7.7(a) sieht man das Ergebnis einer Suchanfrage, wobei zusätzlich auch Empfehlungen ohne konkreten Raumbezug angezeigt werden (drei Bilder auf der linken Seite). Abbildung 7.7(b) erkennt man die Kartenansicht, die speziell für Tripod geschrieben wurde. Dabei lassen sich zum einen die typischsten Ansicht berechnen, zum anderen auch persönliche Ansichten empfehlen. Als geografische Begrenzung der Anfrage wird dabei der momentane Kartenausschnitt verwendet. Eine persönliche Empfehlung ist dabei allerdings nur möglich, falls der Nutzer angemeldet ist.

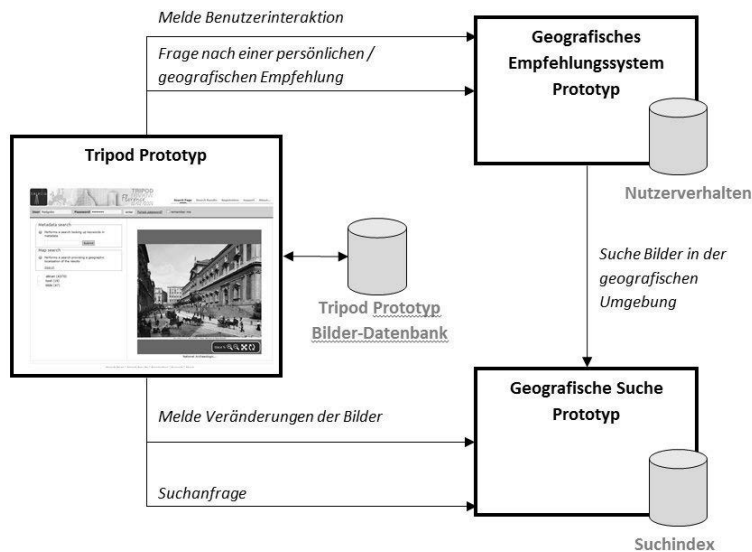
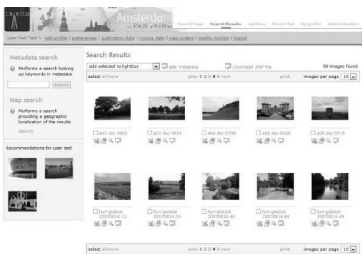
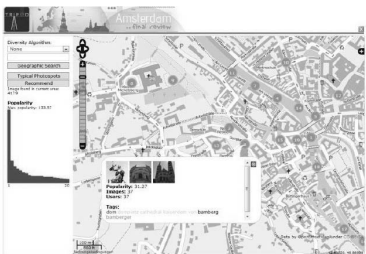


Abbildung 7.6: Integration des Empfehlungssystem in den Tripod Prototyp



(a) Download von Bildern aus einer gemeinsamen Bilddatenbank als räumliche Entscheidung.



(b) Geografische Empfehlung im Tripod Prototypen

Abbildung 7.7: Web Frontend für das Projekt Tripod.

7.3.2 Einbettung in eine Anwendung des Social Web am Beispiel von Facebook

Das Internet hat in den letzten Jahren einige grundlegende Veränderungen durchlebt. Inhalte werden heutzutage nicht mehr hauptsächlich von namenlosen Autoren veröffentlicht, sondern die Nutzer selbst veröffentlichen unter ihrem wirklichen Namen einen Großteil der Inhalte, was den sozialen Charakter des sogenannten Social Web ausmacht. Jeder Nutzer stellt sich dabei selbst durch ein eindeutiges Nutzerprofil dar.

Die Idee, den Nutzer im gesamten Internet durch ein eindeutiges Nutzerprofil zu repräsentieren, ist insbesondere für Empfehlungssysteme interessant, vor allem, wenn es sich dabei um kollaborative Systeme handelt. Zum Beispiel ist es möglich, Entscheidungen des Nutzers auf mehreren Anwendungen mit zu verfolgen. So könnte ein Nutzer, der Bilder auf Panoramio einstellt, eindeutig demselben Nutzer auf Flickr zugeordnet werden. Dadurch könnte ein Bildempfehlungssystem auf der Basis seiner Bilder in beiden Bildsammlungen arbeiten, anstatt zwangsweise beide Bildergruppen zwei unterschiedlichen Nutzern zuzuordnen. Soziale Netzwerke, wie Facebook, haben bereits die Rahmenbedingen dafür geschaffen, indem sie spezielle Authentifizierungs-Lösungen für andere Seiten anbieten. Dazu ermöglicht Facebook den Betreibern einer Internetanwendung eine Authentifizierung der eigenen Nutzer mit Hilfe ihres Facebook-Kontos. Solche Möglichkeiten werden auch als Single-Sign-On bezeichnet. Ein weiterer wichtiger Vertreter einer Single-Sign-On Lösung ist OpenID, dessen Standard von den meisten großen Softwareunternehmen unterstützt wird. Einmal bei einer Single-Sign-On Lösung angemeldet, muss sich der Nutzer auf anderen Seiten, die dieselbe Lösung verwenden, nicht erneut anmelden. Der große Vorteil von Single Sign-On Lösungen, ist dabei die eindeutige Zuordnung von Benutzern über Systemgrenzen hinaus. Der Anbieter der Single Sign-On

Lösung ist dadurch in der Lage das komplette Nutzungsverhalten auf allen Seiten zu überwachen, die die eigene Lösung nutzen.

Gerade Marktführer Facebook stellt dabei eine Vielzahl von unterstützenden Plugins zur Verfügung, um Funktion, wie zum Beispiel Nutzerkommentare, auf Fremdseiten anzubieten. Die Informationen, die über diese Funktionen ausgetauscht werden, sind im Gegenzug auch für Facebook sichtbar. So wird der Strom an persönlichen Neuigkeiten auf Facebook über diese Seiten hoch gehalten, indem diese spezielle Statusmeldungen im Facebook-Profil einfügen. Zudem ermöglicht das die Kommunikation der Facebook-Nutzer außerhalb und innerhalb der eigenen Systemgrenzen.

Als zweiten Anwendungsfall wurde eine Web-Applikation umgesetzt, die die Single Sign-On Lösung von Facebook verwendet. Im Mai 2008 öffnete sich Facebook mit der dritten Iteration ihrer Single-Sign-On Lösung. Diese wird Facebook Connect genannt. Diese bietet Drittseiten auch die Möglichkeit an, auf alle Möglichkeiten und Inhalte der Facebook-Plattform zuzugreifen. Interessant für das geografische Empfehlungssystem ist dabei der einfache Zugriff auf die Nutzerdaten und somit auch auf deren eingestellte Bilder. Im Gegensatz zum Tripod-Prototypen wurden allerdings alle Bilder von den Nutzern selbst erstellt. Ein verortetes Bild bedeutet in diesem Fall auch, dass der Nutzer sich an der entsprechende Stelle aufgehalten hat.

Die Verwendung eines kollaborativen Empfehlungssystems in einem sozialen Netzwerk macht aber auch aus anderen Gründen Sinn. Kernstück eines sozialen Netzwerkes sind die Verbindungen der Nutzer untereinander, die in den meisten Fällen realen Freundschaften entsprechen. Die Berechnung von Nutzerähnlichkeiten über das räumliche Verhalten ermöglicht es daher, sozialen Netzwerken, neue Beziehungen zu berechnen und dadurch eventuell neue Freundschaften zu initiieren. Die

geografischen Ähnlichkeitsbeziehungen könnten somit ergänzend zu den expliziten Freundschaftsbeziehungen genutzt werden.

Freundschaftsbeziehungen werden oft als Graph dargestellt. Facebook spricht daher auch von dem sozialen Graphen, wenn es das Beziehungsgeflecht zwischen den Nutzern meint. Jeder Nutzer wird dabei als Knoten dargestellt und die Beziehung der einzelnen Personen wird als Kante symbolisiert. Mittlerweile gibt es auch die ersten Ansätze, die Informationen aus dem sozialen Graphen für kollaborative Empfehlungsansätze zu verwenden (Spertus u. a., 2005). Die Idee dabei ist häufig, die berechneten Ähnlichkeiten durch explizite Freundschaftsbeziehungen zu ersetzen oder zu ergänzen ((Konstas u. a., 2009), (De Pessemier u. a., 2009), (Ye u. a., 2010)).

Tabelle 7.4 gibt eine kurze Übersicht über verschiedene Social-Web-Anwendungen mit unterschiedlichen Schwerpunkten. In all diesen Anwendungen muss sich der Nutzer zunächst identifizieren, indem er ein eigenes Nutzerprofil erstellt. In Zukunft könnte hier auch eine Single-Sign-On Lösung zum Einsatz kommen, wie Facebook oder OpenID. Für unseren Anwendungsfall sind natürlich vor allem verortete Bilddatensätze interessant. Der grundlegende geografische Empfehlungsansatz ist allerdings auch auf andere ortsgebundene Inhalte anwendbar.

Für eine erste Einbettung in eine Social-Web-Anwendung wurde Facebook ausgewählt. In der Übersichtstabelle 7.4 ist Facebook zwar die einzige Anwendung, die keine Ortsinformationen für seine Inhalte bereit stellt, es ist aber auch die einzige Anwendung, die über eine ausgereifte API verfügt, die den Zugriff auf alle Inhalte und den sozialen Graphen bereithält. Mit 400 Millionen Menschen weltweit ist Facebook zudem die erfolgreichste Social-Web-Anwendung.

Facebook bietet seit Anfang 2010 selbst ein Empfehlungssystem für das komplette Internet an. Dies wird durch ein einfaches Plugin ermöglicht, das es dem Nutzer ermöglicht, über einen beliebigen Inhalt eine Mei-

Name	Räumliche Infor- mationen	Inhalt	Sozial Beziehun- gen
Sharing sites			
Panoramio	POV	Bilder	Nein
Flickr	POV	Bilder	Ja
Locr	POV (also POI)	Bilder	Ja
GeoNotes	Punkt	Informationen	Nein
Social Media / ortsbasierte Netzwerke			
Facebook	-	Bilder, Videos, Informationen	Ja
BrightKite	Punkt	Personal Locati- on, Information, Bilder	Ja
CitySense	Punkt	Location hot- spots	Nein
GyPSii	Punkt	Bilder, Videos, Informationen/- Places	Ja
MobiLuck	Punkt	Informationen	Ja
Loopt	Punkt	Informationen	Ja
Plazes	Punkt	(geplante) per- sönliche Position	Ja
Whrrl	Punkt	Informationen	Ja
iPling	Punkt	Momentane Posi- tion (adhoc social networks)	Ja
Google Latitude	Punkt	Personal Locati- on	Ja
Tripadvisor	Punkt	Position des Nut- zers	Ja
Games with a purpose			
Geograf	POV, Gitter	Bilder	Nein
CityExplorer	POV, POI	Bilder	Ja

Tabelle 7.4: Übersicht über soziale Anwendungen, die räumlichen Informationen und die verwalteten Multimedia-Inhalte

nung abzugeben. Für ein bestimmtes Produkt oder Medieninhalt wird dabei ein Knopf angeboten, auf dem „Gefällt mir“ zu lesen ist. Dadurch gibt der Nutzer letzten Endes ein eindeutiges Feedback ab, auf welcher Seite er welches Produkt besonders gut findet. Dadurch sammelt Facebook nicht nur die Präferenzen eines Nutzers, sondern auch die Präferenzen seiner Freunde, die als Empfehlungen ausgegeben werden können. Der Nutzen dieser einfachen Empfehlungstechnik wurde dabei schnell von den Betreibern diverser Seiten erkannt und so befinden sich auf immer mehr Seiten ein „Gefällt mir“-Knopf (zum Beispiel auf dem offiziellen Levi's-Store ¹⁷).

Für das geografische Empfehlungssystem sind allerdings nur verortete Bilder interessant, die mit dem Nutzerprofil auf Facebook verbunden sind. Im Moment greift die Facebook-Applikation auf die Bilder zu, die über die interne Bild-Applikation verwaltet werden. Allerdings wurden die Bilder und ihre Metadaten aus der Sicht von Facebook auf das Wesentliche reduziert, um die Bilder von ungefähr 400 Millionen Menschen überhaupt verwalten zu können. Dabei wurde die Auflösung der Bilder minimiert und ihre Metadaten komplett entfernt. Das hatte leider zur Folge, dass selbst Bilder, die schon zuvor verortet wurden, keine geografischen Positionen in ihren Metadaten ausweisen. Daher wurde für die Anwendung eine Möglichkeit umgesetzt, die Bilder zu laden und mittels Drag and Drop auf einer Karte zu verorten. Diese mühsame Vorarbeit hat aber auch den Vorteil, dass der Nutzer eventuell schon eine Vorauswahl trifft, für welche Bilder sich eine Georeferenzierung überhaupt lohnt. Jedes Mal, wenn der Nutzer ein Bild auf die Karte setzt, wird der GeoRecommender mit Hilfe einer SOAP-Nachricht über diese Entscheidung aufmerksam gemacht. Dieser verändert dann entweder eine bestehende räumliche Entscheidung oder legt eine neue für den jeweiligen Nutzer und das entsprechende Bild an. Abbildung 7.7 (a)

¹⁷ <http://eu.levi.com/>



(a) GeoTagging mit eigenen Bildern als räumliche Entscheidung. (b) Geografische Empfehlung in der Facebook Anwendung GeoRecommender.

Abbildung 7.8: GeoRecommender als Facebook Applikation.

zeigt die Oberfläche, die für diesen ersten Schritt implementiert wurde, wobei auch eine andere Besonderheit der Facebook-Plattform zu sehen ist.

Nachdem der Nutzer Bilder aus einem Album verortet hat, berechnet die Anwendung den Aufnahmeort der Bilder (verwendet die Möglichkeit des Geocoding über die GoogleMaps API) und bietet dann dem Nutzer an, diese Information in seinem persönlichen Informationsstrom auf seiner Startseite anzuzeigen. Dies ist eine Möglichkeit andere Nutzer, die diesem Informationsstrom folgen, auf die verorteten Bilder aufmerksam zu machen und ihnen einen Link anzubieten mit dem sie die Bilder auf einer Karte betrachten können. Dadurch lassen sich zum Beispiel Bilder, die während einer Wanderung entstanden sind, auf einer Karte nachverfolgen. Diese Nachrichten soll andere Benutzer motivieren, selbst Bilder an diesem Ort einzustellen. In dem Beispiel in Abbildung 7.7 (a) sieht man wie die eigenen Freunde aufgefordert werden, selbst Bilder im Gebiet Potsdam einzupflegen. Dadurch kann eine

schnelle geografische Überlappung unter den Nutzern des Systems entstehen.

Der zweite logische Schritt ist die Berechnung von neuen räumlichen Empfehlungen für ein anderes Gebiet. Dazu wechselt der Nutzer auf eine andere Kartenansicht und wählt einen beliebigen Kartenausschnitt aus, für die er die Empfehlung berechnen lassen will. Eine personalisierte Empfehlung kann allerdings nur berechnet werden, wenn zuvor Bilder über die Facebook-Anwendung verortet wurden. Ist das der Fall versucht das System zunächst, wieder eine persönliche Empfehlung zu berechnen und falls dies scheitert, wird versucht, die Aufnahmen von den typischsten Fotografenstandpunkten zu finden. Als zusätzliche Quelle für räumliche Entscheidungen werden dazu die Bilder aus Panoramio verwendet, um der Anwendung schon bei möglichst wenigen Informationen von Facebook-Nutzern Empfehlungen generieren zu können. Abbildung 7.7 b) zeigt das Ergebnis einer persönlichen Empfehlung für das Innenstadtgebiet von Bamberg. Diese Empfehlung basiert hauptsächlich auf ähnlichen Benutzern aus Panoramio, lediglich die Auswahl der Bilder wird auf die Bilder von Facebook-Nutzern eingeschränkt.

Ausgehend von der Einbettung in das soziale Netzwerk lassen sich noch weitere interessante Anknüpfungspunkte für ein geografisches Empfehlungssystem finden. So ist eine Funktionalität der Bildapplikation, die Möglichkeit, Nutzer auf Bildern zu markieren. Dies bedeutet dementsprechend, dass sich die jeweiligen Personen in der Nähe des Fotografen aufgehalten haben. Dadurch können neue räumliche Entscheidungen entstehen ohne dass ein Nutzer selbst Bilder erstellen muss. Dadurch würde man allerdings diese Verlinkung auf einem Bild auf die gleiche Weise, wie das Erstellen eines Bildes messen, beide würden gleichbedeutend in das Popularitätsmaß einfließen. Durch eine derartige Gleichstellung würden Events, an denen mehrere Personen beteiligt sind, wie-

der recht schnell eine hohe Popularität erhalten. Um dem entgegen zu wirken müsste man diese Verlinkungen daher anders als seither in die Popularitätsberechnungen eines CPVs mit aufnehmen.

Ein weiterer Anknüpfungspunkt für die zukünftige Erweiterungen des Empfehlungssystems ist das Wissen über explizite Freundschaftsbeziehungen, die oft auch mit einer Vertrauensbeziehung gleichgesetzt werden; so wird ein Nutzer einer Empfehlung eventuell mehr vertrauen. Momentan wird nur die Auswahl eines repräsentativen Bildes für eine räumliche Entscheidung durch eine Freundschaftsbeziehung beeinflusst. Stehen für einen Ort mehr als ein Bild zur Verfügung, wählt das System bevorzugt Bilder aus, die von Freunden erstellt wurden. Weitere Möglichkeiten, diese zusätzlichen Informationen aus sozialen Netzwerken für eine Empfehlung zu nutzen, wird an dieser Stelle allerdings zukünftigen Arbeiten überlassen.

7.3.3 Auswertungstool GeoRec in Java

Während der Entwicklung der Methode und den Metriken wurde es notwendig bestimmte Teilschritte zu visualisieren, angefangen von der Ausgangsmenge bis hin zu den Ergebnissen der Empfehlung. Insbesondere da es sich bei Bildern um ein rein visuelles Medium handelt, macht es Sinn, die Ergebnisse bzw. die ausgewählten Bilder auch visuell darzustellen und dadurch die Empfehlung zu überprüfen.

Das Programm bietet dem Nutzer eine Kartenansicht als grundlegendes Bedienelement an. Die Karte weist die Standardfunktionen eines GIS-Systems auf, das heißt man kann die Karte verschieben und zoomen. Zusätzlich lassen sich auf der Karte für bestimmte Zwischenschritte der Empfehlungsberechnung mehrere Overlays anzeigen. Die Anwendung ermöglicht es, zunächst den Fokus auf ein bestimmtes Gebiet aus der verwendeten Partonomie zu setzen. Ausgehend von diesem Gebiet

können nun die Daten weiter untersucht werden. Wichtig für die Berechnung der Empfehlungen sind die Standpunkte der Fotografen, die in den Metadaten der Bilder zu finden sind. Um die Standpunkte der einzelnen Nutzer darzustellen, können über eine Liste von Nutzern diejenigen selektiert werden, für die die Positionen ihrer Bilder angezeigt werden sollen. Da auch die Positionen der Bilder, die außerhalb des fokussierten Gebietes liegen, angezeigt werden, bekommt man schnell einen Überblick über die Gebiete, wo ebenfalls vermehrt Bilder aufgenommen wurden.

Für das gewählte Gebiet kann die entsprechende Heatmap bestimmt werden und über die Anzeige von entsprechend eingefärbten Kreisscheiben sind die Gewichtung der einzelnen Cluster auf der Karte sichtbar. Im gleichen Stil wird auch die Gewichtung durch eine persönliche Empfehlung angezeigt. Wurde durch das System eine Empfehlung berechnet, lassen sich die Nachbarschaftsbeziehungen in einem Graphen darstellen. Die ausgewählten nächsten Nachbarn werden dabei als blaue Knoten angezeigt, und die Städte, an denen sie ähnliche Bilder aufgenommen haben, als grüne Knoten, wie in Abbildung 7.9 B zu sehen. Eine Kante besteht, wenn ein Nutzer in einer Stadt ein Bild aufgenommen hat. Die Stärke der Kante drückt dabei die Ähnlichkeit zum Initiator der Empfehlung aus. Wählt man einen Nutzerknoten aus diesem Graphen aus, kann man auf der Karte wiederum rekonstruieren an welchen Stellen er seine Bilder aufgenommen hat und kann so visuell nachvollziehen wie die letztendliche persönliche Gewichtung durch die nächsten Nachbarn zustande gekommen ist.

Die Bildempfehlungen selbst lassen sich, unabhängig davon ob sie durch die Heatmap oder den kollaborativen Filter berechnet wurden, in einer Bilderleiste einsehen, wobei die Nutzerkennung und die Bildquelle mit angegeben werden (Abbildung 7.9 E).

Es können eine Reihe von zusätzlichen statistischen Informationen berechnet werden, von denen einige in Abbildung 7.9 zu sehen sind. Es lässt sich zum Beispiel die Verteilung der Bilder im Verhältnis zu der Rangfolge der räumlichen Entscheidungen anzeigen, Abbildung 7.9 C. Die Darstellung nennt man auch eine Lorenzkurve und gibt an, wieviele der Bilder man von einem Benutzer bereits gesehen hat, wenn man die räumlichen Entscheidungen absteigend nach ihrer Popularität besucht, wobei die Achsen des Schaubildes jeweils in Prozent angegeben sind. In diesem Beispiel würde man demnach bereits 75% der Bilder gesehen haben, wenn man 5% der populärsten Standpunkte in Bamberg besucht hat. Die Darstellung einer Lorenzkurve wird häufig in der Ökonomie verwendet, um die Konzentration von wirtschaftlicher oder politischer Macht zu messen. Insbesondere wird damit die Verteilung von Geldmitteln auf die Bevölkerung dargestellt und diesbezüglich ein oft recht großes Ungleichgewicht unter der Bevölkerung aufgezeigt. Die erste Winkelhalbierende gibt dabei eine Gleichverteilung an. Orientiert sich die Kurve eines Nutzers an der ersten Winkelhalbierenden, verteilen sich seine Bilder gleichmäßig von den typischsten bis zu den untypischen fotografischen Standpunkten (Abbildung 7.9 C). Mit Hilfe dieser Funktion lässt sich der Gini-Koeffizient berechnen, der die Ungleichverteilung in einer Maßzahl ausdrückt. In Falle der Verteilung der Standpunkte lässt sich am Gini-Koeffizienten ablesen, ob ein Nutzer eher nur typische Aufnahmen aufnimmt oder tendenziell sehr individuelle Standpunkte für seine Bildaufnahmen wählt. Liegt der Gini-Koeffizient unter 1 begnügt sich der Nutzer mehr mit einer Vielzahl von Aufnahmen von typischen Orten; liegt er zwischen -1 und 0 verteilen sich die Standpunkte sehr gleichmäßig, liegt er zwischen 0 und 1 hat der Nutzer eine Vielzahl von Bildern auch an sehr untypischen Orten aufgenommen.

Ansonsten lassen sich noch die typischen Bewegungen der Nutzer anzeigen, welche auf einer einfachen Kookkurenzanalyse beruhen. Berechnet

wird dabei der Faktor aus Anzahl der Nutzer, die beide Orte besuchen, im Verhältnis zu der Anzahl von Nutzern, die nur den Zielort besuchen. Um die Ausgangsmenge von Nutzern mit zu beachten, wird oft die Anzahl der Nutzer am Ausgangsort zum Nenner multipliziert. Meistens wird die Anzahl des Ausgangsort mit einem Faktor $0 < a < 1$ potenziert um den Einfluss dieses Faktors zu schwächen. Durch dieses Verfahren, wird der Einfluss von sehr populären Orten reduziert, die mit höchster Wahrscheinlichkeit Ausgangspunkt von jeder Bewegung sind. Berechnet werden diese Kookkurrenzen sowohl auf der Ebene der räumlichen Entscheidungen (rote Pfeile in Abbildung 7.9 A) als auch auf den höheren Ebenen der Partonomy (rote Kanten in Abbildung 7.9 D).

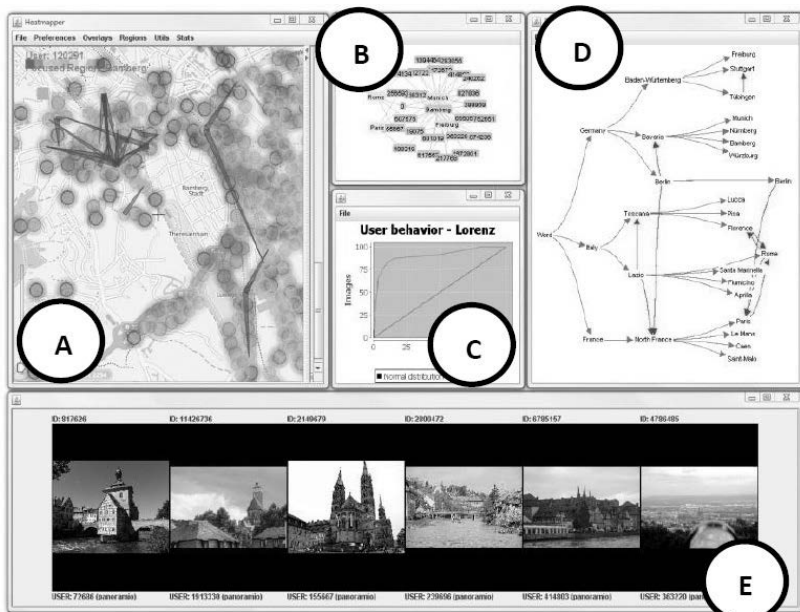


Abbildung 7.9: Java Anwendung GeoRec

7.4 Erweiterungspunkte des Empfehlungssystems

Die Codebasis des Empfehlungssystems ist an einigen Stellen um neue Datenquellen oder Algorithmen erweiterbar. Zunächst ist es über die zuvor angesprochene Mediatoren Architektur der Datenschicht möglich, neue Datenquellen für räumliche Entscheidungen bzw. verortete Bilder zu verwenden. Für die Einbettung in Facebook musste zum Beispiel lediglich ein neuer Translator geschrieben werden, der auf die verorteten Bilder der Anwendung zugreifen konnte. Zudem konnte das Empfehlungssystem so konfiguriert werden, dass er diese Quelle als Bildquelle für die Empfehlungen verwendet.

Ein weiterer wichtiger Erweiterungspunkt ist die verwendete Partonomy. Für den ersten Prototypen wurde eine recht einfache Partonomy gewählt, die die Polygone über der meisten größeren europäischen Städte beinhaltet und daher ein Patchwork nach Definition in Kapitel 6 darstellt. In Zukunft könnten bestehende hierarchischen Raumunterteilungen, wie den NUTS Gebieten, diese einfache Partonomy ersetzen.

Es wurde weiterhin darauf geachtet, dass alle wichtigen Teilschritte der Algorithmen und die darin enthaltenen Maße aus Kapitel 5 und 6, soweit es möglich ist, modularisiert wurden. Dadurch können Details in den Verfahren in Zukunft einfach ersetzt werden können. Ein Beispiel ist die Berechnung der Benutzerähnlichkeit, bei der die Kombination der drei Teilmengen durch ein alternatives Kombinationsverfahren zu dem Tanimot-Verfahren ersetzt werden kann.

8 Auswertung

Die Genauigkeit der beiden Ansätze zur Berechnung von typischen und personalisierten Bildgruppen einer Region werden im Folgenden durch mehrere Untersuchungen evaluiert. Dabei werden die berechneten Bildgruppen mit bestehenden verglichen. Die Stichproben dieser Bildgruppen ergeben sich dabei aus folgenden Ausgangsmengen:

- **Gedruckte Patchwork-Bild-Postkarten:** Für eine Auswahl von europäischen Städten wurden gedruckte Patchwork-Bild-Postkarten gesammelt. Durch die Häufigkeit der Sehenswürdigkeiten auf den Postkarten entsteht eine Rangfolge, die mit der Ordnung durch das Popularitätsmaß vergleichbar ist. Die Top 20 aus beiden Rangfolgen wird mit Hilfe von Precision und Recall verglichen - Abschnitt 8.1.
- **Durch Nutzer manuell ausgewählte Bildgruppen für Patchwork-Bild-Postkarten:** In dieser Onlinestudie wurden verschiedene Personen gebeten, für Paris und Bamberg Bildpostkarten zu erstellen. Anschließend wurde getestet, wie sich Nutzerähnlichkeiten in unterschiedlichen Städten verhalten. Für die Erhebung wurde eine Webanwendung mit dem Namen Tripost implementiert - Abschnitt 8.2.
- **Bilder aus geografischen Bildsammlungen aus unterschiedlichen Städten:** Der Unterschied zu den Bildern aus der Postkartenstudie bestand darin, dass der Ersteller dieser Bilder sich tatsächlich im Raum bewegen musste, um dann selbst ein Bild an diesem Ort aufzunehmen. Diese festgehaltenen räumlichen Entscheidungen

der Nutzer wurden mit den berechneten Empfehlungen verglichen. Gerade für diese Art der Nutzerentscheidungen lieferte das vorgestellte räumliche Empfehlungsverfahren deutlich bessere Ergebnisse als klassische Empfehlungsverfahren - Abschnitt 8.3.

Da es sich bei Empfehlungen um den Versuch handelt, dem persönlichen Geschmack eines Nutzers gerecht zu werden, ist eine Evaluation dieser Systeme nicht trivial, zumal im Grunde nur der Nutzer selbst die Güte einer Empfehlungen bewerten kann. Auf die Schwierigkeit bei der Evaluation von Empfehlungssystemen wird im Abschnitt 8.3 eingegangen. Im selben Abschnitt werden auch die gängigen Evaluationsmaße vorgestellt.

8.1 Vergleich typischer Bildgruppen zu Patchwork-Postkarten

Um einen objektiven Eindruck der Qualität des Heatmapverfahrens zu erhalten, wurden die identifizierten Sehenswürdigkeiten, die von den populärsten Standpunkten aus fotografiert wurden, mit den Objekten verglichen, die auf gekauften Patchwork-Postkarten zu sehen sind. Das heißt, es wurde für die Städte, aus denen im Vorfeld Postkarten gesammelt wurden, eine Heatmap mit Fokus auf den Stadtkern berechnet. Aus dieser Heatmap wurden dann die ersten 20 CPVs für den Vergleich ausgewählt. Die Ergebnisse entsprechen für Amsterdam, Bamberg, Cardiff und Dublin den Ergebnissen, die in Kapitel 5 diskutiert wurden. Daher wurden die ersten 20 Ergebnisse aus der Heatmap gewählt, da im Durchschnitt insgesamt 20 verschiedene Sehenswürdigkeiten auf den Postkarten abgebildet waren, siehe Tabelle 5.1. Dadurch ließen sich 20 berechnete Orte mit den ungefähr 20 möglichen Orten aus den Postkarten vergleichen.

Analysiert wurde, wie viele der repräsentativen Ansichten aus den 20 CPVs ebenfalls auf den Postkarten zu finden waren. Dieser Wert entspricht dem Gütemaß *Precision*. Ergänzend dazu wurde berechnet, wie viele der Sehenswürdigkeiten auf den Postkarten in den ersten 20 CPVs zu finden waren, dieses Maß wird auch *Recall* genannt. Im Normalfall werden sowohl Precision als auch Recall berechnet, da beide Werte abhängig voneinander sind. Es ist leicht möglich einen Wert auf Kosten des anderen zu erhöhen. Beide Werte zusammen ermöglichen eine Aussage, ob die Ergebnisse aus der Top 20 sowohl relevant, als auch vollständig, im Vergleich zu den Bildern auf den Postkarten, sind. Beide Werte liegen zwischen 0 und 1, wobei ein Wert von 1 dem besten Ergebnis entspricht.

Aus den Ergebnissen aller acht Regionen wurde das Mittelmaß von Precision und Recall berechnet. Als Endergebnis lag der Recall Wert durchschnittlich bei 0,44 und Precision bei 0,92. Das heißt, beinahe jedes Objekt aus den Top-20-Listen wurde auch von den Erstellern der Bildpostkarten als typisch für die jeweilige Stadt angesehen. Auf der anderen Seite kann man sagen, dass nicht alle Objekte, die auf Postkarten zu sehen sind, von der Heatmap in deren oberen Ränge aufgenommen werden. Das liegt unter anderem an Duplikaten, die bei der Heatmapanalyse der POVs auftreten können. Ein prominentes Beispiel findet man im Datensatz von Paris. Der Eiffelturm nimmt gleich 5 Plätze in der Top 20 von Paris ein und nach der vorherigen Definition wird dasselbe Objekt nicht doppelt in die Berechnung des Recalls mit aufgenommen. Die Aufnahmen des Eiffelturms stammten von sehr weit auseinanderliegenden Orten, wobei der populärste Ort für eine Aufnahme direkt am Eiffelturm lag, der zweitpopulärste lag 500 Meter entfernt im nord-westlich gelegenen Palais de Chailot und der drittpopulärste Standpunkt, um den Eiffelturm zu fotografieren, befand sich im Park du Champs de Mars, der sich hinter dem Eiffelturm nach Südosten

erstreckt. Alle Duplikate haben daher eine sehr deutliche räumliche Trennung, so dass bei diesen Duplikaten eine andere Perspektive auf den Eiffelturm gewählt wurde. Ein weiterer Grund für den niedrigen Recall ist auch, dass bestimmte typische Sichtwinkel nicht von jeder Person eingenommen werden können. So findet man in den Postkarten aus Bamberg eine sehr typische Sicht, die von einem Turm aus aufgenommen wurde, zu dem nur sehr wenige Personen Zutritt bekommen. Die restlichen Sehenswürdigkeiten nehmen in den Ranglisten der Heatmap Rangplätze über 20 ein. Tabelle 8.1 fasst die Precision- und Recall-Werte der einzelnen Städte zusammen. Zusätzlich wurde auch das F1-Maß mit angegeben, dass ein gewichtetes Mittel von Precision und Recall bestimmt. Im Falle von F1 (Rijsbergen, 1979) werden dabei Precision und Recall gleichgewichtet. Der Durchschnitt des F1-Maßes liegt bei 0,59.

	Warwick	Cardiff	Barcelona	Paris	Potsdam	Edinburgh	Bamberg	Freiburg	Average
Precision	0,75	0,86	0,92	1,00	0,91	1,00	0,91	1,00	0,92
Recall	0,14	0,38	0,52	0,45	0,38	0,67	0,45	0,56	0,44
F1	0,23	0,52	0,67	0,62	0,54	0,80	0,60	0,71	0,59

Tabelle 8.1: Precision und Recall für den Vergleich von Heatmap und Postkarten

Abschließend sind die Ergebnisse gut genug, um von einer sinnvollen Auswahl an Bilder durch die Heatmap zu sprechen, die das typische Konzept einer geografischen Region mit Bildern illustrieren kann. Selbst bei einer zufälligen Bildauswahl aus den Bildclustern an den stereotypischen Fotografenstandpunkten, können wir davon ausgehen, dass typische Ansichten ausgewählt werden. Bei dem Vergleich mit den ge-

druckten Patchwork-Bild-Postkarten wurde von den Postkartenherstellern festgelegt, welche Bilder typisch für eine Stadt sind.

8.2 Tripost

Das vorgestellte Empfehlungsverfahren, wie auch alle anderen kollaborativen Filtersysteme, sind stark davon abhängig, wie gut eine Prädiktion über ähnliche Nutzer in einem neuen Kontext möglich ist. Aus diesem Grunde wurde eine Vorstudie umgesetzt, die untersuchte, ob sich ähnliche Vorlieben von zwei Nutzern in einer Stadt auch in einer zweiten Stadt beobachtbar sind. Dazu wurde untersucht wie stark sich die Nutzergemeinschaft eines Nutzers in beiden Städten veränderte. Im besten Fall sollten nächste Nachbarn eines Nutzers in einer anderen Stadt nächste Nachbarn bleiben. Der Versuch wurde als Webapplikation umgesetzt, die unter www.kinf.wiai.uni-bamberg.de/tp5/experiment/ erreichbar ist. Sie trägt in Anlehnung an das Projekt Tripod, in dessen Zusammenhang viele Teile dieser Arbeit entstanden sind, den Namen Tripost. Zu Beginn des Experiments wurden 25 Personen persönlich angeschrieben, um ihnen die Aufgabenstellung zu erläutern. Ein Teil der Aufgabe war es, die vier erstellten Patchwork-Bild-Postkarten als E-Mail an weitere Personen zu senden, mit der Bitte, sich ebenfalls an dem Versuch zu beteiligen. Durch dieses Schneeballverfahren wurden 101 Personen angeschrieben. 39 Personen erstellten vier vollständige Postkarten und lieferten dadurch auswertbare Ergebnisse.

8.2.1 Aufgabenstellung

Um eine Postkarte zu erstellen, wählten die Teilnehmer jeweils sechs Bilder aus einer Menge von ungefähr 100 Bildern für Bamberg und Paris aus. Für die beiden Aufgabenstellungen in einer Stadt wurden dieselben Bilder verwendet. Sowohl die Bilder als auch deren Sortierung blieben

Step 1: Create a postcard containing typical views in Bamberg

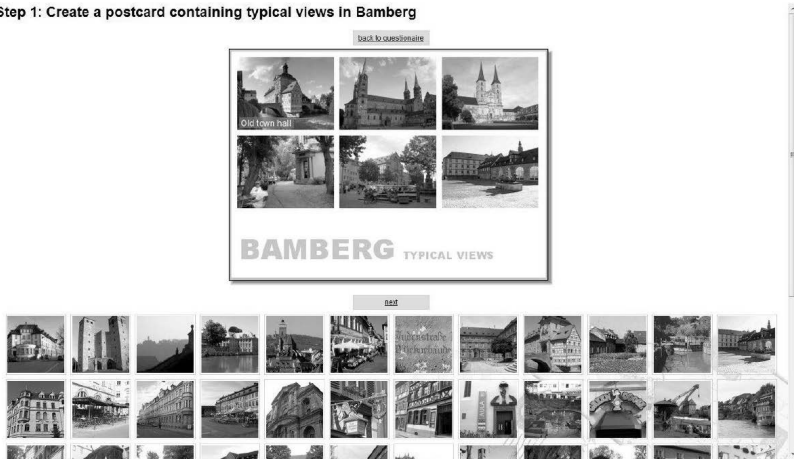


Abbildung 8.1: Zusammenstellung einer Postkarte in TripPost.

für den Teilnehmer unverändert. Damit wurde sichergestellt, dass für alle Testpersonen die gleiche Ausgangsbedingung hatten. Ein Screenshot der Oberfläche ist in Abbildung 8.1 zu sehen. Die Bilder wurden im Vorfeld gemischt und es lässt sich beobachten, dass auf den Postkarten nicht nur aus den ersten n Bildern zu finden sind, sondern auch Bilder ausgewählt wurden, die sehr spät in der Sortierung zu finden sind. Die Nutzer haben sich daher durchaus Zeit genommen, alle Bilder nach der besten Auswahl zu durchsuchen. Die folgende Übersicht zeigt die Bilder aus Bamberg in der Reihenfolge, wie sie den Nutzern in der Anwendung präsentiert wurden. Die Zahl vor dem Bild gibt die Anzahl der Postkarten an, auf denen das jeweilige Bild zu finden ist. Es lässt sich keine Korrelation zwischen Sortierreihenfolge und Auswahlhäufigkeit feststellen.



Die Aufgabenstellung beinhaltete ebenfalls, dass die Auswahl der Bilder dem jeweiligen Bildtitel der Postkarte gerecht werden sollte. Dieser

Titel war in Überschrift und Unterschrift aufgeteilt, wobei die erste Ebene den Namen der jeweiligen Stadt ausmachte und die zweite Ebene die Auswahl der Bilder beschrieb. Folgende Untertitel wurden für die Postkarten gewählt:

- *TYPICAL VIEWS*. Bei dieser Postkarte sollten die Teilnehmer Bilder auswählen, die ihrer Meinung nach die typischsten Ansichten der jeweiligen Stadt zeigen. Hier waren ähnliche Ergebnisse zu erwarten, wie wir sie schon bei den offiziellen touristischen Patchwork-Postkarten gesehen haben.
- *MY FAVORITE PLACES*. Auf dieser Postkarte sollten die Nutzer ihre eigenen Lieblingsplätze auswählen. Es war zu erwarten, dass gerade Menschen, die bereits in den Städten gelebt haben oder dort noch leben, nicht nur die bekannten touristischen Orte auswählen werden.

Aus diesen Vorgaben für die Postkarte ergeben sich vier verschiedene Szenarien, für die der Benutzer eine Postkarte erzeugen sollte: Bamberg/Typical, Bamberg/Personal, Paris/Typical und Paris/Personal. Im ersten Anschreiben wurde eine Link auf die Webseite versendet, auf der sich der Nutzer selbst anmelden konnte. Zusätzlich wurde für jeden Empfänger einer E-Mail, der noch nicht Nutzer der Anwendung war, ein neuer Account angelegt und ein Link versendet, über den er sich mit einem Klick in das System einloggen konnte. Dadurch war eine schnelle und unkomplizierte Anmeldung für die angeschriebene Personen möglich. Die Anschreiben, wie auch die Oberfläche, waren in deutscher, englischer und französischer Sprache verfügbar.

Nachdem sich eine Testperson eingeloggt hatte, wurden ihm/ihr zu Beginn des Experiments ein paar wenige Fragen zu seiner/ihrer Person gestellt, um zumindest abschätzen zu können, wie vertraut die jeweilige Person mit den beiden Städten war. Es wurden dabei Menschen unterschieden, die schon einmal in der Stadt gelebt haben oder nur als

Männ- lich	Weiblich	Bam- berg - lebt in	Paris - lebt in	Bam- berg - hat besucht	Paris - hat besucht
15	24	24	10	4,0256	5,0256

Tabelle 8.2: Persönliche Angaben der Testpersonen

Tourist die Stadt besucht hatten. Vor allem bei der Auswahl der Lieblingsplätze war zwischen diesen Gruppen ein deutlicher Unterschied zu erwarten. Folgende Fragen wurden den Testpersonen gestellt:

- Wie alt sind Sie?
- Welches Geschlecht haben Sie? (männlich/weiblich)
- In welcher Stadt haben Sie bereits gelebt? (Bamberg/Paris)
- Wie viele Wochen haben Sie als Tourist in der jeweiligen Stadt verbracht?

Danach wurde er/sie gebeten, die zuvor beschriebenen vier Postkarten zu erstellen, um sie dann alle zusammen einer weiteren Person als E-Mail zu schicken. Im Idealfall sollte der Empfänger der Nachricht eine Person sein, die in der jeweiligen anderen Stadt wohnt beziehungsweise jemand, der diese Stadt besser kennt als man selbst. Das System versuchte dabei, doppelte Anfragen an dieselben Nutzer zu vermeiden, in dem es den Nutzer darauf hinweist, wenn der Empfänger bereits von einer anderen Person kontaktiert wurde. Tabelle 8.2 gibt eine Übersicht über die Teilnehmer der Tripot-Studie. Abbildung 8.2 zeigt den E-Mail-Kontakt zwischen den Nutzern der Anwendung. Es ist gut zu erkennen, dass die meisten Empfänger keine weitere Person kontaktierten und selbst nicht aktiv am Versuch teilgenommen hatten (oktagonförmige Knoten).

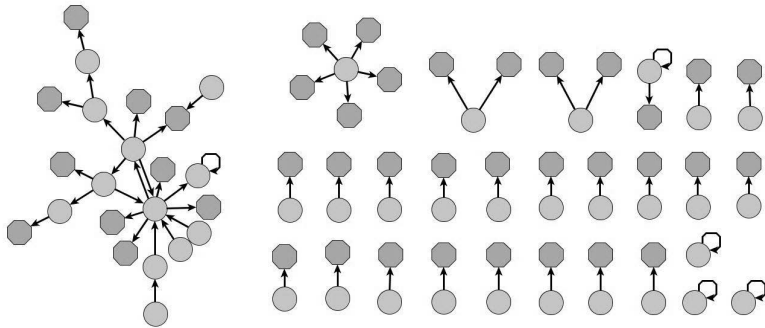


Abbildung 8.2: E-Mailverkehr durch Tripost. Jeder Knoten steht für einen Nutzer der Anwendung. Eine Kante symbolisiert den E-Mail-Kontakt zwischen zwei Nutzern. Ein Oktagon steht für einen Nutzer, der nicht am Versuch teilgenommen hat.

8.2.2 Ausgangsfragen

Basierend auf den Ergebnissen des vorgestellten Versuchs wurden die folgenden Fragestellungen untersucht.

Hypothese 1: Die Nutzer, die sich in einer Stadt ähnlich sind, sind sich auch in der anderen Stadt ähnlich.

Weiterhin wurden folgende Behauptungen untersucht:

Hypothese 2: Die Postkarten mit den typischen Ansichten einer Stadt zeigen vermehrt dieselben Aufnahmen, während die persönlicheren Postkarten ganz verschiedene Facetten der Stadt zeigen.

- Für die typischen Bilder ist eine Longtail-Verteilung zu erwarten. Die Häufigkeitsverteilung wird sich einer Potenzfunktion annähern.

- Für die Bilder der persönlichen Lieblingsplätze sollten sich die Häufigkeiten der Bilder gleichmäßiger und auf deutlich mehr Aufnahmen verteilen. Die Kurve der Häufigkeitsverteilung sollte daher flacher verlaufen, als die Kurve der Häufigkeitsverteilung der typischen Aufnahmen.

Hypothese 3: Besucher einer Stadt sind sich in ihren Entscheidungen ähnlicher als die Einwohner der Städte.

8.2.3 Methoden

Zunächst wurden aus dem Datensatz alle Personen entfernt, die nicht für alle Postkarten sechs Bilder ausgewählt hatten. Um die Ähnlichkeit von zwei Nutzern zu bestimmen, wurde als einfaches Maß die Anzahl der Bilder bestimmt, die in beiden Postkarten verwendet wurden. Enthalten die Postkarten von zwei Nutzern für ein bestimmtes Nutzungsszenario ein gemeinsames Bild, erhielten die Nutzer den Ähnlichkeitswert 1, bei zwei gemeinsamen Bildern den Ähnlichkeitswert 2 und so weiter. Aufgrund der Nutzerähnlichkeiten ergibt sich für jeden Nutzer und jedes Nutzungsszenario eine bestimmte Nutzergemeinschaft, wie sie in Kapitel 6 Abschnitt 6.2 definiert wurde. Bestätigt sich die Haupthypothese, müsste diese Gemeinschaft eines Nutzers relativ stabil bei einem Wechsel zwischen den Städten bleiben. Besonders interessant ist hierbei der Wechsel zwischen den Städten mit denselben Untertiteln, demnach von Bamberg/Typical nach Paris/Typical und Bamberg/Personal nach Paris/Personal. Gerade bei den persönlichen Vorlieben in beiden Städte ist es interessant zu sehen, ob hier ähnliches Nutzerverhalten bestehen bleibt. Zumindest sollte es deutlich beständiger bleiben als bei einer zufälligen Zusammenstellung, die wir als Baseline verwendet haben.

8.2.4 Baseline

Bei dem Versuchsaufbau besteht die Möglichkeit, dass bestimmte Beobachtungen zufällig auftreten. Daher wurde zunächst berechnet, wie wahrscheinlich es ist, dass zwei Personen dasselbe Bild aus einer Menge von 100 Bildern ziehen und dieses Ereignis auch bei einer davon unabhängigen zweiten Ziehung auftritt. Zusätzlich wurde eine zufällige Ziehung von Bildgruppen implementiert. Die Zufallspostkarten wurden mit den denselben Maßen, wie die Postkarten der Nutzer, analysiert und die Ergebnisse miteinander verglichen.

Für die Berechnungen dieser Baseline helfen Überlegungen aus der Kombinatorik. Die Wahrscheinlichkeit für die oben beschriebenen Beobachtungen können auf einfache Weise berechnet werden, wenn man annimmt, dass die Auswahl der Bilder rein zufällig zustande kommen. Man spricht dabei von einer Laplace-Verteilung, was bedeutet, dass jedes Bild mit derselben Wahrscheinlichkeit von $\frac{1}{100}$ für eine Postkarte ausgewählt wird. Das entspricht der gleichen Annahme, dass mit einem perfekten Würfel eine bestimmte Zahl mit einer Wahrscheinlichkeit von $\frac{1}{6}$ gewürfelt wird.

Zu überlegen ist nun, wie wahrscheinlich es ist, dass zwei Nutzer bei einer sechsmaligen Ziehung aus 100 Bildern mindestens ein gemeinsames Motiv auf ihren beiden Postkarten auswählen. Der erste Nutzer vergibt durch seine Auswahl sechs Bildern die Eigenschaft „ausgewählt“ und 94 Bildern die Eigenschaft „nicht ausgewählt“. Es ergibt sich eine dichotome Menge an Bildern. Es muss nun berechnet werden wie hoch die Wahrscheinlichkeit ist, dass mindestens k Bilder, mit der Eigenschaft „ausgewählt“, zufällig von einem zweiten Nutzer ebenfalls ausgesucht werden. Die Wahrscheinlichkeit dieser Ziehung lässt sich über eine hypergeometrische Verteilung berechnen (Irle, 2005). Die Wahr-

scheinlichkeit bei einer Lottoziehung k „Richtige“ kann zum Beispiel ebenfalls mit einer hypergeometrischen Verteilung bestimmt.

Nach der allgemeinen Definition einer hypergeometrischen Verteilung muss eine dichotome Gesamtmenge N vorliegen. Die Elemente dieser Menge können daher nur eine von zwei mögliche Eigenschaften aufzeigen, im Fall der Postkarten „ausgewählt“ oder „nicht ausgewählt“. Die Menge der Elemente mit der erwünschten Eigenschaft sei M . Die Fragestellung ist nun wie wahrscheinlich es ist, dass man k Elemente mit der gewünschten Eigenschaft erhält, wenn man n mal ohne zurücklegen aus der Gesamtmenge N auswählt. Die Ausgangssituation für die Berechnung mit Hilfe einer hypergeometrischen Verteilung lässt sich folgendermaßen zusammenfassen:

- Gesamtmenge N besteht aus 100 Bildern.
- Anzahl der Bilder M , die von dem ersten Nutzer bereits für eine Postkarte ausgewählt wurde ($M = 6$).
- Anzahl der Bilder n , die für eine Postkarte ausgewählt werden ($n = 6$).
- k ist Anzahl der Bilder, die in beiden Postkarten identisch sein müssen.

Die Formel für die hypergeometrische Verteilung lautet wie folgt:

$$P(k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}} \quad (8.1)$$

Die Formel für die Wahrscheinlichkeit, dass mindestens k Elemente in der Auswahl gleich sind ergibt sich aus der Summe der Einzelwahrscheinlichkeiten:

$$P(x \geq k) = \frac{\sum_{a=x}^M \binom{M}{a} \cdot \binom{N-M}{n-a}}{\binom{N}{n}} \quad (8.2)$$

Rechnerisch ergibt sich daraus eine Wahrscheinlichkeit von

$$P(x \geq 1) = \frac{\sum_{a=1}^6 \binom{6}{a} \cdot \binom{100-6}{6-a}}{\binom{100}{6}} \approx 0,3169. \quad (8.3)$$

Es werden demnach ungefähr 31% aller Nutzerpaare als ähnlich eingestuft, wenn dafür jeweils nur mindestens ein gleiches Bild auf ihren Postkarten zu finden sein muss. Geht man davon aus, dass die Nutzer zwei gleiche Bilder ausgewählt haben müssen, liegt die Wahrscheinlichkeit bereits bei ungefähr 4%, bei mindestens drei Bildern bei 0,2%.

Bei dem vollständigen Versuch stellt sich allerdings die Frage, ob Nutzer, die in einer Stadt ähnlich waren, dies auch in einer anderen Stadt sind. Daher müssen ihre Postkarten auch in dieser zweiten Stadt eine Ähnlichkeit aufweisen. Rechnerisch ergibt sich die kumulative Wahrscheinlichkeit aus dem Produkt der Einzelwahrscheinlichkeiten. Bei der Vorgabe, dass mindestens ein Bild gleich sein muss, liegt die Wahrscheinlichkeit bei $0,3169 \cdot 0,3169 = 0,1004$, also ungefähr bei 10%. Ein Zehntel der beobachteten Fälle werden sich bereits zufällig ergeben. Es ist nun also zu zeigen, dass sich deutlich mehr Fälle als die berechnete Baseline von 10% ergeben. Bei zwei Bildern liegt die Baseline bei 0,2%.

8.2.5 Ergebnisse der Analyse

Hypothese 1: Stabilität von Nutzergemeinschaften

Für einen weiteren Vergleich der Ergebnisse wurde eine weitere Baseline durch die Implementierung einer zufälligen Ziehung der Bilder eingeführt. Für jeden Teilnehmer des Experiments wurde zusätzlich eine zufällige Postkarte erstellt und die Evaluierung analog zu den nutzergenerierten Postkarten durchgeführt. Die Bilder auf den zufällig „gezogenen“ Postkarten sind nicht gleichverteilt, sondern folgen eher der Häufigkeitsverteilung, die man bei den Lieblingsplätzen beobachten kann.

Die Zufallsziehung bieten daher eine gute Vergleichsmöglichkeit zwischen den Ergebnissen. Abbildung 8.3 gibt einen Überblick über die Häufigkeitsverteilungen der Bilder in den einzelnen Postkartensets. Die x-Achse gibt den Rang des Bildes an und an der y-Achse kann man die Anzahl der Postkarten ablesen, auf denen das entsprechende Bild zu sehen war.

Bei den typischen Entscheidungen folgt die Kurve dem erwarteten exponentiellen Verlauf. Das Bestimmtheitsmaß R^2 zur angenommenen Potenzfunktion liegt, wie bei den vergleichbaren Beobachtungen, bei über 0,9. Bei den Lieblingsplätzen verläuft die Kurve sehr viel flacher und verhält sich recht ähnlich zu den willkürlichen Ziehungen der Zufallspostkarten. Einzige Ausnahme ist hierbei der Eiffelturm, der auch bei den Lieblingsplätzen in Paris eindeutig an erster Position liegt.

Da selbst bei der zufälligen Ziehung keine Gleichverteilung vorliegt, liegen die Prozentwerte der simulierten Nutzergemeinschaften höher als dem berechneten stochastischen Wert, der über das Geburtstagsphänomen berechnet wurden. Tabelle 8.3 vergleicht diese Ergebnisse mit den Ergebnissen des Nutzer-Experiments. Berechnet wurde eine Wahrscheinlichkeit von 14%, dass zwei Menschen bei jeweils sechs aus 100 Bildern, ein Bild gemeinsam haben. Bei der zufälligen Ziehung kam dieser Fall um die 25-29% der Fälle vor, wenn man fordert, dass ein Bild auf den Postkarten dasselbe sein muss. Bei den Nutzern der Anwendung liegt der Prozentsatz zwischen 57% und 63%.

Um die Hypothese zu bestätigen ist es nun interessant, ob die ähnlichen Nutzer auch nach dem Wechsel in eine andere Stadt konstant bleiben. Aus diesem Grund wurde analysiert, wie hoch der Anteil der gemeinsamen Nutzer in den Nutzergemeinschaften aus zwei verschiedenen Städten ist. Bei dieser Untersuchung sind nur Nutzergemeinschaften aus Set 2 (Lieblingsplätze aus Bamberg) und Set 4 (Lieblingsplätze aus Paris) relevant, da man hier von persönlichen Entscheidungen aus-

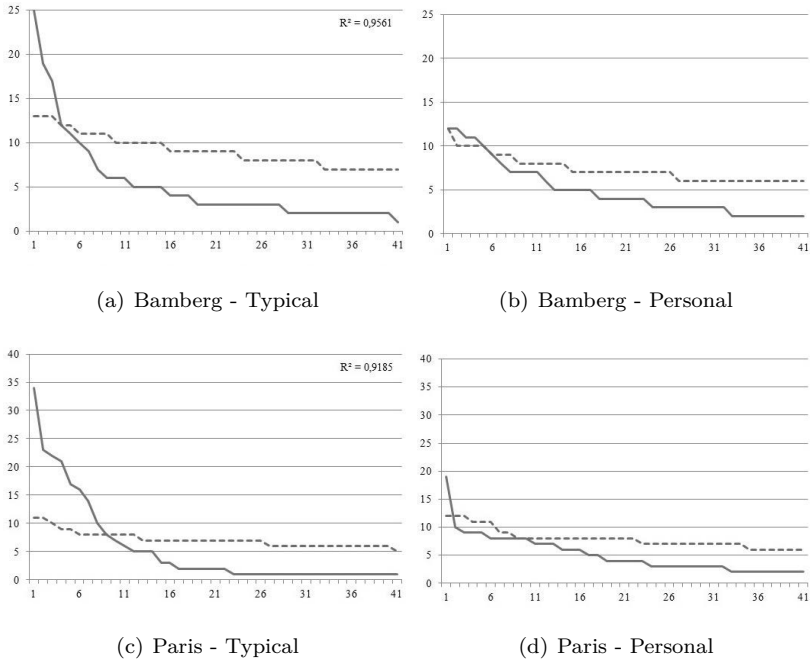


Abbildung 8.3: Häufigkeitsverteilung der einzelnen Bilder innerhalb der verschiedenen Sets. Gestrichelte Linie beschreibt die Häufigkeiten bei einer zufälligen Ziehung von Bildern.

Geforderte Ähnlichkeit	Set	Postkarten der Nutzer	Zufällige Postkarten
1	2	57,62%	25,51%
	4	63,29%	29,42%
2	2	17,27%	2,70%
	4	22,13%	4,18%
3	2	3,10%	0%
	4	3,29%	0,27%
4	2	0,27%	0%
	4	0,27%	0%

Tabelle 8.3: durchschnittliche Anteile der Nutzergemeinschaft an der gesamten Nutzerschaft bei steigender Mindestähnlichkeit

gehen kann, die man letztenendes über einen Empfehlungsansatz vorherzusagen will. Nach der Untersuchung dieser Postkarten ergab sich, dass ungefähr dreimal so viele Nutzer in beiden Nutzergemeinschaften dieselben waren, als das bei einer rein zufälligen Auswahl zu beobachten war. Im Vergleich zur Baseline waren es sogar mehr als vier mal so viele. Die Ergebnisse wurden in Tabelle 8.4 zusammengefasst.

Vergleichsets	Geforderte Ähnlichkeit	Postkarten der Nutzer	Zufällige Postkarten	Baseline
2 – 4	1	44,08%	15,45%	10%
2 – 4	2	13,26%	1,05%	0,2%
1 – 3	1	75,56%	16,43%	10%
1 – 3	2	39,59%	0,78%	0,2%
1 – 3	3	13,88%	0%	0%

Tabelle 8.4: Nachbarschaftsverhältnisse bei steigender Anzahl von Bilder, die übereinstimmen müssen, um zwei Nutzer als Nachbarn auszuzeichnen.

Verlangt man nun, dass mindestens zwei Bilder auf den Postkarten übereinstimmen müssen, damit zwei Nutzer eine Nachbarschaftsbeziehung eingehen, ergeben sich die Werte, wie in die Tabelle 8.4 zu sehen. In beiden Fällen ist ein deutlicher Unterschied, zwischen den simulierten Nutzergemeinschaften und den Nutzergemeinschaften, die durch das Nutzerverhalten während des Experiments entstanden sind, erkennbar.

Um die Hypothese weiter zu bestärken, wurde zusätzlich untersucht, ob die Nutzergemeinschaften sich nur auf der Popularität der einzelnen Bildern beruhen. Daher wurden die populärsten Bilder schrittweise für die Bestimmung der Ähnlichkeit ausgeschlossen. Im ersten Schritt waren sich Nutzer nicht mehr ähnlich, wenn sie nur das „Alte Rathaus“ oder den „Eiffelturm“ gemeinsam für ihre Postkarten ausgewählt hatten. Würde diese Vermutung sich bestätigen, müssten die Nutzergemeinschaften und ihre Überlappungen in Relation zu der Popularität der Bilder abnehmen: im Fall von Bamberg demnach um die 65% und in Paris um die 87%. In Abbildung 8.4 sieht man den Einfluss dieser Restriktion auf die Nutzergemeinschaften unter der Bedingung, dass mindestens ein Bild gleich sein musste. Abbildung 8.5 zeigt das gleiche Nutzungsszenario, allerdings unter der Bedingung, dass mindestens zwei gleiche Bilder ausgewählt wurden. Auf der x-Achse werden dabei die Anzahl der Bilder eingezeichnet, die in der Top x der Häufigkeitsverteilung zu finden sind und die für die Ähnlichkeitsbestimmung nicht mehr beachtet wurden.

Es ist zu beobachten, dass der Einfluss auf die Nutzergemeinschaften sehr viel geringer ausfällt, als die Popularität der Bilder vermuten lässt. So bilden sich Nutzergemeinschaften nicht nur ausschließlich wegen der Bilder, die in den Top-N der jeweiligen Stadt zu finden sind.

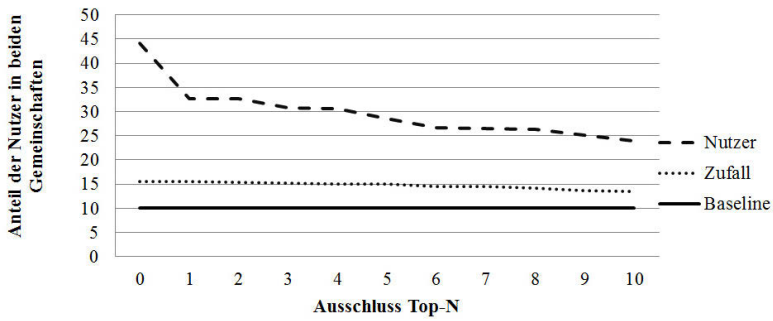


Abbildung 8.4: Stabilität von Nutzergemeinschaften bei mindestens einem gemeinsamen Bild.

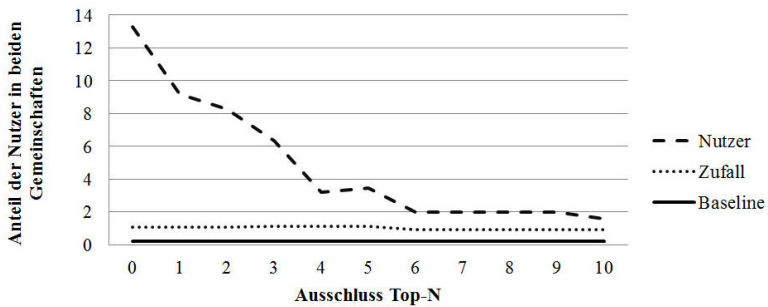


Abbildung 8.5: Stabilität von Nutzergemeinschaften bei mindestens zwei gleichen Bildern.

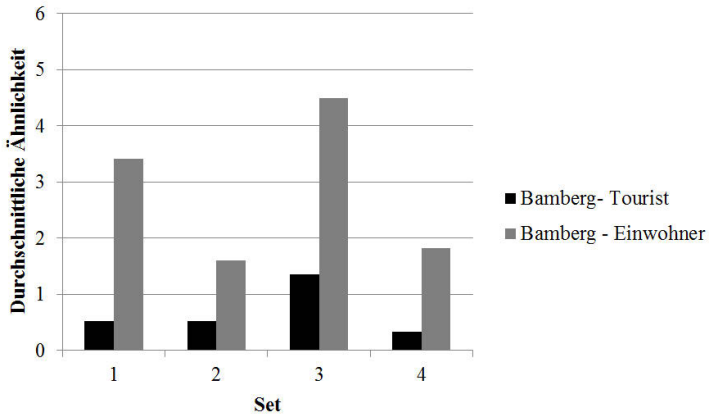
Hypothese 2: Auswahl der Bilder ist vergleichbar mit einer Auswahl durch die Heatmap

Wie bereits erwartet, ließ sich bei den Szenarien Bamberg/Typical und Paris/Typical eine Longtail-Verteilung feststellen. Die Häufigkeitsverteilungen beider Städte ließ sich wieder an eine Potenzfunktion mit einem sehr guten Bestimmtheitsmaß ($R^2 > 0,9$) angleichen. Interessant sind die Ergebnisse im Vergleich zu der Rangfolge der Heatmap. So beginnt in Bamberg die Verteilung, wie schon die Heatmap mit Altes Rathaus, Klein Venedig und Dom. Auch in den Postkarten von Paris stimmen die ersten 5 Ränge Plätze mit der Orten aus der Heatmap überein, allerdings in etwas anderer Reihenfolge, bis auf Rang 1 der eindeutig vom Eiffelturm eingenommen wird. Als Hypothese wurde weiterhin angenommen, dass bei den persönlichen Postkarten sich die Auswahl der Bilder auf deutlich mehr unterschiedliche Bildern verteilen würden. Für Bamberg war die Anzahl der unterschiedlichen Bilder bei der Auswahl für die persönlichen Postkarten um 10% (60/66 Bilder) höher, auf den Postkarten von Paris lag dieser Anteil bei 42% (45/64 Bilder).

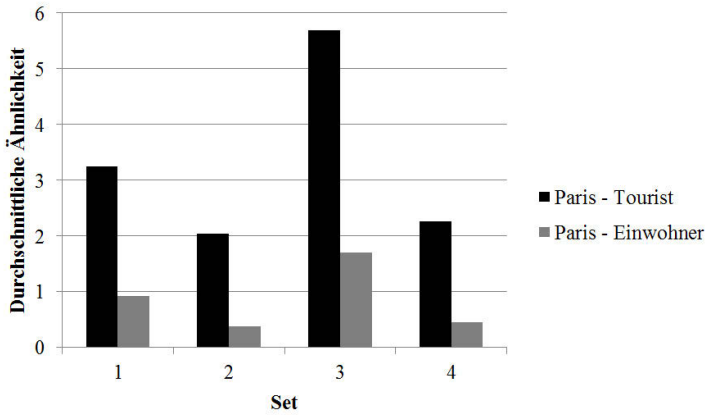
Hypothese 3: Touristen sind sich ähnlicher als Einwohner

Tabelle 8.5 gibt zunächst eine Übersicht über die Personen, die sich im Datensatz finden lassen, aufgeteilt in Tourist und Einwohner. Definiert wurde ein Einwohner als eine Person, die die Frage „In welcher Stadt haben Sie bereits gelebt?“ für die jeweilige Stadt mit ja beantwortet hat. Als Gast bzw. Tourist wurde jemand eingestuft, der noch nie in der jeweiligen Stadt gelebt hat, aber mindestens ein paar Tage dort verbracht hatte.

Die Abbildung 8.6 zeigt die durchschnittlichen Ähnlichkeiten der Nutzer in den beiden Städten auf, aufgeteilt in Touristen und Einwohner.



(a) Nutzer



(b) Zufall

Abbildung 8.6: durchschnittliche Ähnlichkeit von Einwohnern und Touristen in Bamberg und Paris

	Bamberg	Paris
Anzahl Einwohner	24	10
Anzahl Gäste	7	29

Tabelle 8.5: Anzahl Einwohner und Gäste in den Städten Paris und Bamberg.

Es lässt sich beobachten, dass sich in Bamberg die Einwohner ähnlicher sind, als die Gruppe der Touristen. In Paris tritt allerdings der umgekehrte Fall auf, hier sind sich die Touristen ähnlicher als die Einwohner. Bei den Touristen in diesen beiden Städten lässt sich daher nicht ableiten, dass sich Touristen immer gleich verhalten.

Eine wichtige Rolle spielt dabei sicher der unterschiedliche Charakter der beiden gewählten Städte. Paris ist mit einer Fläche von 105,40 Quadratkilometern, im Vergleich zu den 54,58 Quadratkilometern von Bamberg, deutlich größer und nicht mehr in einer Tagestour erkundbar. Außerdem verfügt Paris über einige sehr herausragende einzelne Gebäude, die sich vom restlichen Stadtbild deutlich abheben. Dahingegen verschmilzt Bamberg zu einer großen mittelalterlichen Stadtlandschaft, die man als Tourist auch ohne feste Ziele selbst erkunden kann um sich in den engen Gassen auch mal zu verlieren. Zwar hat Bamberg ebenfalls einzelne bedeutende Gebäude, die aber eher denjenigen Einwohnern bewusst sind, die sich selbst länger mit Bamberg beschäftigt haben. Mit seinen über 1500 denkmalgeschützten Häusern in der Altstadt (siehe Abbildung 8.7) wurde diese 1993 als schützenswerter Stadtteil in die Liste der Unesco Weltkulturerben aufgenommen. In Paris findet man hingegen wiederum nur die populären Einzelgebäude. Nicht zuletzt wird in Reiseführern, wie dem Lonely Planet - Travel Book (Planet, 2010), Bamberg als „perfect little small town“ beschrieben, in der man nicht einzelne markante Ziele besucht, sondern in der man sich „verlieren“ soll. Zusammenfassend lässt sich sagen, dass es Orte gibt,

die durch bestimmte Landmarken typikalisiert sind, wohingegen andere als spezifisch geprägte Stadtlandschaft wahrgenommen werden.

8.2.6 Ergebnis

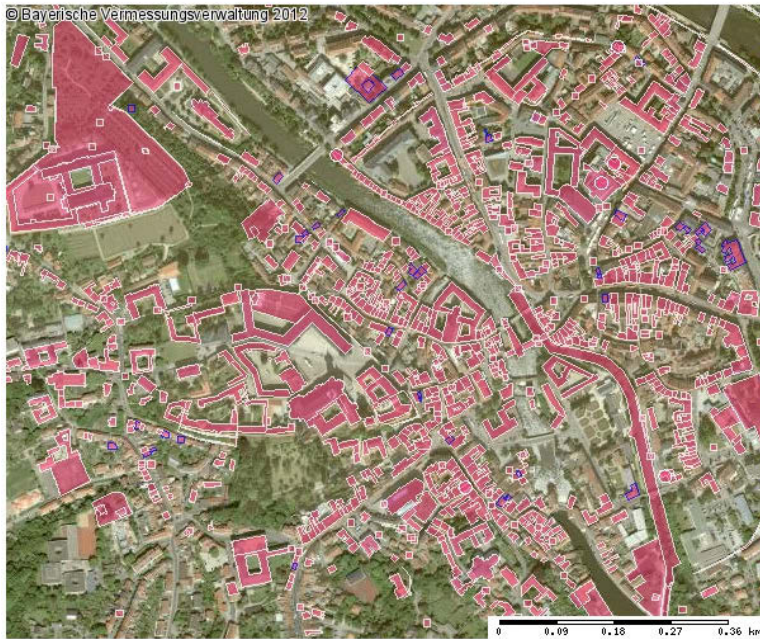
Die zu Anfang geforderte Konstanz von Nutzerähnlichkeiten, auch nach einem Ortswechsel, lässt sich zumindest in diesem recht einfachen Versuch bestätigen. Es ließ sich beobachten, dass Nutzergemeinschaften nicht ganz zufällig entstehen und sich dieselben Menschen auch in einer anderen Stadt ähnlich verhalten. Zusätzlich konnten auch weitere Eigenarten der unterschiedlichen Stadtlandschaften in den Datensätzen der beiden Städte gefunden werden. Diese sollen im Ausblick noch einmal aufgenommen werden, um weitere zukünftige Forschungsfragen aufzuwerfen.

Interessant für den Empfehlungsansatz ist nun, wenn wir uns auf Bilder beschränken, die durch die Bewegungen des Fotografen im Raum entstanden sind. Einige dieser Datensätze wurden bereits in Kapitel 3 vorgestellt und werden im nächsten Abschnitt dazu genutzt, den vollständigen Empfehlungsansatz zu evaluieren.

8.3 Offline-Evaluierung

Die Qualität einer Empfehlung zu bestimmen, ist keine triviale Aufgabe, da es selbst für den Benutzer schwierig einzuschätzen ist, ob ein Vorschlag für den Nutzer von Interesse ist. Hier spielen sehr subjektive Einschätzungen des Benutzers eine Rolle. Selbst Versuche, bei denen Benutzer die gleichen Objekte wiederholt bewerten mussten haben gezeigt, dass es zu einer deutlichen Varianz zwischen der ersten und der wiederholten Bewertung gibt (Herlocker u. a., 2004); kein Verfahren

Auszug aus dem BayernViewer



Legende

-  DOP (40cm)
-  Einzeldenkmäler (überprüft)
-  Einzeldenkmäler (nicht überprüft)

Abbildung 8.7: Denkmalgeschützte Häuser im Innenstadtbereich von Bamberg (rot umrahmt). Quelle: Bayerisches Landesamt für Denkmalpflege

wird aller Voraussicht nach besser sein, als diese natürliche Varianz von menschlichen Bewertungen.

Discovering the Inherent Variability in Recommender Datasets. We speculate above that algorithms trying to make better predictions on movie datasets may have reached the optimal level of error given human variability. Such variability can be explored using test-retest situations and analyses of taste change over time. If we can find effective ways to analyze datasets and learn the inherent variability, we can discover sooner when researchers have mined as much data as possible from a dataset, and thus when they should shift their attention from accuracy to other attributes of recommender systems. – Herlocker 2004

Eine Evaluation lässt sich auf zwei Arten durchführen: online und offline. Bei Online-Verfahren wird das System während dem eigentlichen Betrieb überprüft, zum Beispiel, indem der Nutzer die Gegenstände in seiner Empfehlung bewertet, nachdem sie ihm empfohlen wurden. Systeme wie Movielens oder auch Amazon bieten ihren Nutzern zum Beispiel an, Gegenstände jederzeit zu bewerten. Dadurch entsteht nicht nur neues Wissen für eine neue Empfehlung, sondern es werden auch indirekt die bereits berechneten Ergebnisse validiert. Alternativ sind auch kontrollierte Nutzer-Experimente möglich, welche die Nutzer zufällig in verschiedene Nutzungsbedingungen einteilt.

Eine andere Art der Evaluierung bieten Offline-Verfahren, die nur auf gespeicherte Transaktionen ausgeführt werden. Bieten die einzelnen Informationen im Evaluationsdatensatz einen Zeitstempel, kann eine normale Nutzungssituation auch nachgespielt werden und somit eine Onlinesituation simuliert werden. Die Evaluationsdaten bieten durch die räumliche Verortung die Möglichkeit, folgende Nutzungsszenario nachzuvollziehen.

1. Nutzer A besucht verschiedene Städte und macht dort mehrere Bildaufnahmen, die er zusammen mit den Aufnahmeorten dem

Empfehlungssystem mitteilt. Daraus ergeben sich räumliche Entscheidungen nach der Definition in Abschnitt 5.4.

2. Nutzer A lässt sich für Stadt S eine Empfehlung $ES(A)$ generieren, die eine Menge von räumlichen Entscheidungen beinhaltet.
3. Nutzer A besucht nun tatsächlich die Stadt A und macht dort wieder Bildaufnahmen, daraus ergeben sich die räumlichen Entscheidungen.

8.3.1 Vergleich und mit Hilfe eines passenden Evaluationsmaßes

Im Offline-Verfahren liegen die Bilddaten bereits komplett vor, so dass Schritt 2 und 3 vertauscht sind. Die Daten für Schritt 3 werden daher mit Hilfe ihrer geografischen Metadaten aus der Gesamtmenge der Bilder ausgefiltert.

Beide Evaluationsarten haben das Problem, dass keine Aussagen über Gegenstände getroffen werden können, die der Nutzer noch nicht bewertet hat. Im Allgemeinen werden diese Gegenstände als nicht relevant für den Nutzer festgelegt, was generell zu einer Unterschätzung der entsprechenden Evaluationsmaße führen wird. Eine Übersicht über die häufigsten Evaluationsmaße finden wir im nächsten Abschnitt.

8.3.2 Evaluationsmaße

Um die Genauigkeit und die Qualität von Empfehlungssystemen zu messen, wurde eine Vielzahl verschiedener Maße erarbeitet. Tabelle 8.6 gibt eine kurze Übersicht über die Maße, die in anderen Arbeiten zu nutzerbasierten Empfehlungssystemen zu finden sind, um die breite Masse an verschiedenen Evaluationen zu zeigen. Generell lassen sich die meisten Maße in zwei Arten einteilen: Maße, die die Genauigkeit einer Empfehlung bestimmen, und Maße, die die Qualität der Empfehlung

bestimmen. Gerade die Genauigkeitsmaße unterstützten die Entwicklung der Empfehlungssysteme, wie sie bereits in Kapitel 4 als Systeme mit fehlender Variabilität kritisiert wurde.

Genauigkeit

Empfehlungsansätze, die mit vorgegebenen Bewertungen r arbeiten, verwenden gerne statistische Genauigkeitsmaße, um die berechneten Bewertungen r_{pred} mit diesen zu vergleichen. In der Literatur findet man vor allem das Fehlermaß „Mean Absolute Error“ (MAE).

$$MAE = \frac{1}{n} \sum_{i=1}^n |r(i) - r_{pred}(i)| \quad (8.4)$$

Alternative dazu wird auch häufig der „Root mean square deviation“ oder auch „Root mean square error“ als Fehlerabschätzung verwendet. Dieses Maß bewertet im Vergleich zu MAE durch die Quadratbildung kleinere Fehler weniger stark, als größere.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (r(i) - r_{pred}(i))^2}{n}} \quad (8.5)$$

Vorhandene und berechnete Bewertungen lassen sich auch mit Hilfe von Korrelationsfaktoren miteinander vergleichen. Beispiele für Korrelationsfaktoren findet man in (Herlocker u. a., 2004), wobei er Pearsons Korrelationsfaktor als Stellvertreter für eine lineare Relation anführt, sowie Spearmans ρ und Kendalls Tau als Vertreter für die Korrelation der Ränge. Für alle drei Maße sind ähnliche Aussagen zu beobachten und je nach Situation können diese Maße alternativ verwendet werden. Im Falle des geografischen Empfehlungssystems kann die Genauigkeit über die sonst üblichen Maße, wie MAE , nicht definiert werden, da weder der Nutzer noch das Empfehlungssystem eine Bewertung auf einer

ansonsten gängigen Skala bestimmt, sondern die Gegenstände durch eine Gewichtung sortiert werden. Für eine Berechnung von Korrelationsfaktoren, wie Spearmans ρ und Kendalls Tau, müsste der Nutzer im Vorfeld eine Rangfolge seiner räumlichen Entscheidungen festlegen, was in unserem Fall nicht machbar war. Zudem sollte Genauigkeit bei einer Mehr-Objekt-Empfehlung nicht das einzige Evaluationskriterium sein.

Qualität

Um die Qualität einer Empfehlung zu berechnen, werden andere Maße verwendet. Ein Qualitätsmerkmal eines Empfehlungssystems kann zum Beispiel der Prozentsatz an Objekten sein, die überhaupt die in einer Empfehlung enthalten sein können. Die Frage dabei ist, ob für ein bestimmtes Nutzer-Gegenstands-Paar eine Empfehlung ausgesprochen werden kann. Es wird also die mögliche Abdeckung (Coverage – Sarwar u. a. (2000)) der Gegenstandsmenge durch den Empfehlungsansatz bestimmt. Bei nutzerbasierten Ansätzen ist die Abdeckung sehr von der Dichte der Nutzerbeziehungen und deren Bewertungen abhängig. In unserem Fall entsteht eine Bewertung durch die Aufnahme einer bestimmten Ansicht. Die Abdeckung für einen einzelnen Nutzer ist demnach immer die Gesamtsumme von räumlichen Entscheidungen seiner nächsten Nachbarn.

$$coverage_{georec}(u) = \frac{\bigcup_{v \in U_{sim}} cpv(v)}{\bigcup_{v \in U} cpv(v)} \quad (8.6)$$

Da unsere geografische Empfehlung immer für ein konkretes Gebiet ausgewählt wird, kann die Menge an CPVs auch geografisch eingeschränkt werden und dadurch die mögliche Abdeckung in dem jeweiligen Zielgebiet bestimmt werden.

Wird eine Menge von Gegenständen vorgeschlagen, sei es durch eine Top N Empfehlung oder eine Mehr-Objekt-Empfehlung, werden oft Maße aus dem Bereich des Information Retrievals verwendet, um deren Qualität zu bestimmen. Nach der ursprünglichen Definition unterscheidet man in der Menge aller Dokumente zwischen relevanten und nicht relevanten Dokumenten. Die Vereinigung dieser beiden Mengen ergibt demnach wieder die Gesamtmenge aller Dokumente.

$$Recall = \frac{|\{relevanteDokumente\} \cap \{empfohleneDokumente\}|}{|\{relevanteDokumente\}|} \quad (8.7)$$

$$Precision = \frac{|\{relevanteDokumente\} \cap \{empfohleneDokumente\}|}{|\{empfohleneDokumente\}|} \quad (8.8)$$

Da Precision und Recall voneinander abhängig sind, wird oft auch das F1-Maß verwendet, das als gewichtetes Mittel der beiden Maßzahlen gesehen werden kann. Das F1-Maß definiert sich wie folgt:

$$F_1 = 2 \frac{precision * recall}{precision + recall} \quad (8.9)$$

In den meisten Fällen werden bei kollaborativen Filtern Kreuzvalidierungsverfahren eingesetzt. Dazu werden die Bewertungen in zwei Teilmengen aufgeteilt. Mit Hilfe der einen Teilmenge werden die Empfehlungen berechnet, um dann die Ergebnisse mit der zweiten Teilmenge zu überprüfen. Man bezeichnet die erste Teilmenge als Trainingsmenge und die zweite als Testmenge. Für die Berechnung von Precision und Recall wird ein Dokument nur dann als relevant gewertet, wenn es in diesem Datensatz zu finden ist. Dadurch wird ein Dokument, das nicht in dieser Menge liegt, automatisch als nicht-relevant gewertet und somit würde die Menge der relevanten Objekte ständig unterschätzt sein.

Precision und Recall lassen sich für diese Annahmen folgendermaßen umschreiben (Sarwar u. a., 2000):

$$Recall_{rec} = \frac{|\{Testsammlung\} \cap \{Empfehlung\}|}{|\{Testsammlung\}|} \quad (8.10)$$

$$Precision_{rec} = \frac{|\{Testsammlung\} \cap \{Empfehlung\}|}{|\{Empfehlung\}|} \quad (8.11)$$

Ähnlich wie Precision und Recall verhalten sich Maße, die bestimmen, wie gut der Nutzer bei einer Entscheidung durch das Empfehlungssystem unterstützt wird. ROC-Kurven (Herlocker u. a., 2004) sind ein beliebtes Mittel um diese Fähigkeit der Systems auszudrücken, ROC steht für „relative operating characteristic“ oder auch „receiver operating characteristic“.

Da die Qualität von Mehr-Objekt-Empfehlungen ganz neue Anforderungen an die Empfehlungen aufgeworfen hat, wurden neue Qualitätsmaße definiert, die in den meisten Fällen die Variabilität der Ergebnisse beschreibt. Zhang und Hurley (2008) definiert dabei die Maße mit der Bezeichnung „Novelty“ und „Diversity“, Ziegler (2005) verwendet die Begriffe „diversification“ und „Intra-list similarity“, um zu bestimmen, wie überraschend bzw. wie durchmischt eine Ergebnismenge ist.

In unserem Fall werden wir kein direktes Variabilitätsmaß verwenden, da wir die Annahme in Abschnitt 4.2 aufgestellt haben, dass jeder Nutzer eine andere Vorstellung von Variabilität aufzeigt. Variabilität ergibt sich eher aus der präzisen Vorhersage von Gegenständen, die im langen Schwanz der Popularitätsverteilung liegen. Daher wird für die folgende Evaluation ein Maß definiert, dass Precision und Recall nur für Gegenstände nach einem bestimmten Rang beachtet. Dazu wurden aus der berechneten Empfehlung alle Elemente entfernt, die einen niedrigeren Rang als diese obere Schranke haben. In den folgenden Berechnungen

lag diese Schranke bei Rang 10, gemessen auf der Popularitätsverteilung der Heatmap. Wir bezeichnen dieses Maß mit Precision behind 10 beziehungsweise Recall behind 10 oder kurz Precision@10 bzw. Recall@10.

$$Recall@10 = \frac{|\{Testsammlung\} \cap \{Empfehlung\} \setminus \{Top10\}|}{|\{Testsammlung\}|} \quad (8.12)$$

$$Precision@10 = \frac{|\{Testsammlung\} \cap \{Empfehlung\} \setminus \{Top10\}|}{|\{Empfehlung\} \setminus \{Top10\}|} \quad (8.13)$$

Datensätze

Abgesehen davon ein geeignetes Maß zu finden, um ein Empfehlungssystem zu evaluieren, wird für die Offline-Berechnung ein passender Datensatz benötigt. Dieser muss eine entsprechende Breite an Informationen beinhalten, um die üblichen Anlaufschwierigkeiten von Empfehlungstechniken zu überwinden. Einige Datensätze, die für die Evaluierung von kommerziellen Seiten verwendet werden (z.B.: Amazon oder Yahoo), sind leider nicht öffentlich zugänglich.

Einer der am meisten verwendeten Datensätze für die Evaluierung von kollaborativen Filtern ist eine Untermenge der Daten, die mit dem MovieLens Empfehlungssystem gesammelt wurden. Die Daten werden in bereinigter Form von der Forschungsgruppe GroupLens um Paul Resnick zum Download zur Verfügung gestellt. Die Daten werden in drei unterschiedlichen Größenordnungen an Nutzerpräferenzen angeboten (100 Tausend, 1 Million oder 10 Millionen). Jeder Datensatz wurde von GroupLens in kleinere Datensätze für eine Kreuzvalidierung in Trainings- und Testdatensätze aufgeteilt, sodass verschiedene Arbeiten mit den gleichen Ausgangsdatenmengen arbeiten können, um die

Referenz	Evaluationsmaß	Methode	Datensatz
Resnick, et al. 1994	Befragung		Eigene Daten
Hill, et al. 1995	Pearson correlation	CV 90/10	Datensatz 7
Shardanand and Maes 1995	MAE, Standart deviation		Eigene Daten
Karypsis 2001	Precision	CV	Eigen Daten, Datensatz 2
Burke 2002	Accuracy	CV 50/50	EntryC prototype test set (50.000 sessions)
Linden 2003	-	-	Datensatz 9
MacLaughlin and Herlocker 2004	Precision, MAE	CV 75/25	Datensatz 2
Ziegler, et al. 2004	Precision, Recall	CV 80/20	Datensatz 8, 9
Siersdorfer, et al. 2007	Nutzerstudien, IR quality		Eigene Daten
Berkvosky, et al. 2007	MAE (against user sparsity)		Datensatz 5
Mehta, et al. 2007	MAE	CV 90/10	Datensatz 3 + noise and profile attacks
Zhang and Pu 2007	MAE	CV 80/20	Datensatz 2
Zhang, et al. 2008	Precision, Item-novelty, Diversity	CV 50/50	Datensatz 3
Park and Tuzhilin 2008	MAE, RMSE		Datensatz 2, 6
Shani, et al. 2008	ED	CV 90/10	Datensatz 1
Yildirim and Mukkai 2008	Hit-rate (=recall or precision), MAE, RMSE	CV 99/, CV 20/80	Datensatz 3

Tabelle 8.6: Überblick über Evaluationsverfahren, die in der Literatur über Empfehlungstechniken zu finden sind. CV steht für eine Kreuzvalidierung (Cross Validation) der Daten.

Ergebnisse vergleichbarer zu machen. Leider hat sich diese Aufteilung nicht als Benchmark für die verschiedenen Empfehlungssysteme durchgesetzt. In der Literatur zu Empfehlungssystemen wird zwar häufig der MovieLens-Datensatz verwendet, allerdings mit einer Vielzahl von Aufteilungen und verwendeten Evaluationsmaßen. Für den geografischen kollaborativen Filter ist leider keiner der vorgestellten Datensätze verwendbar, da in den Metadaten keine geografischen Informationen enthalten sind. Mögliche Datensätze für die Evaluation finden sich allerdings in den vorgestellten Bildsammlungen aus Kapitel 3.

8.4 Offline-Evaluierung - Panoramio/Flickr

In diesem Abschnitt wird das vorgestellte Empfehlungsverfahren mit Hilfe von bestehenden Datensätzen zu evaluieren. Ziel war die Verwendung einer Evaluationsmethode, mit denen Empfehlungssysteme üblicherweise analysiert werden. Zusätzlich musste ein möglichst großer Datensatz erzeugt werden, der für den geografischen Empfehlungsansatz geeignet war. Für das ausgewählte Kreuzevaluierungsverfahren musste der Datensatz möglichst viele Nutzer beinhalten, die genügend Bildaufnahmen in unterschiedlichen Gegenden zur Verfügung stellen konnten, um die Empfehlungsergebnisse mit einer Teilmenge von diesen Bildern zu vergleichen.

8.4.1 Evaluationsmethode

Die Evaluation der geografischen Empfehlungsmethode wird mit Hilfe einer Kreuzvalidierung durchgeführt (Herlocker u. a., 2004). Dazu wurde der Datensatz in zwei getrennte Mengen unterteilt: in einen Trainingsdatensatz, mit dessen Unterstützung man eine Empfehlung berechnet, und einem Testdatensatz, mit dem man die Güte der Ergebnisse kontrolliert. Im Gegensatz zu anderen Verfahren wurde diese

	Name	Benutzer	Objekte	Instanzen
1	Netflix ¹⁸	480.000	17.000	1.000.000
2	100K movie-lens ¹⁹	943	1.682	100.000
3	1M movie-lens ²⁰	6.040	3.900	1.000.000
4	10M movie-lens ²¹	71567	10681	10.000.000
5	eachmovie ²²	72.916	1628	2.811.983
6	Bookcrossing ²³			
7	video at bell-core.com ²⁴	95291	1750	55.000
8	All consu- ming ²⁵	2.783	9.237	14.591
9	amazon.com ²⁶	2003	29.000.000	Several milli- on
10	Yahoo webs- cope ²⁷	Various	Various	Various
11	Jester Joke Data Set ²⁸	73.496	100	4.100.000

Tabelle 8.7: Datasätze für eine Offline-Evaluierung.

Teilung nicht zufällig oder zeitlich durchgeführt, sondern räumlich. Für die Evaluation der Empfehlungen für Ort O wird dabei folgendermaßen vorgegangen: Zunächst wählt man einen Nutzer aus, der mindestens fünf Bilder in O aufgenommen hat und verwendet diese Bilder als Testmenge für diesen Ort. Diese Bilder bzw. die räumlichen Entscheidungen werden aus der Gesamtmenge entfernt und die restlichen räumlichen Entscheidungen als Trainingmenge verwendet, aus der die Empfehlung für O berechnet wird. Wenn der Nutzer keine Aufnahmen außerhalb des Zielortes aufgenommen hat und daher keine Empfehlung berechnet werden konnte, wird der Nutzer nicht weiter beachtet und es wird mit dem nächsten Nutzer weiterverfahren. Abbildung 8.8 gibt einen Überblick über das Vorgehen.

In Abschnitt 6.3.2 wurden vier verschiedene Maße vorgestellt, um die geografische Ähnlichkeit der verschiedenen Nutzer zu berechnen. Die Hypothese war dabei, dass die Verwendung von zusätzlichen geografischen Randinformationen zu einer Verbesserung der geografischen Empfehlungen führt und damit den nutzerbasierten Empfehlungsansatz besser unterstützen kann. Drei der vier vorgestellten Verfahren verlangen für eine Aussage über die Ähnlichkeit von zwei Nutzern, dass sich diese mindestens eine gemeinsame Entscheidung an einem CSD aufweisen. Das heißt, sie haben mindestens an einem Ort eine gemeinsame räumliche Entscheidung. In allen anderen Fällen ist ihre Ähnlichkeit gleich 0. Weist der Nutzer zu allen anderen Nutzern eine Ähnlichkeit von 0 auf, ist die Ergebnismenge leer. Die Maße sind die Single-Layer-feature-similarity (SFS), die Two-Layer-feature-similarity (TFS) und die Two-Layer-information-similarity (TIS). Im Falle der Geografic-coverage-similarity-Ähnlichkeit (GCS) ist es immer möglich, eine Ähnlichkeit zwischen zwei Nutzern festzustellen, da das Verfahren alle Ebenen der Partonomy in Betracht zieht und die Nutzer immer

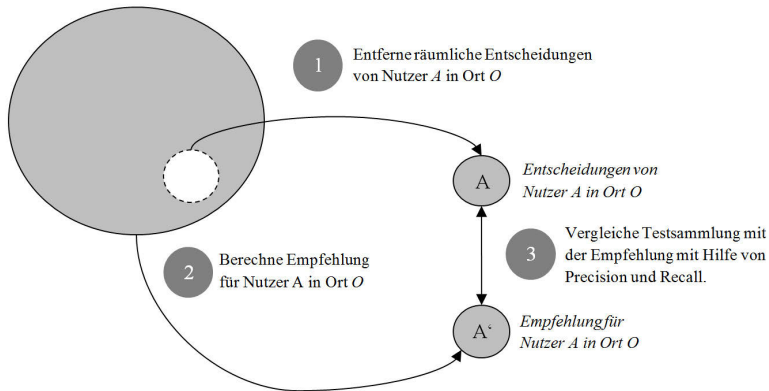


Abbildung 8.8: Kreuzvalidierung mit geografischen Daten für die Güte des geografischen Empfehlungssystems im Detail. (1) Auswahl des Zielgebiets für das die Empfehlung berechnet werden soll. Aus diesem Gebiet werden alle räumlichen Entscheidungen von Nutzer A als Testsammlung aussortiert. (2) Berechnung einer Empfehlung mit Hilfe der restlichen räumlichen Entscheidungen. (3) Vergleich der Empfehlung mit der Testsammlung nach Precision und Recall

eine Gemeinsamkeit im Wurzelknoten, der alle Bilder auf der gesamten Welt beinhaltet, aufzeigen.

Mit der beschriebenen Methode wurden Empfehlungen für 31 Nutzer berechnet, die ihrerseits 500 Bilder in Bamberg aufgenommen hatten. Im Durchschnitt hat jeder dieser Nutzer 15 Bilder aufgenommen, die sich in der Regel auf über acht räumliche Entscheidungen verteilt haben.

8.4.2 Evaluationsdaten

Für die Evaluation wurden Bilder aus Panoramio (siehe Abschnitt 3.1.2) verwendet, da sie eine gute Abdeckung über die gesamte Welt haben und eine sehr aktive Gemeinschaft von Nutzer aufweist, die georeferenzierte Bilder von vielen verschiedenen Orten der Welt zu Verfügung stellen. Um das Verfahren anwenden zu können wurde ein Datensatz benötigt, in dem die Nutzer ihr räumliches Verhalten offenbarten, was sie im Fall von Panoramio durch das Einstellen von georeferenzierten Bildern taten. Als Zielort für die Empfehlungen wurde die Stadt Bamberg gewählt. 261 Nutzer hatten Bilder für Bamberg zu dem Zeitpunkt der Untersuchung über Panoramio online gestellt. Ausgehend von den Nutzern dieser Stadt wurden 18 weitere Städte in 3 verschiedenen Ländern ausgewählt, in denen die Nutzer aus Bamberg verstärkt Bilder aufgenommen hatten. Aus diesen 19 Städten wurden alle Bilder ausgelesen, die über die öffentliche Schnittstelle der Webseite zugänglich waren. Insgesamt wurden dadurch 33.947 Bilder von 5.766 Nutzern in den Evaluationsdatensatz übernommen. Tabelle 8.8 gibt eine Übersicht über alle Gebiete, sowie die jeweilige Anzahl der Nutzer und der Bilder. Für dieselben Gebiete wurden zusätzlich Bilddaten aus Flickr ausgelesen, um das Verfahren auf Bilddaten zu testen, deren Fokus nicht so sehr auf Aufnahmen mit einem starken Ortsbezug lag.

Region	Bilder	Nutzer
World		5.766
Germany		2.763
Baden-Württemberg		1.410
Stuttgart	3.471	808
Freiburg	2.047	505
Tübingen	638	171
Bavaria		1.170
Munich	2977	637
Nürnberg	195	44
Bamberg	4.692	261
Würzburg	1.209	298
Berlin		350
Berlin	3.258	350
Italy		2042
Toscana		1.693
Pisa	793	494
Florence	3.988	876
Lucca	1.569	498
Lazio		508
Rome	2.858	430
Santa Marinella	49	24
Fiumicino	99	45
Aprilia	297	18
France		1461
North France		1.461
Paris	3.784	598
Le Mans	390	89
Caen	691	182
Saint-Malo	1.735	658

Tabelle 8.8: Testdaten aus Panoramio für die Offline-Evaluation

8.4.3 Evaluationsergebnisse

Panoramio

Die unterschiedlichen Ähnlichkeitsmaße aus Kapitel 6 wurden mit Hilfe des Kreuzvalidierungsverfahrens auf Precision und Recall überprüft. Als Baseline wurde die Top10 der Heatmap mit den persönlichen räumlichen Entscheidungen der Nutzer verglichen. Es ist interessant festzustellen, dass bereits 25% der persönlichen Entscheidungen in der Top10 der Heatmap vorhanden waren, was nicht zu sehr verwundert, da diese Orte die beliebtesten im Raum darstellen. Dennoch ist es eine bemerkenswerte Baseline, da der Empfehlungsansatz in der Lage sein müsste, zusätzlich zu diesen „Bestsellern“ Orte zu empfehlen, die abseits der Top 10 zu finden sind. In Abbildung 8.9 ist eine sichtbare Steigerung gegenüber dieser Baseline zu erkennen, da die anfängliche Precision von 0,24 auf 0,41 (blaue Kurve) und der Recallwert von 0,35 auf 0,62 ansteigt (rote Kurve). Gerade die Steigerung von beiden Werten verdeutlicht eine echte Verbesserung der Baseline. Die Nutzerähnlichkeit, die die höchste Steigerungsrate (58% für Precision und 56% für Recall) bot, war die Two-Layer-information-similarity, das auch am meisten von dem berechneten geografischen Zusatzwissen profitiert.

Als zweites Evaluationmaß wurde Precision@10 und Recall@10 verwendet, das im Abschnitt zuvor definiert wurde. Dieses Maß sollte aufzeigen, wie gut das Verfahren in der Lage ist, Gegenstände zu empfehlen, die nicht im Kopf der Verteilung zu finden sind. Auch hier wurden die besten Ergebnisse mit Hilfe des Two-Layer-information-similarity Maßes erreicht. Mit diesem Maß war es möglich ungefähr 50% der Orte, die in der persönlichen Top 10 eines Nutzers vorhanden waren, zu empfehlen und insgesamt 40% der Orte, die in seiner Gesamtauswahl zu finden war. Der Vergleichswert in der Baseline ist 0, da die Ausgangsmenge

an Bildern ausschließlich aus den zehn populärsten Rängen besteht, die wir bei diesem Maß ausgeschlossen hatten.

Dasselbe Evaluationsverfahren wurde zusätzlich auf zwei anderen Regionen, Würzburg (siehe Abbildung 8.10) und Berlin (siehe Abbildung 8.11), angewendet, wohlwissend, dass die Testdaten nicht in für diese Region ausgesucht wurden. Das heißt, viele Nutzer aus diesen Gebieten haben wahrscheinlich noch in anderen Regionen Bilder aufgenommen, die nicht durch den Testdatensatz abgedeckt werden. Für die Region Bamberg hatten wir ja im Vorfeld explizit nach anderen Regionen gesucht, in denen sich gute Nachbarschaften vermuten ließen. Auf diese zusätzliche Auswahl wurde für diese beiden Gebiete verzichtet, um das Verfahren unter weniger optimalen Bedingungen zu testen. Erneut waren die Empfehlungen besser als die Baseline, 60% bis 75% in Precision und circa 70% in Recall.

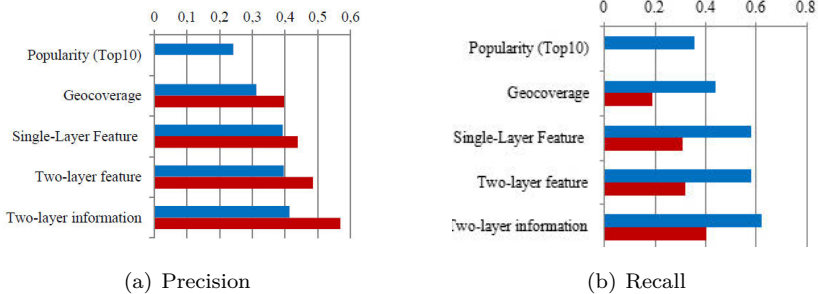


Abbildung 8.9: Precision und Recall für Bamberg. Blaue Balken geben die Precision und den Recall bei Rang 10 an, rote Balken zeigen das Ergebnis von Precision und Recall an, wenn man zuvor alle Elemente, die bei einer typischen Rangordnung unterhalb von Rang 10 lagen, entfernt.

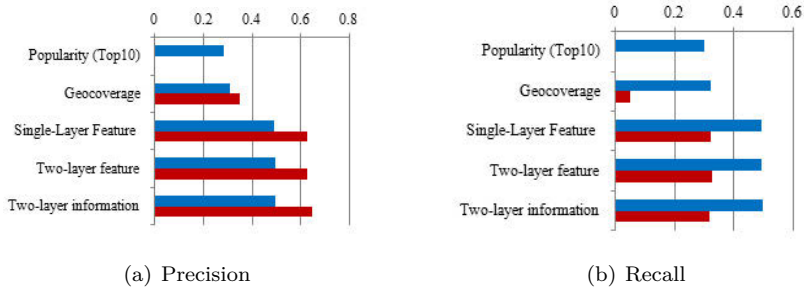


Abbildung 8.10: Precision und Recall für Würzburg.

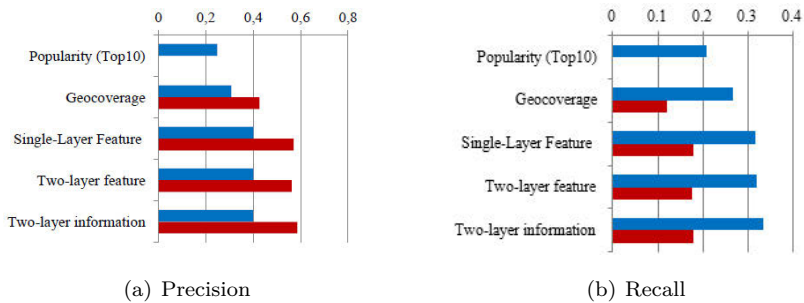


Abbildung 8.11: Precision und Recall für Berlin.

Flickr

Als weitere Kontrolle der Ergebnisse wurde dasselbe Evaluationsverfahren auf eine andere Datenquelle angewendet. In einem zweiten Testverfahren wurden 52.118 Bilder von Flickr für die Analyse verwendet und ebenso wie im ersten Fall wurden die verschiedenen Ähnlichkeitsmaße mit Hilfe von Precision und Recall mit einander verglichen. Tabelle 8.9 gibt eine Übersicht über die Daten, die von Flickr für den Testlauf übernommen werden konnten. Im Vergleich zu dem Datensatz aus Panoramio ist die Anzahl der Benutzer geringer, wobei jeder von

ihnen insgesamt mehr Bilder eingestellt hat. Im Gegensatz zu Panoramio hat Flickr keinen Focus auf Bilder von Orten und geografischen Landmarken, so dass die Auswahl der Bilder, thematisch gesehen, viel offener ist.

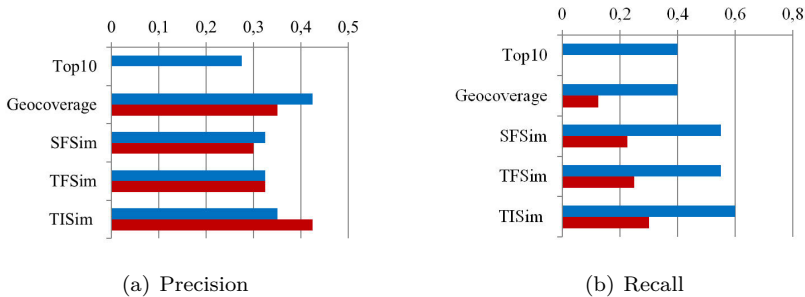


Abbildung 8.12: Precision und Recall für Bamberg – geografische Empfehlung für Bilder aus Flickr

Die Ergebnisse der Evaluierung selbst waren nicht ganz dieselben, wie man sie noch bei der Auswertung für Panoramio beobachten konnte. In dem Fall von Flickr brachte das Geocoverage-Verfahren die besten Ergebnisse, wenn man alleine die Precision der Antwort betrachtet, allerdings zu Kosten des Recalls.

Ein ausgeglichenes Ergebnis lieferte allerdings wieder das TIS Maß. Auch im Vergleich zu den anderen Verfahren liefert dieses Maß die besten Ergebnisse. Eine Steigerung ist gerade Abseits von der Top 10 erkennbar, gemessen an den Werten für Precision@10 und Recall@10. Abschließend zeigen die Ergebnisse, dass auch in Fällen, in denen nicht alle Bilder den gewünschten geografischen Bezug aufzeigen, eine genauere Empfehlung möglich ist. Interessant wäre es das Verhalten des Empfehlungssystems zu testen, wenn die Bildsammlung zuvor mit Hilfe von inhaltsbasierten Verfahren nach diesen Aufnahmen vorgefiltert würde. Bilder, wie Portraits oder andere Personenaufnahmen könnten

Region	Bilder	Nutzer
World		3.362
Germany		1.516
Baden-Württemberg		476
Stuttgart	4.613	247
Freiburg	2.966	211
Tübingen	502	29
Bavaria		751
Munich	4.514	288
Nürnberg	4.707	294
Bamberg	2.544	150
Würzburg	2.133	139
Berlin		350
Berlin	4.296	385
Italy		1.193
Toscana		823
Pisa	3.858	318
Florence	4.296	315
Lucca	3.045	285
Lazio		406
Rome	4.154	338
Santa Marinella	90	11
Fiumicino	307	58
Aprilia	18	9
France		736
North France		736
Paris	4.291	436
Le Mans	933	70
Caen	3.071	85
Saint-Malo	1.521	173

Tabelle 8.9: Testdaten aus Flickr für Offline-Evaluation

mit Hilfe dieser Verfahren aussortiert werden. Während des Projekts Tripod wurde zwar nach Methoden einer Vorfilterung geforscht, allerdings war die Genauigkeit der Ergebnisse zu gering, um sie sinnvoll in dieser Arbeit nutzen zu können. Im Ausblick wird noch einmal verstärkt auf die Möglichkeiten durch inhaltsbasierte Verfahren eingegangen.

Abschließend lässt sich sagen, dass die berechneten Konzeptionalisierungen durch die Heatmap im Vergleich zu professionell erstellten Bildgruppen durchaus vergleichbare Ergebnisse liefern. Außerdem hat sich gezeigt, dass das Empfehlungsverfahren für Datenquellen, die unserem Anwendungsfall entsprechen, deutlich bessere Ergebnisse liefert, als Standardverfahren.

9 Ergebnisse und Ausblick

9.1 Ergebnisse

Ziel dieser Arbeit war es ein Empfehlungssystem zu beschreiben, das auf Basis von räumlichen Entscheidungen neue Entscheidungen vorher-sagen kann. Der Grundgedanke des vorgestellten Ansatzes ist es, die Georeferenzierung einer Information nicht nur als räumliche Metainfor-mation über den Gegenstand zu sehen, sondern auch als Feedback des Nutzers für ein Empfehlungssystem. Diese Art des impliziten Feedbacks wurde bislang weitgehend vernachlässigt, obwohl sie bereits seit 2006 in öffentlichen Bildgalerien, wie Flickr oder Panoramio, gesammelt wer-den. Seither fasste man diese Daten lediglich für eine gemeinschaftliche Sicht auf die Welt zusammen, aus der man zum Beispiel touristische Verhaltensmuster rekonstruiert hat. Eine Mehr-Objekt Empfehlung für einen einzelnen Nutzer, durch eine personalisierte Auswahl von Bildern auf Grundlage seiner eigenen georeferenzierten Bilder, zu berechnen, ist meines Wissens ein neuer Ansatz. Dieser wurde in dem vorgestellten geografischen Bild-Empfehlungssystem umgesetzt und evaluiert.

In dieser Arbeit wurde in diesem Zusammenhang ein geografisches Ähn-lichkeitsmaß entwickelt. Das Ähnlichkeitsmaß selbst ergibt sich aus dem Vergleich von räumlichen Entscheidungen. Eine besondere Rolle spielt dabei die Popularität eines Ortes, die sich über die Anzahl der Nut-zer und der Bilder berechnet. Dies ergibt eine implizite Bewertung der Nutzer für den Raum selbst. Ausgehend von den Überlegungen eines diagnostischen Wertes nach Tversky wurde die berechnete Po-

pularität dazu verwendet, Nutzerentscheidungen unterschiedlich stark für die Ähnlichkeitsberechnung zu werten. Zusätzlich wurden mehrere räumliche Bezugsebenen eingeführt, auf denen sich eine Nutzerähnlichkeit berechnen lässt. Ähnliche, frequenzbasierte Ansätze wären auch mit anderen Informationen möglich, solange sich die Empfehlungsgegenstände gruppieren und mit Hilfe einer Partonomy aufteilen lassen. Konzeptionell würde man dann nicht mehr über eine Übertragung von Nutzerähnlichkeiten von einem Gebiet ins Nächste sprechen, sondern man würde versuchen Nutzergemeinschaften von einem Genres oder einer Kategorien in die Nächste zu überführen. Es müsste allerdings, wie in dem Fall der räumlichen Entscheidungen, getestet werden, ob so eine Übertragung sinnvoll ist. Wie sinnvoll dies im Fall von Bildempfehlungen ist, wurde mit Hilfe von selbsterstellten Postkarten aus zwei unterschiedlichen Städten (Tripost) getestet. Der Empfehlungsansatz selbst wurde mit Bildern aus Flickr und Panoramio evaluiert. Im Vergleich zu den üblichen Empfehlungsansätzen konnte der geometrische Ansatz den Wert von Precision und Recall um ungefähr 60% übertreffen. Interessant wäre es im nächsten Schritt den geografischen Empfehlungsansatz mit Bildquellen zu evaluieren, bei denen keine Vorauswahl durch den Nutzer stattgefunden hat. Es konnte beobachtet werden, dass die Nutzer von beiden Onlinegalerien nicht alle ihre Aufnahmen zur Verfügung gestellt haben. Allerdings sind komplette Bildsammlungen von Touristen nicht online verfügbar und nur über gezielte Befragungen greifbar. In Schlieder und Kremer (2011) zum Beispiel wurden Touristen auf der Straße angesprochen und ihnen GPS Geräte und Fotoapparate zur Verfügung gestellt, unter der Auflage die aufgenommen Bilder für eigene Studien weiterverwenden zu dürfen.

Das vorgestellte Empfehlungsverfahren trennt, durch die Definition der räumlichen Entscheidungen, generell den Inhalt (in unserem Fall das Bild) von dem persönlichen Feedback des Nutzers. Der eigentliche In-

halt der Bilder spielt in dem geografischen Empfehlungsverfahren keine entscheidende Rolle. Dadurch war es möglich, ansonsten heterogene Datenquellen gemeinsam für die Bewertung des Raumes zu nutzen. Das System ist in der Lage, die Webseiten Flickr, Panoramio, Locr, Facebook, sowie Datensätze, die während des Projekts Tripods entstanden sind, anzusprechen, aber auch andere georeferenzierte Bildsammlungen lassen sich leicht integrieren. Möglich macht das nicht nur die verwendeten mediationsbasierte Architektur, sondern auch die strikten Definition einer geografischen Koordinate, die bis auf formale Unterschiede immer dieselbe Aussagekraft hat. Andere Metadaten, zum Beispiel zusätzliche Textinformationen, verhalten sich sehr viel problematischer. Alleine Datenquellen aus mehreren Sprachkulturen sind ein bekanntes Problem und erfordern sehr viel Hintergrundwissen für eine sinnvolle Integration. Ein beliebtes Beispiel stammt aus der Übersetzung von zwei verwandten europäischen Sprachen (Englisch und Französisch) und somit nach der Suche nach einem gemeinsamen Konzept (Mark (1993)). So wird das englische Wort „Lake“ auf Französisch mit „Lac“ übersetzt, was jedoch einen leichte Nuance innerhalb der Bedeutung hat. Wasserflächen, die im englischen Sprachgebrauch als „Lake“ bezeichnet werden, würden im Französischen nicht zwangsläufig als „Lac“ benannt. Die unterschiedliche Konzeptualisierung des Raums mit Hilfe von Text und geografischen Positionsangaben, wurde am Beispiel von Bildern aus der Stadt Bamberg und Texten über die Stadt Bamberg diskutiert.

Abschließend lässt sich sagen, dass die neuen mobilen internetfähigen Endgeräte ein neues Verständnis zum momentanen Standpunkt mit sich bringen. Es ist nicht mehr nur interessant, *wer* eine Information einstellt oder abfragt, sondern auch von *wo*. Für dieses neue Umfeld zeigt die hier vorliegende Arbeit neue Anwendungsfälle auf und bietet auch erste Lösungsansätze an. Schon heute entstehen die nötigen In-

frastrukturen, die für solche Services nötig sind. Zum einen werden die technischen Voraussetzungen (wie Smartphones) immer verbreiteter, zum anderen entstehen aber auch die benötigten Datensätze, da digitale Gemeinschaften (wie ortsbasierte Netzwerke) immer beliebter werden. Nutzer werden heutzutage immer transparenter und das schließt auch Ortsangaben mit ein. Wie man diese zusätzlichen Informationen über die Nutzer sinnvoll nutzt bleibt ein lebendiges Forschungsfeld, nicht nur im Teilgebiet der geografischen Empfehlungssysteme. Offene Fragestellungen und mögliche Anknüpfungspunkte sollen im folgenden Abschnitt als kurzen Ausblick diskutiert werden.

9.2 Ausblick

9.2.1 Ausblick auf weitere Analyse von raum-zeitlichen Entscheidungen

Wie auch andere Ansätze zeigen, können georeferenzierte Informationen sehr unterschiedlich analysiert und ausgewertet werden. Obwohl es sich anbieten würde auch die zeitliche Komponente von Bildern zu betrachten, beschränkt sich die vorliegende Arbeit ausschließlich auf die Ortsinformationen. Projekte, wie Citysense versuchen, die zeitlichen Abläufe mit in eine Empfehlung zu übernehmen, allerdings machen sie die Empfehlung nur abhängig vom augenblicklichen Standpunkt und der momentanen Urzeit. Dadurch empfiehlt das System Orte, die auffällig oft zu einem späteren Zeitpunkt aufgesucht wurden. Diese Empfehlung kommt ohne das Wissen über den Nutzer selbst aus. Zusätzlich könnte man versuchen, wiederum den Nutzer als zusätzlichen Parameter in die Empfehlungsberechnung mit aufzunehmen. Erste Ansätze, touristische Aktivitäten mit Hilfe vom Faktor Zeit weiter zu analysieren, findet man zum Beispiel in (Schlieder und Kremer, 2011).

Unabhängig von den verwendeten nutzerbasierten kollaborativen Verfahren, die für diese Arbeit eine Design-Entscheidung war, um die Personalisierung der Empfehlungen zu garantieren, könnten auch gegenstands-basierte kollaborative Filter mit Hilfe von räumlichen Entscheidungen interessante Einsichten bringen. Diese Verfahren würden berechnen, wie Orte sich durch den Besuch von ähnlichen Nutzern vernetzen. Die Ergebnisse könnten in Zukunft für ein ortsbasiertes Empfehlungssystem genutzt werden.

9.2.2 Ausblick auf Sensorik in Kameras

Neue Sensoren wie GPS für Kameras sind auf der weltgrößten Messe für Fotografie Photokina immer wieder ein Thema. Während die meisten Hersteller neue GPS-Kameras ankündigen, wird auch das Zubehör selbst mit immer mehr Sensoren ausgestattet, die nicht zuletzt auch das Geo-Tagging erleichtern sollen. Ein Beispiel ist die Eye-Fi²⁹ Speicherkarte. Diese ist in der Lage über ein integriertes WIFI-Modul die Position der Kamera, zusätzlich zu den Bildaufnahmen, auf der Speicherkarte abzuspeichern. Tabellen 9.1- 9.3 zeigen eine kurze Übersicht über die Relevanz von bestimmten Trends in Zusammenhang mit dem Stichwort Photokina aus den Jahren 2010 - 2012. Die mögliche Treffermenge wurde über die Expertensuche von Google auf die jeweiligen Jahre eingeschränkt.

GPS Sensoren in Digitalkameras scheinen weiterhin ein Nischendasein zu führen. Allerdings zeichnet es sich ab, dass in die kommenden Generationen von GPS-Kameras weitere Lage-Sensorik eingebaut werden. Digitale Kameras boten im Januar 2011 nur die Möglichkeit zusätzlich Kompassdaten zu erfassen. Die einzige Kamera mit eingebautem Kompass war damals von Sony erhältlich; sie versucht, sich durch die-

²⁹ <http://de.eye.fi/>

Trend	Anzahl Artikel	Anteil in Prozent
Photokina 2010	939.000	
Photokina 2010 DSLR	281.000	29,93
Photokina 2010 3D	246.000	26,20
Photokina 2010 GPS	172.000	18,32
Photokina 2010 Systemkamera	33.400	3,56

Tabelle 9.1: Trends der Photokina 2010 - Googlesuche für das Jahr 2010

Trend	Anzahl Artikel	Anteil in Prozent
Photokina 2011	1.140.000	
Photokina 2011 DSLR	326.000	28,60
Photokina 2011 3D	288.000	25,26
Photokina 2011 GPS	192.000	16,84
Photokina 2011 Systemkamera	40.900	3,59

Tabelle 9.2: Trends der Photokina 2011 - Googlesuche für das Jahr 2011

Trend	Anzahl Artikel	Anteil in Prozent
Photokina 2012	1.430.000	
Photokina 2012 Systemkamera	210.000	14,68
Photokina 2012 3D	124.000	8,67
Photokina 2012 DSLR	121.000	8,46
Photokina 2012 GPS	93.100	6,51

Tabelle 9.3: Trends der Photokina 2012 - Googlesuche für das Jahr 2012

ses zusätzliche Feature von anderen Geotagging-Kameras abzuheben. Im Laufe von 2011 waren es hauptsächlich die robusteren Outdoor-Kameras, die zusätzlich mit weiteren Lagesensoren ausgestattet wurden (Olympus mju tough TG-810, Nikon COOLPIX AW100, Panasonic Lumix DMC-FT3, ...).

Gerade für Gelegenheitsfotografien verwenden die meisten Menschen allerdings eher ein Smartphone. Gerade die neuen Generationen von Smartphones bieten standardmäßig die Lagesensoren, die sich auf dem Markt der Digitalkameras nicht durchzusetzen scheinen. So werden sowohl die GPS-Position, Kompassdaten und auch Neigungswinkel für jede Bildaufnahme mitprotokolliert. Daraus lässt sich der Blick der Kamera komplett rekonstruieren und die Aufnahme lässt sich dadurch mit vorhandenen Datensätzen auch im Dreidimensionalen abgleichen, wie es zum Beispiel in Tripod umgesetzt wurde.

Für einen Empfehlungsansatz besteht der Vorteil dieser zusätzlichen räumlichen Daten darin, dass sich die Bildaufnahmen besser gruppieren lassen. Man entwickelt eine genauere Vorstellung, welches Objekt der Nutzer von seinem Standpunkt aus wirklich fotografiert hat. So könnte man viele der Bildaufnahmen, die in einem Cluster liegen, zusätzlich nach Hauptblickrichtungen trennen, wobei bestimmte topologische Überlegungen beachtet werden müssten (das angenommene Motiv kann zum Beispiel innerhalb oder außerhalb eines Clusters liegen). Alternativ kann man mit Hilfe der Blickrichtung versuchen, das Motiv des Bildes zu rekonstruieren, in dem man vom Standpunkt aus der Blickrichtung solange folgt, bis man auf ein Objekt auf einem zwei- oder dreidimensionalen Raumbelegungsplan trifft (Quinn u. a., 2009). In Zukunft könnte man dadurch nicht nur die Popularität des Benutzerstandpunktes bestimmen, sondern auch die Popularität der fotografierten Objekte. Die Empfehlungen würden dadurch nicht mehr auf dem POVs sondern auf den POIs beruhen.

9.2.3 Ausblick auf Evaluierungsframeworks

Auf die besonderen Herausforderungen bei der Evaluierung von Empfehlungssystemen wurde bereits in Kapitel 8 im Detail eingegangen. Ein Hauptproblem dabei ist die Vergleichbarkeit von unterschiedlichen Empfehlungsalgorithmen. Meistens wird mit speziellen Datensätzen, eigenen Implementationen von Empfehlungsalgorithmen und Evaluationsmethoden gearbeitet. Seit Ende 2011 entwickelt daher einer der führenden Forschungsgruppen (GroupLens) unter dem Namen Lenskit an einem offenen Evaluations-Framework, das eine bessere Vergleichbarkeit ermöglichen soll (Ekstrand u. a., 2011). LensKit ist nicht das einzige Framework dieser Art, das gerade in der Entwicklung ist. Es wird sich in Zukunft zeigen, ob sich eines von ihnen als Standard in der Wissenschaft durchsetzen wird. Weitere Frameworks für Empfehlungssysteme sind:

- SUGGEST³⁰
- COFI³¹
- Apache Mahout³²
- MyMediaLite³³
- EasyRec³⁴
- jCOLIBRI³⁵

³⁰ <http://glaros.dtc.umn.edu/gkhome/suggest/overview/>

³¹ <http://savannah.nongnu.org/projects/cofi/>

³² <http://mahout.apache.org>

³³ <http://www.ismll.uni-hildesheim.de/mymedialite/>

³⁴ <http://www.easyrec.org/>

³⁵ <http://gaia.fdi.ucm.es/projects/jcolibri/>

- myCBR³⁶

Leider umfassen keines der implementierten Anwendungsfälle und die zugrunde liegenden Evaluierungsdatensätze eine Möglichkeit, ein geografisches Empfehlungssystem zu evaluieren. Es ist dennoch zu hoffen, dass sich in Zukunft eines dieser Frameworks auch für die Evaluierung eines geografischen Anwendungsfalls nutzen lässt. Abgesehen von einem fehlenden Framework und einheitlichen Testdatensätzen ist es dennoch interessant zu sehen, dass bereits andere Forscher die erarbeiteten Ergebnisse als Grundlage für ihre eigene Arbeit (Marinho u. a., 2012) verwendet haben und zu vergleichbaren Ergebnissen gekommen sind.

³⁶ <http://mycbr-project.net>

Danksagung

Ich danke allen, die mich über die letzten Jahre begleitet, betreut und unterstützt haben. Die mir zugehört haben und ein offenes Ohr für mich hatten, sei es bei wissenschaftlichen oder privaten Themen.

An erster Stelle danke ich meinem Betreuer, der mir mit seiner langjährigen Erfahrung und seinem vielseitigen Wissen immer wieder den richtigen Weg gezeigt hat. In den letzten Jahren habe ich ihn als Lehrstuhlvater, Projektleiter und Betreuer sehr zu schätzen gelernt. Ohne ihn wäre ich nicht so weit gekommen, vielen Dank. Dem ganzen Team im Lehrstuhl will ich ebenso danken, all den guten Kollegen und Freunden mit denen man so trefflich diskutieren konnte. Ich will auch der guten Seele am Lehrstuhl danken, die mir sehr viel bedeutet. Auch wenn ich schon längst nicht mehr am Lehrstuhl bin, will ich mich für die schöne Zeit dort bedanken.

Insbesondere danke ich allen, die diese Arbeit Korrektur gelesen oder mir Feedback und Verbesserungsvorschläge zu diesem Text gegeben haben. Ich danke auch allen Personen, die ich auf Konferenzen kennen lernen durfte und für die interessanten Gespräche rund um das Thema geografische Empfehlungssysteme. Ich bin auch froh darüber, dass ich diese Arbeit im Rahmen eines EU-Projektes durchführen konnte. Es hat mir die Möglichkeit gegeben, konzentriert und fokussiert an den Ergebnissen dieser Arbeit zu forschen. Außerdem konnte ich dadurch Teile meine Arbeit immer wieder einem Team aus anerkannten europäischen Forschern vorstellen und so hilfreiches Feedback erhalten. Für diese Chance und die nötige finanzielle Unterstützung danke ich der

Universität, dem Lehrstuhl und den Menschen, die damals dafür gesorgt haben, dass der Forschungsantrag damals bewilligt wurde.

Nicht zuletzt danke ich noch meinen Eltern, die immer mit großem Interesse den Fortschritt dieser Arbeit verfolgt haben. Ich will mich auch für den bedingungslosen Rückhalt und Unterstützung bedanken. Es ist gut zu wissen, dass meine Eltern immer für mich da waren und sein werden.

Literaturverzeichnis

- [Ahern u. a. 2007] AHERN, Shane ; NAAMAN, Mor ; NAIR, Rahul ; YANG, Jeannie Hui-I: World explorer: visualizing aggregate data from unstructured text in geo-referenced collections. In: *JCDL '07: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*. New York, NY, USA : ACM, 2007, S. 1–10. – ISBN 978-1-59593-644-8
- [Al-Khalifa und Davis 2007] AL-KHALIFA, Hend S. ; DAVIS, Hugh C.: Exploring The Value Of Folksonomies For Creating Semantic Meta-data. (2007). – URL <http://eprints.ecs.soton.ac.uk/13555/>
- [Ali und van Stam 2004] ALI, Kamal ; STAM, Wijnand van: TiVo: making show recommendations using a distributed collaborative filtering architecture. In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA : ACM, 2004 (KDD '04), S. 394–401. – URL <http://doi.acm.org/10.1145/1014052.1014097>. – ISBN 1-58113-888-1
- [Anand und Mobasher 2003] ANAND, Sarabjot S. ; MOBASHER, Bamshad: Intelligent Techniques for Web Personalization. In: MOBASHER, Bamshad (Hrsg.) ; ANAND, Sarabjot S. (Hrsg.): *ITWP* Bd. 3169, Springer, 2003, S. 1–36. – ISBN 3-540-29846-0
- [Anand u. a. 2007] ANAND, Sarabjot S. ; KEARNEY, Patricia ; SHAPCOTT, Mary: Generating semantically enriched user profiles for Web

- personalization. In: *ACM Trans. Internet Technol.* 7 (2007), Nr. 4, S. 22. – ISSN 1533-5399
- [Anderson 2006] ANDERSON, Chris: *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion, 2006. – ISBN 1401302378
- [Ashbrook und Starner 2002] ASHBROOK, Daniel ; STARNER, Thad: Learning Significant Locations and Predicting User Movement with GPS. In: *ISWC*, 2002, S. 101–108
- [Ashbrook und Starner 2003] ASHBROOK, Daniel ; STARNER, Thad: Using GPS to learn significant locations and predict movement across multiple users. In: *Personal Ubiquitous Comput.* 7 (2003), Nr. 5, S. 275–286. – ISSN 1617-4909
- [Auge 1995] AUGE, Marc.: *Non-places : introduction to an anthropology of supermodernity / Marc Auge ; translated by John Howe*. Verso, London ; New York :, 1995. – 122 p. ; S. – ISBN 1859840515 1859849563
- [Bae-Hee u. a. 2006] BAE-HEE, Lee ; HEUNG-NAM, Kim ; JIN-GUK, Jung ; GEUN-SIK, Jo: Location-Based Service with Context Data for a Restaurant Recommendation. In: *Database and Expert Systems Applications*. Heidelberg, Germany : Springer Berlin, 2006, S. 430–438. – ISBN 978-3-540-37871-6
- [Belongie u. a. 2001] BELONGIE, Serge ; MALIK, Jitendra ; PUZICHA, Jan: Shape Matching and Object Recognition Using Shape Contexts. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2001), S. 509–522
- [Berjani und Strufe 2011] BERJANI, Betim ; STRUFE, Thorsten: A recommendation system for spots in location-based online social networks. In: *Proceedings of the 4th Workshop on Social Network Systems*. New York, NY, USA : ACM, 2011 (SNS '11), S. 4:1–4:6.

– URL <http://doi.acm.org/10.1145/1989656.1989660>. – ISBN 978-1-4503-0728-4

[Billsus und Pazzani 2000] BILLSUS, Daniel ; PAZZANI, Michael J.: User Modeling for Adaptive News Access. In: *User Modeling and User-Adapted Interaction* 10 (2000), Nr. 2-3, S. 147–180. – ISSN 0924-1868

[Breese u. a. 1998] BREESE, John S. ; HECKERMAN, David ; KADIE, Carl: Empirical Analysis of Predictive Algorithms for Collaborative Filtering, URL <http://citeseer.ist.psu.edu/breese98empirical.html>, 1998, S. 43–52

[Burke 2002] BURKE, Robin: Hybrid Recommender Systems: Survey and Experiments. In: *User Modeling and User-Adapted Interaction* 12 (2002), Nr. 4, S. 331–370. – ISSN 0924-1868

[Carbonell und Goldstein 1998] CARBONELL, Jaime ; GOLDSTEIN, Jade: The use of MMR, diversity-based reranking for reordering documents and producing summaries. In: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA : ACM, 1998 (SIGIR '98), S. 335–336. – URL <http://doi.acm.org/10.1145/290941.291025>. – ISBN 1-58113-015-5

[Christiaens 2006] CHRISTIAENS, Stijn: Metadata Mechanisms: From Ontology to Folksonomy ... and Back. In: *Lecture Notes in Computer Science: On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, Springer, 2006. – URL <http://www.springerlink.com/content/m370107220473394>

[Clarke u. a. 2008] CLARKE, Charles L. ; KOLLA, Maheedhar ; CORMACK, Gordon V. ; VECHTOMOVA, Olga ; ASHKAN, Azin ; BÜTTCHER, Stefan ; MACKINNON, Ian: Novelty and diversity in infor-

- mation retrieval evaluation. In: *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA : ACM, 2008, S. 659–666. – ISBN 978-1-60558-164-4
- [Claypool u. a. 1999] CLAYPOOL, Mark ; GOKHALE, Anuja ; MIRANDA, Tim ; MURNIKOV, Pavel ; NETES, Dmitry ; SARTIN, Matthew: Combining Content-Based and Collaborative Filters in an Online Newspaper. In: *In Proceedings of ACM SIGIR Workshop on Recommender Systems*, 1999
- [Clough u. a. 2006] CLOUGH, Paul ; JOHO, Hideo ; PURVES, Ross: Judging the Spatial Relevance of Documents for GIR. In: LALMAS, Mounia (Hrsg.) ; MACFARLANE, Andy (Hrsg.) ; RÜGER, Stefan M. (Hrsg.) ; TOMBROS, Anastasios (Hrsg.) ; TSIKRIKA, Theodora (Hrsg.) ; YAVLINSKY, Alexei (Hrsg.): *ECIR* Bd. 3936, Springer, 2006, S. 548–552. – ISBN 3-540-33347-9
- [Coleman u. a. 2009] COLEMAN, David ; GEORGIADOU, Yola ; LABONTE, Jeff: Volunteered Geographic Information: the nature and motivation of producers. In: *International Journal of Spatial Data Infrastructure Research* (2009)
- [De Pessemier u. a. 2009] DE PESSEMIER, Toon ; DERYCKERE, Tom ; MARTENS, Luc: Context aware recommendations for user-generated content on a social network site. In: *EuroITV '09: Proceedings of the seventh european conference on European interactive television conference*. New York, NY, USA : ACM, 2009, S. 133–136. – ISBN 978-1-60558-340-2
- [Deselaers u. a. 2009] DESELAERS, Thomas ; GASS, Tobias ; DREUW, Philippe ; NEY, Hermann: Jointly optimising relevance and diversity in image retrieval. In: MARCHAND-MAILLET, Stéphane (Hrsg.) ;

KOMPATSIARIS, Yiannis (Hrsg.): *CIVR*, ACM, 2009. – ISBN 978-1-60558-480-5

[Dunn u. a. 2009] DUNN, Greg ; WIERSEMA, Jurgen ; HAM, Jaap ; AROYO, Lora: Evaluating Interface Variants on Personality Acquisition for Recommender Systems. In: *UMAP '09: Proceedings of the 17th International Conference on User Modeling, Adaptation, and Personalization*. Berlin, Heidelberg : Springer-Verlag, 2009, S. 259–270. – ISBN 978-3-642-02246-3

[Dykes u. a. 2008] DYKES, J. ; PURVES, R.S. ; EDWARDES, A. J. ; WOOD, J.: Exploring Volunteered Geographic Information to Describe Place: Visualization of the 'Geograph British Isles' Collection. In: *GIS Research UK (GISRUK 2008)* Manchester Metropolitan University (Veranst.), 2008

[Edwardes und Purves 2007] EDWARDES, Alistair J. ; PURVES, Ross S.: A theoretical grounding for semantic descriptions of place. In: *W2GIS'07: Proceedings of the 7th international conference on Web and wireless geographical information systems*. Berlin, Heidelberg : Springer-Verlag, 2007, S. 106–120. – ISBN 3-540-76923-4, 978-3-540-76923-1

[Ekstrand u. a. 2011] EKSTRAND, M. D. ; LUDWIG, M. ; KONSTAN, J. A. ; RIEDL, J. T.: Rethinking the Recommender Research Ecosystem: Reproducibility, Openness, and LensKit. In: *The Fifth ACM Conference on Recommender Systems*. Chicago, IL : Association of Computing Machinery, 10/23/2011 2011, S. 133–140. – URL <http://doi.acm.org/10.1145/2043932.2043958>

[Espinoza u. a. 2001] ESPINOZA, Fredrik ; PERSSON, Per ; SANDIN, Anna ; NYSTRÖM, Hanna ; CACCIATORE, Elenor ; BYLUND, Markus: GeoNotes: Social and Navigational Aspects of Location-Based

- Information Systems. In: *Proceedings of the 3rd international conference on Ubiquitous Computing*. London, UK, UK : Springer-Verlag, 2001 (UbiComp '01), S. 2–17. – URL <http://portal.acm.org/citation.cfm?id=647987.741330>. – ISBN 3-540-42614-0
- [Fan u. a. 2009] FAN, Jianping ; KEIM, Daniel A. ; GAO, Yuli ; LUO, Hangzai ; LI, Zongmin: JustClick: personalized image recommendation via exploratory search from large-scale Flickr images. In: *IEEE Trans. Cir. and Sys. for Video Technol.* 19 (2009), Nr. 2, S. 273–288. – ISSN 1051-8215
- [Fan u. a. 2010] FAN, Xin ; AKER, Ahmet ; TOMKO, Martin ; SMART, Philip ; SANDERSON, Mark ; GAIZAUSKAS, Robert: Automatic image captioning from the web for GPS photographs. In: *Proceedings of the international conference on Multimedia information retrieval*. New York, NY, USA : ACM, 2010 (MIR '10), S. 445–448. – URL <http://doi.acm.org/10.1145/1743384.1743463>. – ISBN 978-1-60558-815-5
- [Girardin u. a. 2007] GIRARDIN, F. ; FIORE, F. D. ; BLAT, J. ; RATTI, C.: Understanding of Tourist Dynamics from Explicitly Disclosed Location Information. In: *The 4th International Symposium on LBS & TeleCartography*, 2007
- [Girardin u. a. 2008] GIRARDIN, Fabien ; CALABRESE, Francesco ; FIORE, Filippo D. ; RATTI, Carlo ; BLAT, Josep: Digital Footprinting: Uncovering Tourists with User-Generated Content. In: *IEEE Pervasive Computing* 7 (2008), October, S. 36–43. – URL <http://portal.acm.org/citation.cfm?id=1477057.1477252>. – ISSN 1536-1268
- [Girardin u. a. 2009] GIRARDIN, Fabien ; DILLENBOURG, Pierre ; NOVA, Nicolas: Detecting air travel to survey passengers on a worldwide

- scale. In: *J. Location Based Services* 3 (2009), Nr. 3, S. 210–226.
– URL <http://dblp.uni-trier.de/db/journals/jlbs/jlbs3.html#GirardinDN09>
- [Goffman 1964] GOFFMAN, William: On relevance as a measure. In: *Information Storage and Retrieval* 2 (1964), Nr. 3, S. 201–203
- [Goldberg u. a. 1992] GOLDBERG, David ; NICHOLS, David ; OKI, Brian M. ; TERRY, Douglas: Using collaborative filtering to weave an information tapestry. In: *Commun. ACM* 35 (1992), Nr. 12, S. 61–70. – ISSN 0001-0782
- [Golder und Huberman 2006] GOLDER, Scott A. ; HUBERMAN, Bernardo A.: Usage patterns of collaborative tagging systems. In: *J. Inf. Sci.* 32 (2006), Nr. 2, S. 198–208. – ISSN 0165-5515
- [Goodchild 2007] GOODCHILD, Michael: Citizens as sensors: the world of volunteered geography. In: *GeoJournal* 69 (2007), Nr. 4, S. 211–221. – URL <http://dx.doi.org/10.1007/s10708-007-9111-y>
- [Gruber 2005] GRUBER, Tom: *Ontology of Folksonomy: A Mash-up of Apples and Oranges*. 2005. – URL <http://tomgruber.org/writing/mtsr05-ontology-of-folksonomy.htm>
- [Gunawardana und Meek 2009] GUNAWARDANA, Asela ; MEEK, Christopher: A unified approach to building hybrid recommender systems. In: *RecSys '09: Proceedings of the third ACM conference on Recommender systems*. New York, NY, USA : ACM, 2009, S. 117–124. – ISBN 978-1-60558-435-5
- [Hagemann und Vossen 2009] HAGEMANN, Stephan ; VOSSEN, Gottfried: ActiveTags: making tags more useful anywhere on the web. In: *Proceedings of the Twentieth Australasian Conference on Australasian Database - Volume 92*. Darlinghurst, Australia, Australia :

- Australian Computer Society, Inc., 2009 (ADC '09), S. 39–46. – URL <http://dl.acm.org/citation.cfm?id=1862681.1862689>. – ISBN 978-1-920682-73-6
- [Herlocker u. a. 2004] HERLOCKER, Jonathan L. ; KONSTAN, Joseph A. ; TERVEEN, Loren G. ; RIEDL, John T.: Evaluating collaborative filtering recommender systems. In: *ACM Trans. Inf. Syst.* 22 (2004), Nr. 1, S. 5–53. – ISSN 1046-8188
- [Horozov u. a. 2006] HOROZOV, Tzvetan ; NARASIMHAN, Nitya ; VASUDEVAN, Venu: Using Location for Personalized POI Recommendations in Mobile Environments. In: *SAINT '06: Proceedings of the International Symposium on Applications on Internet*. Washington, DC, USA : IEEE Computer Society, 2006, S. 124–129. – ISBN 0-7695-2508-3
- [Hsu u. a. 2002] HSU, Rein-Lien ; ABDEL-MOTTALEB, Mohamed ; JAIN, Anil K.: Face Detection In Color Images. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002), S. 696–706
- [Hu und Pu 2009a] HU, Rong ; PU, Pearl: Acceptance issues of personality-based recommender systems. In: *RecSys '09: Proceedings of the third ACM conference on Recommender systems*. New York, NY, USA : ACM, 2009, S. 221–224. – ISBN 978-1-60558-435-5
- [Hu und Pu 2009b] HU, Rong ; PU, Pearl: A comparative user study on rating vs. personality quiz based preference elicitation methods. In: *IUI '09: Proceedings of the 13th international conference on Intelligent user interfaces*. New York, NY, USA : ACM, 2009, S. 367–372. – ISBN 978-1-60558-168-2
- [Irle 2005] IRLE, A.: *Wahrscheinlichkeitstheorie und Statistik: Grundlagen- Resultate- Anwendungen*. Teubner, 2005 (Teub-

ner Studienbücher). – URL <http://books.google.de/books?id=-mGRLA9w6VYC>. – ISBN 9783519123958

[Karypis 2001] KARYPIS, George: Evaluation of Item-Based Top-N Recommendation Algorithms. In: *CIKM '01: Proceedings of the tenth international conference on Information and knowledge management*. New York, NY, USA : ACM, 2001, S. 247–254. – ISBN 1-58113-436-3

[Kemp 2004] KEMP, Anthony E.: *The musical temperament: Psychology and personality of musicians*. Oxford : Oxford Univ. Press, 2004. – URL <http://www.gbv.de/dms/bowker/toc/9780198523635.pdf>. – ISBN 9780198523628

[Kennedy u. a. 2007] KENNEDY, Lyndon ; NAAMAN, Mor ; AHERN, Shane ; NAIR, Rahul ; RATTENBURY, Tye: How flickr helps us make sense of the world: context and content in community-contributed media collections. In: *Proceedings of the 15th international conference on Multimedia*. New York, NY, USA : ACM, 2007 (MULTIMEDIA '07), S. 631–640. – URL <http://doi.acm.org/10.1145/1291233.1291384>. – ISBN 978-1-59593-702-5

[Konstan u. a. 2006] KONSTAN, Joseph A. ; MCNEE, Sean M. ; ZIEGLER, Cai-Nicolas ; TORRES, Roberto ; KAPOOR, Nishikant ; RIEDL, John T.: Lessons on applying automated recommender systems to information-seeking tasks. In: *proceedings of the 21st national conference on Artificial intelligence - Volume 2*, AAAI Press, 2006, S. 1630–1633. – URL <http://dl.acm.org/citation.cfm?id=1597348.1597458>. – ISBN 978-1-57735-281-5

[Konstas u. a. 2009] KONSTAS, Ioannis ; STATHOPOULOS, Vassilios ; JOSE, Joemon M.: On social networks and collaborative recommendation. In: *SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information re-*

- trieval*. New York, NY, USA : ACM, 2009, S. 195–202. – ISBN 978-1-60558-483-6
- [Kuhn 2007] KUHN, Werner: Volunteered Geographic Information and GIScience. In: *Europe* (2007), S. 13–14. – URL http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Kuhn_paper.pdf
- [Landgrebe 2005] LANDGREBE, Silke: *Städtetourismus*. München : Oldenbourg, 2005 (Lehr- und Handbücher zu Tourismus, Verkehr und Freizeit). – URL <http://www.gbv.de/dms/hbz/toc/ht014184137.pdf>. – ISBN 3486576771
- [Lavrenko u. a. 2003] LAVRENKO, V. ; MANMATHA, R. ; JEON, J.: A model for learning the semantics of pictures. In: *in NIPS*, MIT Press, 2003
- [Linden u. a. 2003] LINDEN, Greg ; SMITH, Brent ; YORK, Jeremy: Amazon.com Recommendations: Item-to-Item Collaborative Filtering. In: *IEEE Internet Computing* 7 (2003), Nr. 1, S. 76–80. – ISSN 1089-7801
- [Lowe 1999] LOWE, David G.: Object Recognition from Local Scale-Invariant Features. In: *Proc. of the International Conference on Computer Vision ICCV, Corfu*, URL <http://citeseer.ist.psu.edu/lowe99object.html>, 1999, S. 1150–1157
- [Maidel u. a. 2008] MAIDEL, Veronica ; SHOVAL, Peretz ; SHAPIRA, Bracha ; TAIEB-MAIMON, Meirav: Evaluation of an ontology-content based filtering method for a personalized newspaper. In: *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*. New York, NY, USA : ACM, 2008, S. 91–98. – ISBN 978-1-60558-093-7
- [Marinho u. a. 2012] MARINHO, Leandro B. ; SOUZA BAPTISTA, Cláudio de ; SANDHOLM, Thomas ; NUNES, Iury ; NÓBREGA, Caio ;

- ARAÚJO, Jordão: Extracting Geospatial Preferences Using Relational Neighbors. In: *CoRR* abs/1204.1528 (2012)
- [Mark 1993] MARK, David: Toward a theoretical framework for geographic entity types. In: FRANK, Andrew (Hrsg.) ; CAMPARI, Irene (Hrsg.): *Spatial Information Theory A Theoretical Basis for GIS* Bd. 716. Springer Berlin / Heidelberg, 1993, S. 270–283
- [Markov 2007] MARKOV, Daniel: *Data Mining the Web: Uncovering Patterns in Web Content, Structure, and Usage*. 1. Wiley-Interscience, 2007. – 218 S. – URL <http://www.amazon.com/Data-Mining-Web-Uncovering-Structure/dp/0471666556>
- [Matyas und Schlieder 2009] MATYAS, Christian ; SCHLIEDER, Christoph: A Spatial User Similarity Measure for Geographic Recommender Systems. In: *Proceedings of the 3rd International Conference on GeoSpatial Semantics*. Berlin, Heidelberg : Springer-Verlag, 2009 (GeoS '09), S. 122–139. – URL http://dx.doi.org/10.1007/978-3-642-10436-7_8. – ISBN 978-3-642-10435-0
- [Matyas 2011] MATYAS, Sebastian: *Gemeinschaftliche Qualitätsgesicherte Erhebung und Semantische Integration von Raumbezogenen Daten*. Bamberg, Universitätsbibliothek der Universität Bamberg, Dissertation, 2011. – URL <http://www.opus-bayern.de/uni-bamberg/volltexte/2011/309>
- [Matyas u. a. 2008] MATYAS, Sebastian ; MATYAS, Christian ; SCHLIEDER, Christoph ; KIEFER, Peter ; MITARAI, Hiroko ; KAMATA, Maiko: Designing location-based mobile games with a purpose: collecting geospatial data with CityExplorer. In: *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*. New York, NY, USA : ACM, 2008 (ACE '08), S. 244–247. – URL <http://doi.acm.org/10.1145/1501750.1501806>. – ISBN 978-1-60558-393-8

- [McGinty und Smyth 2003] MCGINTY, Lorraine ; SMYTH, Barry: On the role of diversity in conversational recommender systems. In: *ICCBR'03: Proceedings of the 5th international conference on Case-based reasoning*. Berlin, Heidelberg : Springer-Verlag, 2003, S. 276–290. – ISBN 3-540-40433-3
- [McLaughlin und Herlocker 2004] MCLAUGHLIN, Matthew R. ; HERLOCKER, Jonathan L.: A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In: *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA : ACM, 2004, S. 329–336. – ISBN 1-58113-881-4
- [McNee u. a. 2006] MCNEE, Sean M. ; RIEDL, John ; KONSTAN, Joseph A.: Being accurate is not enough: how accuracy metrics have hurt recommender systems. In: *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*. New York, NY, USA : ACM, 2006, S. 1097–1101. – ISBN 1-59593-298-4
- [Mertens und Back 2001] MERTENS, Peter ; BACK, Andrea: *Lexikon der Wirtschaftsinformatik*. 4., vollst. neu bearb. und erw. Aufl. Berlin : Springer, 2001. – URL <http://www.gbv.de/dms/bsz/toc/bsz094337675inh.pdf>. – ISBN 3540423397
- [Michael u. a. 2010] MICHAEL, Batty ; ANDREW, Hudson-Smith ; RICHARD, Milton ; ANDREW, Crooks: Map mashups, Web 2.0 and the GIS revolution. 16 (2010), Nr. 1, S. 1–13. – URL <http://www.informaworld.com/openurl?genre=article&doi=10.1080/19475681003700831&magic=crossref>
- [Miyahara und Pazzani 2000] MIYAHARA, Koji ; PAZZANI, Michael J.: Collaborative filtering with the simple Bayesian classifier. In: *PRICAI'00: Proceedings of the 6th Pacific Rim international conference*

rence on Artificial intelligence. Berlin, Heidelberg : Springer-Verlag, 2000, S. 679–689. – ISBN 3-540-67925-1

[Naaman u. a. 2004] NAAMAN, Mor ; SONG, Yee J. ; PAEPCKE, Andreas ; GARCIA-MOLINA, Hector: Automatic organization for digital photographs with geographic coordinates. In: *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*. New York, NY, USA : ACM, 2004, S. 53–62. – ISBN 1-58113-832-6

[Page und Kobsa 2009] PAGE, Xinru ; KOBASA, Alfred: The Circles of Latitude: Adoption and Usage of Location Tracking in Online Social Networking. In: *Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04*. Washington, DC, USA : IEEE Computer Society, 2009, S. 1027–1030. – URL <http://portal.acm.org/citation.cfm?id=1632710.1633650>. – ISBN 978-0-7695-3823-5

[Paramita u. a. 2009] PARAMITA, Monica L. ; TANG, Jiayu ; SANDERSON, Mark: Generic and Spatial Approaches to Image Search Results Diversification. In: *ECIR '09: Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval*. Berlin, Heidelberg : Springer-Verlag, 2009, S. 603–610. – ISBN 978-3-642-00957-0

[Park und Tuzhilin 2008] PARK, Yoon-Joo ; TUZHILIN, Alexander: The long tail of recommender systems and how to leverage it. In: *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*. New York, NY, USA : ACM, 2008, S. 11–18. – ISBN 978-1-60558-093-7

[Parsons 2008] PARSONS, Ed: Next-Generation Digital Earth: A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science. 3 (2008), S. 146–167.

– URL <http://ijmdir.jrc.ec.europa.eu/index.php/ijmdir/article/view/119/99>

[Planet 2010] PLANET, Lonely: *The Travel Book: A Journey Through Every Country in the World*. Lonely Planet, 2010 (Travel Book). – URL <http://books.google.de/books?id=bG7Rrb-HdQAC>. – ISBN 9781741792119

[Purves und Jones 2006] PURVES, Ross ; JONES, Chris: Geographic Information Retrieval (GIR). In: *Computers, Environment and Urban Systems* 30 (2006), Nr. 4, S. 375–377

[Quinn u. a. 2009] QUINN, J. ; SMART, P. ; JONES, C.: 3D city registration and enrichment. In: *CityModels09*, 2009

[Rashid u. a. 2006] RASHID, Al M. ; LING, Kimberly ; TASSONE, Regina D. ; RESNICK, Paul ; KRAUT, Robert ; RIEDL, John: Motivating participation by displaying the value of contribution. In: *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*. New York, NY, USA : ACM, 2006, S. 955–958. – ISBN 1-59593-372-7

[Rashid u. a. 2002] RASHID, Al M. ; ALBERT, Istvan ; COSLEY, Dan ; LAM, Shyong K. ; MCNEE, Sean M. ; KONSTAN, Joseph A. ; RIEDL, John: Getting to know you: learning new user preferences in recommender systems. In: *Proceedings of the 7th international conference on Intelligent user interfaces*. New York, NY, USA : ACM, 2002 (IUI '02), S. 127–134. – URL <http://doi.acm.org/10.1145/502716.502737>. – ISBN 1-58113-459-2

[Rattenbury und Naaman 2009] RATTENBURY, Tye ; NAAMAN, Mor: Methods for extracting place semantics from Flickr tags. In: *ACM Trans. Web* 3 (2009), Nr. 1, S. 1–30. – ISSN 1559-1131

- [Relph 1976] RELPH, E. C.: *Place and placelessness / [by] E. Relph*. Pion, London :, 1976. – [10], 156 p. : S. – ISBN 978-0-85086-055-9
- [Rentfrow u. a. 2003] RENTFROW, Peter J. ; GOSLING, Samuel D. ; RENTFROW, Peter J. ; GOSLING, Samuel D. ; PSYCHOLOGY, Department O.: *Personality processes and individual differences The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences*. 2003
- [Resnick u. a. 1994] RESNICK, Paul ; IACOVOU, Neophytos ; SUCHAK, Mitesh ; BERGSTROM, Peter ; RIEDL, John: GroupLens: an open architecture for collaborative filtering of netnews. In: *CSCW '94: Proceedings of the 1994 ACM conference on Computer supported cooperative work*. New York, NY, USA : ACM, 1994, S. 175–186. – ISBN 0-89791-689-1
- [Resnick und Varian 1997] RESNICK, Paul ; VARIAN, Hal R.: Recommender systems. In: *Commun. ACM* 40 (1997), Nr. 3, S. 56–58. – ISSN 0001-0782
- [Rijsbergen 1979] RIJSBERGEN, C. J. V.: *Information Retrieval*. 2nd. Newton, MA, USA : Butterworth-Heinemann, 1979. – ISBN 0408709294
- [Robertson u. a. 2009] ROBERTSON, Stephen ; VOJNOVIC, Milan ; WEBER, Ingmar: Rethinking the ESP game. In: *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*. New York, NY, USA : ACM, 2009 (CHI EA '09), S. 3937–3942. – URL <http://doi.acm.org/10.1145/1520340.1520597>. – ISBN 978-1-60558-247-4
- [Sarwar u. a. 2001] SARWAR, Badrul ; KARYPIS, George ; KONSTAN, Joseph ; REIDL, John: Item-based collaborative filtering recommendation algorithms. In: *WWW '01: Proceedings of the 10th interna-*

- tional conference on World Wide Web*. New York, NY, USA : ACM, 2001, S. 285–295. – ISBN 1-58113-348-0
- [Sarwar u. a. 2000] SARWAR, Badrul M. ; KARYPIS, George ; KONSTAN, Joseph A. ; RIEDL, John T.: Application of Dimensionality Reduction in Recommender System – A Case Study. In: *IN ACM WEBKDD WORKSHOP*, 2000
- [Schafer u. a. 1999] SCHAFFER, J. B. ; KONSTAN, Joseph ; RIEDL, John: Recommender systems in e-commerce. In: *EC '99: Proceedings of the 1st ACM conference on Electronic commerce*. New York, NY, USA : ACM, 1999, S. 158–166. – ISBN 1-58113-176-3
- [Schenk 2007] SCHENK, Michael ; SIEBECK, Mohr (Hrsg.): *Medienwirkungsforschung*. 3. Tübingen : Mohr Siebeck, 2007. – 847 S
- [Schlieder 2007] SCHLIEDER, Christoph: Modeling Collaborative Semantics with a Geographic Recommender. In: *ER Workshops* Bd. 4802, Springer, 2007, S. 338–347. – ISBN 978-3-540-76291-1
- [Schlieder und Kremer 2011] SCHLIEDER, Christoph ; KREMER, Dominik: Visiting the Same Place but Seeing Different Things: Place Models of Touristic Behavior. In: *KI 2011* (2011)
- [Schlieder und Matyas 2009] SCHLIEDER, Christoph ; MATYAS, Christian: Photographing a City: An Analysis of Place Concepts Based on Spatial Choices. In: *Spatial Cognition & Computation* 9 (2009), Nr. 3, S. 212 – 228
- [Seneviratne und Izquierdo 2010] SENEVIRATNE, Lasantha ; IZQUIERDO, Ebroul: An interactive framework for image annotation through gaming. In: *Proceedings of the international conference on Multimedia information retrieval*. New York, NY, USA : ACM, 2010 (MIR '10), S. 517–526. – URL <http://doi.acm.org/10.1145/1743384.1743473>. – ISBN 978-1-60558-815-5

- [Shardanand und Maes 1995] SHARDANAND, Upendra ; MAES, Patie: Social information filtering: algorithms for automating “word of mouth”. In: *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1995, S. 210–217. – ISBN 0-201-84705-1
- [Shi u. a. 2009] SHI, Yue ; LARSON, Martha ; HANJALIC, Alan: Exploiting user similarity based on rated-item pools for improved user-based collaborative filtering. In: *RecSys '09: Proceedings of the third ACM conference on Recommender systems*. New York, NY, USA : ACM, 2009, S. 125–132. – ISBN 978-1-60558-435-5
- [Snaveley u. a. 2006] SNAVELY, Noah ; SEITZ, Steven M. ; SZELISKI, Richard: Photo tourism: exploring photo collections in 3D. In: *ACM SIGGRAPH 2006 Papers*. New York, NY, USA : ACM, 2006 (SIGGRAPH '06), S. 835–846. – URL <http://doi.acm.org/10.1145/1179352.1141964>. – ISBN 1-59593-364-6
- [Snaveley u. a. 2008] SNAVELY, Noah ; SEITZ, Steven M. ; SZELISKI, Richard: Modeling the World from Internet Photo Collections. In: *Int. J. Comput. Vision* 80 (2008), Nr. 2, S. 189–210. – ISSN 0920-5691
- [Spertus u. a. 2005] SPERTUS, Ellen ; SAHAMI, Mehran ; BUYUKKOKTEN, Orkut: Evaluating similarity measures: a large-scale study in the orkut social network. In: *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. New York, NY, USA : ACM, 2005, S. 678–684. – ISBN 1-59593-135-X
- [Su und Khoshgoftaar 2009] SU, Xiaoyuan ; KHOSHGOFTAAR, Taghi M.: A survey of collaborative filtering techniques. In: *Adv. in Artif. Intell.* 2009 (2009), S. 2–2. – ISSN 1687-7470

- [Tang und Sanderson 2010] TANG, Jiayu ; SANDERSON, Mark: Spatial diversity, do users appreciate it? In: *GIR '10: Proceedings of the 6th Workshop on Geographic Information Retrieval*. New York, NY, USA : ACM, 2010, S. 1–2. – ISBN 978-1-60558-826-1
- [Terveen und Hill 2001] TERVEEN, L.G. ; HILL, W.: *Human-Computer Collaboration in Recommender Systems*, URL <http://www.grouplens.org/papers/pdf/rec-sys-overview.pdf>, 2001. – <p>in "HCI In the New Millenium"by Carroll, J. Addison-Wesley</p>
- [Yang u. a. 2002] YANG, Ming-Hsuan ; KRIEGMAN, David J. ; AHUJA, Narendra: Detecting Faces in Images: A Survey. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002), January, S. 34–58. – URL <http://dx.doi.org/10.1109/34.982883>. – ISSN 0162-8828
- [Ye u. a. 2010] YE, Mao ; YIN, Peifeng ; LEE, Wang-Chien: Location recommendation for location-based social networks. In: *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York, NY, USA : ACM, 2010 (GIS '10), S. 458–461. – URL <http://doi.acm.org/10.1145/1869790.1869861>. – ISBN 978-1-4503-0428-3
- [Zhang und Pu 2007] ZHANG, Jiyong ; PU, Pearl: A recursive prediction algorithm for collaborative filtering recommender systems. In: *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*. New York, NY, USA : ACM, 2007, S. 57–64. – ISBN 978-1-59593-730-8
- [Zhang 2009] ZHANG, Mi: Enhancing diversity in Top-N recommendation. In: *RecSys '09: Proceedings of the third ACM conference on Recommender systems*. New York, NY, USA : ACM, 2009, S. 397–400. – ISBN 978-1-60558-435-5

- [Zhang und Hurley 2008] ZHANG, Mi ; HURLEY, Neil: Avoiding monotony: improving the diversity of recommendation lists. In: *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*. New York, NY, USA : ACM, 2008, S. 123–130. – ISBN 978-1-60558-093-7
- [Zhang und Hurley 2009] ZHANG, Mi ; HURLEY, Neil: Novel Item Recommendation by User Profile Partitioning. In: *WI-IAT '09: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*. Washington, DC, USA : IEEE Computer Society, 2009, S. 508–515. – ISBN 978-0-7695-3801-3
- [Zhou u. a. 2007] ZHOU, Changqing ; FRANKOWSKI, Dan ; LUDFORD, Pamela ; SHEKHAR, Shashi ; TERVEEN, Loren: Discovering personally meaningful places: An interactive clustering approach. In: *ACM Trans. Inf. Syst.* 25 (2007), Nr. 3, S. 12. – ISSN 1046-8188
- [Ziegler u. a. 2008] ZIEGLER, Cai-Nicolas ; LAUSEN, Georg ; KONS-TAN, Joseph A.: On exploiting classification taxonomies in recommender systems. In: *AI Commun.* 21 (2008), April, S. 97–125. – URL <http://portal.acm.org/citation.cfm?id=1460172.1460177>. – ISSN 0921-7126
- [Ziegler u. a. 2004] ZIEGLER, Cai-Nicolas ; LAUSEN, Georg ; SCHMIDT-THIEME, Lars: Taxonomy-driven computation of product recommendations. In: *CIKM '04: Proceedings of the thirteenth ACM international conference on Information and knowledge management*. New York, NY, USA : ACM, 2004, S. 406–415. – ISBN 1-58113-874-1



University
of Bamberg
Press

Mobile Geräte werden immer häufiger mit Sensoren zur Bestimmung der eigenen Position ausgestattet, zum Beispiel mit GPS. Mit Hilfe der Ortsinformationen dieser Sensoren können beispielsweise moderne Bildmanagementanwendungen digitale Fotos automatisch nach geografischen Regionen gruppieren oder passende Schlagworte generieren. Dies führt unter anderem zu einer besseren Suchbarkeit dieser digitalen Daten. Grundsätzlich geben Ortsinformationen in digitalen Fotos nicht nur Hinweise auf das Foto selbst, sondern machen auch sichtbar, welche geografischen Entscheidungen der Fotograf bei deren Erstellung getroffen hat. Diese Arbeit nutzt diese Entscheidungen für die Berechnung von weiteren Empfehlungen für den Nutzer, beispielsweise einer Bildmanagementanwendung. Ein konkreter Anwendungsfall lautet folgendermaßen: Einem Nutzer sollen für eine frei wählbare geographische Region (z.B. einer Stadt), mehrere Bilder empfohlen werden, die zum einen typisch für diese Region sind, zum anderen aber auch für ihn persönlich interessant sein könnte.



eISBN 978-3-86309-205-4



9 783863 092054

www.uni-bamberg.de/ubp