Faculty of Science - University of Zurich
Department of Geography
GIScience Center: Geographic Information Visualization and Analysis

Master thesis - Geo511

Development of a Mobile Application for a User-generated Collection of Landmarks

30 September 2014

Author
Marius Wolfensberger
marius.wolfensberger@uzh.ch
08-915-795

Supervisor
Dr. Kai-Florian Richter

Faculty Member
Prof. Dr. Sara I. Fabrikant

Abstract

The present thesis introduces a method for a user-generated collection of landmarks in form of a mobile application for smartphones. This program provides a means for the manual gathering and sharing of landmark information. The developed application demands the user to take a photo of the desired landmark with his mobile phone in order to label it. In the moment of taking the picture, the software computes a field of view representing the visible area of the user. The required positioning details are provided by the integrated GPS sensor and the compass. By querying OpenStreetMap data of the respective area, the program derives the required geographical information. In order to calculate the probability of the involved candidates of being the desired landmark, a ranking system is presented, estimating the visual- and semantic suitability of the examined objects. This system considers, among other factors, nearness to the user, background information and visibility of the object. The suggested candidates have to be manually confirmed. Subsequently, the collected information is uploaded to the OpenStreetMap servers.

During the course of this thesis, the performance of the application was evaluated by capturing landmarks under varying environmental conditions and different ranking settings. Furthermore, a user study was conducted investigating the neutrality of the obtained results. Two different regions were examined in the frame of this evaluation: One region represented a typical urban environment (Zurich) whereas the other showed a predominantly rural character (Zumikon). The results exhibited significant improvements in detecting the intended landmark by combining the visual- and semantic landmark characteristics compared to only including one of these two factors. On the condition of an existing OpenStreetMap representation, the application was able to detect the investigated landmarks in 62.5 % (Zurich) and 90.0 % (Zumikon) respectively as the most probable candidate while being in a distance of 5 - 15 meters to the object. In a distance of 15 - 35 meters, this rate decreased to 47.5 % / 80.0 %. In all cases from 5 to 15 meters and in 85.0 % (Zurich) / 100.0 % (Zumikon) of all tests from 15 to 35 metres, the application managed to include the desired landmark in the four most plausible objects.

A key feature of this approach is that the data requirements can be kept at a very low level. In view of the established performance, this thesis has shown that a manual collection of landmarks using the presented approach is a viable solution in at least more densely mapped areas. Hence, this study paves the way for introducing a widespread user-generated landmark collection.

Acknowledgments

I would like to thank my supervisor Dr. Kai-Florian Richter for his invaluable guidance and his time which made this Master thesis possible in the first place.

I'd like to thank numerous other people who assisted me during this process:

The persons from the OpenLandmarks project for their code to authorize and perform OpenStreetMap uploads from third-party applications. A part of this code was used in this the developed application.

Max Wolfensberger, Bernadette Arbogast, Kevin Winter, Ladina Glaus and Simon Etter for the proofreading of the thesis.

Clio Cocina for the participation in the user study.

Larissa Hauser for providing me with a table of the most widespread category-tags from OSM in Switzerland.

Prof. Dr. Sara I. Fabrikant, Dr. Arzu Çöltekin and Benjamin Rohrbach for the valuable input during and after the concept presentation

Contents

List of Figures

List of Tables

List of Code-Fragments

List of Abbreviations

- API: Application Programming Interface
- CBIR: Content-Based Image Retrieval
- FOV: Field of View
- GPS: Global Positioning System
- OAuth: Open Standard for Authorization
- OSM: OpenStreetMap
- TTFF: Time to First Fix
- UMTS: Universal Mobile Telecommunications System
- VGI: Volunteered Geographic Information
- WGS84: World Geodetic System 1984
- WPS: Wi-Fi-based Positioning System
- XML: Extensible Markup Language

1 Introduction

This section provides an overview about the motivation of this thesis and explains the significance of this work. In order to specify the scope of study, the research questions ought to be answered within this thesis are listed.

1.1 Problem Definition

Imagine yourself travelling in a foreign city on foot while using a navigational system to be guided around. Instead of retrieving familiar routing instructions, such as ”Turn right after 100 meters” the system refers to salient objects on your way imitating path descriptions from local experts. For that matter, these guidances do not only consist of simple turn-by-turn directions; they also refer to objects along the route [Michon and Denis, 2001], providing an indication whether you are following the correct path.

These referenced objects are so-called ”landmarks”. By including such landmarks in route instructions, the previous example may turn into: ”Immediately after the traffic signal turn right, where you see a small church on the left side”. This provides several benefits: These landmarks support the traveller at decision points where a reorientation is needed or provide verification of being on the right path [Sadeghian and Kantardzic, 2008].

Although there are several approaches of including these landmarks in routing instructions (see for example [Duckham et al., 2010; Raubal and Winter, 2002]) they have not been used in widespread commercial navigational systems until today. One of the remaining challenges is the identification of suitable landmark objects. Previous methods for the extraction of landmarks out of existing data were hampered by vast data prerequisites, an uneven distribution of landmark candidates or are limited to acquire only prominent representatives [Richter and Winter, 2011]. Thus, Richter [2013] suggests to introduce a user-generated method based on the concept of ”Volunteered Geographic Information” (VGI). By integrating users in the process of landmark collecting, one might be able to overcome these shortcomings.

1.2 Motivation and Scope of the Study

With the advent of smartphones and the associated broad dissemination of positioning techniques such as GPS, location-based services achieved an unprecedented level of deployment. The most frequently adopted usage are wayfinding services, providing routing instructions to assist users in their orientation. Current navigational services construct their guidances exclusively based on road geometry, orientation and street names [Raubal and Winter, 2002]. The lack of a widely applicable method to extract landmarks from existing data and the absence of a common practice for storing landmark-related information has prevented a broad application in navigational services to this day [Richter and Winter, 2011].

The primary goal of this master thesis is to **develop a user-generated method to collect landmarks**. Involving users in the process of landmark identification has the advantage of creating a dataset based on human cognition. The planned program aims to allow the labelling of objects as landmarks while inspecting them. This will be achieved through the creation of an application for current smartphones. The software will be developed for Android phones due to its high market share [Lomas, 2014] and the openness of its system.

On mobile screens, a precise marking of a location using the finger as input is rather difficult. Instead use is made of the integrated camera in order to prevent an inaccurate placing of points. The planned application demands the user to take a picture of the desired landmark in order to label them. The integrated smartphone sensors, in this case the GPS and the compass, provide the information to determine the user's field of view. By querying OpenStreetMap (OSM) data, possible landmark candidates in this area are obtained. The received candidates are then to be ranked by their probability of being the intended object. Subsequently, the ranked objects are presented to the user to confirm or decline the suggested candidates. At the same time, the requirement to take a picture of the landmark enables a photo-based collection of landmarks.

During this work, it will be concluded whether a smartphone application is a suitable approach for user-generated collection of landmarks. This leads to the following four research questions:

1. *Is a camera-based interface an appropriate way to collect landmarks with the help of smartphone sensors?*
2. *Are smartphone sensors accurate enough to identify the desired landmarks?*
3. *Is the data density of OpenStreetMap high enough to provide sufficient landmark candidates for a user-generated landmark collection?*
4. *Does OpenStreetMap provide a valid method to store and access landmark information?*

In order to answer these questions, following points need to be examined:

- How accurate is the recognition of the suggested landmarks and how can it be improved?
- Which factors should be included to determine the probability that a certain object in the field of view is the intended landmark?
- Are the GPS- and the compass sensors of smartphones accurate enough to identify the intended landmarks?
- What is the influence of the OpenStreetMap data density?

- How should landmark-information be stored on OpenStreetMap?

The planned process is the following:

1. Testing of mobile sensors and determining the degree of their inaccuracy
2. Planning and programming of the application
3. Developing and implementing the ranking of landmark candidates
4. Determining of suitable weighting factors
5. Performing a field test to review the mobile application
6. Conducting a user study to investigate the neutrality of the results
7. Evaluating the data and comparing the success rate of detecting the intended landmarks with different ranking options

1.3 Thesis Organization

This thesis has the following structure:

- Section 2 is dedicated to the theoretical background of landmarks and devotes attention to the idea of ”Volunteered Geographic Information” and OpenStreetMap as a particular example of this concept.
- Section 3 focuses on the planning and developing of the application. This includes an estimation of sensor accuracies, strategies to reduce the influence of possible inaccuracies thereof as well as a proper definition of the process of converting OpenStreetMap objects to landmarks.
- Section 4 explains the evaluation of the application and presents the resulting outcomes.
- In Section 5 the acquired cognitions during the process of this thesis are discussed. Additionally, benefits and shortcomings of the chosen approach are to be debated.
- Section 6 illustrates the findings and conclusions of this thesis and gives a possible outlook about how future research should address this topic.

2 Background

The present thesis is based on a series of theoretical concepts concerning the definition and possible use of landmarks. In order to understand the background and the benefits of the planned application, the following aspects are dedicated to the purpose of providing a clear definition about landmarks and ”Volunteered Geographic Information” and presents existing related work. The concept of VGI is of high importance in this work. The envisaged program will be based on the geographical data of OpenStreetMap and will further store the received results on the OSM servers. Therefore, a summary of the map service OpenStreetMap is given in this section.

2.1 Landmarks

Landmarks are distinguishable environmental features that are unique or in contrast with their neighbourhood [Siegel and White, 1975]. Lynch [1960] defined geographical landmarks as any element, which may serve as reference points. This implies that landmarks can be associated to navigational actions to indicate the position and time where these actions should be taken [Vinson, 1999]. Therefore, a landmark can be any element, which can be used to define the location of other objects or locations [Sorrows and Hirtle, 1999].

Landmarks are crucial in mental representations of space [Hirtle and Jonides, 1985] and are widely used in human wayfinding and communication about space [Duckham et al., 2010]. Commercial navigational systems can implement landmarks to mimic human principles of direction giving [Dale et al., 2005]. May et al. [2003] showed, that the inclusion of such reference points in navigational instructions increased the satisfaction as well as the efficiency of pedestrians in orientation tasks. Hence, landmarks are not only of theoretical interest but also of great practical relevance [Raubal and Winter, 2002].

Sorrows and Hirtle [1999] specified typical features of ”good” landmarks: *Singularity, Prominence, Meaning* and *Prototypicality.* The feature *Singularity* symbolizes a sharp visual contrast with the landmarks surroundings. (Visual) *Prominence* describes the significance of the spatial location, such as a high visibility from many different positions or a prominent position for instance at a junction of roads. *Meaning* refers to the landmark’s content i.e., its cultural or historical importance. If a landmark is a typical representative of a category, it may be salient due to their *prototypicality* [Richter, 2007].

Based on these features, the authors Sorrows and Hirtle derived three main categories of landmarks:

- *Visual Landmark:* Visual landmarks are considered as landmarks due to their visual prominence. This may involve a sharp contrast with the object’s surroundings or ”eye-catching” visual properties (e.g. shape, texture).

- *Cognitive / Semantic Landmark:* Cognitive landmarks stick out through their historical or semantic importance. These types of landmarks stand out for their typical or atypical characteristics in their environment. They tend to be more influenced by a personal background and can easily be missed by those not familiar with the surroundings.

- *Structural Landmark:* Structural landmarks are characterized by the importance of their location or their role in space (e.g. traffic junctions).

It is crucial to mention, that these are not discrete classifications. Important landmarks can belong to multiple categories at the same time [Sorrows and Hirtle, 1999]. Figure 2.1 summarizes these categories and lists attributes describing the respective characteristic.

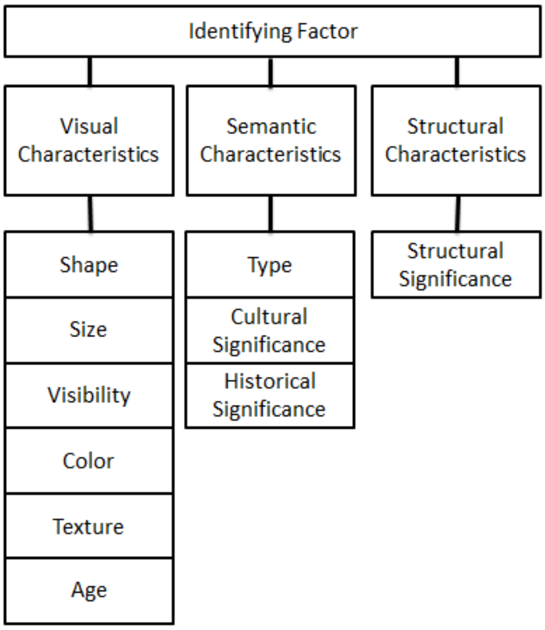


Figure 2.1: The identifying factors defining the ”landmarkness” of a geographic object [Original by Richter and Winter, 2011]

In addition, the distinction between *local* and *global* landmarks is required for the subject matter of this thesis. Global landmarks function as overall orientation points and are visible from a large-scale point of view. If used for navigational purposes, they can act as reference points to provide directional information [Sadeghian and Kantardzic, 2008]. Local landmarks are typically close to a route, and therefore suited to assist navigational tasks at decision points or along route segments [Winter et al., 2008]. This thesis will focus primarily on local landmarks.

2.1.1 Including Landmarks in Routing Instructions

Local landmarks can be used to complement routing instructions both for pedestrians [Tomko, 2004] and for car navigation [Burnett, 2000; Duckham et al., 2010]. Automatically generated instructions, for example "After 100 meters turn right", are typically only based on orientation and metrics (time or distance) and partially street names [Raubal and Winter, 2002]. However, the inclusion of metrics is not an effective way of indicating an upcoming decision point, as the estimation of distances without any further tools constitutes a complex task [Beeharee and Steed, 2006] and can easily be twisted due to outside influences (traffic lights, crowded pathways) [Tomko, 2004]. To overcome these deficiencies, landmarks can be included in routing instructions. Particularly at decision points, where a reorientation is needed, they increase the performance and efficiency of users [May et al., 2003]. Moreover, landmarks can also be included along straight segments in order to confirm being on the right path [Sadeghian and Kantardzic, 2008].

According to Raubal and Winter [2002], a typical routing instruction including landmarks can be constructed as follows:

> *"AT previous landmark TURN LEFT ONTO Stephansplatz UNTIL Haas building, a dark building of architectural significance containing a (signed) Zara shop at the right"*

This instruction includes the "'Haas"-building as its destination landmark. An internet search describes this object as an "eye-catching example of modern architecture with a curved glass facade" [Tripomatic, 2014]. Therefore, this object exhibits salient visual characteristics making it a suitable landmark.

Based on such examples, Burnett et al. [2001] developed a list of relevant characteristics of landmarks qualifying objects for the use in navigational purposes:

- **Permanence:** The likelihood of a landmark of being consistent over a longer period of time
- **Visibility:** Whether the landmark is visible from different locations and under different conditions (night, rain).
- **Usefulness of Location:** Whether the landmark is close to navigational decision points
- **Uniqueness:** The likelihood of the landmark of not being confused with other objects in its environment due to an individual appearance or due to distance to objects with the same type
- **Brevity:** The simplicity of referring to the landmark i.e., the number of words used to describe the landmark

For a widespread use in navigational systems, the challenge remains to find landmarks fulfilling these requirements on the one hand and on the other hand excluding those geographic objects which do not fit into routing instructions [Richter and Winter, 2011].

2.1.2 Collecting Landmarks

Currently, there are very few commercial navigation systems which include landmarks in their route instructions. The primary reason for this fact is the lack of available landmark data [Duckham et al., 2010]. Until today there are neither widespread possibilities to access and store them [Tomko, 2004] nor standardized characteristics defining landmarks [Duckham et al., 2010]. For a wider application range of landmarks in such systems, there is a need for widely applicable methods to extract them from existing data and an appropriate possibility to store and access this information [Richter and Winter, 2011].

Until now, several automated methods have been developed for this purpose. The first method was contributed by Raubal and Winter [2002]. Their approach was, to translate the three main characteristics of landmarks (see Section 2.1) into specific attributes which can be extracted from available data sources. These sources consisted of digital city maps for structural information, geo-referenced images for visual properties and databases such as yellow pages for cognitive data. The extracted attributes were compared to properties of surrounding objects to decide whether something can be seen as a potential landmark. However, the vast amount of data hindered this method from a broad use [Richter, 2013].

Other approaches used data mining approaches for the identification of landmark. Tomko [2004] for example filtered search engine results to look for potential landmark-candidates along a pre-calculated route. Yet, this approach has struggled with the issue that search and filtering results were still obtained manually. The generated routing instructions were therefore labour-intensive [Tomko, 2004; Richter, 2013]. Other methods searched through popular photosharing services for instance Flickr[1] to detect clusters where landmarks are suspected [Papadopoulos et al., 2010]. This approach, however, provided particularly famous landmarks for example tourist attractions. Hence, any potential benefit of them in routing instructions cannot be guaranteed.

Since landmarks should be unique in their respective environment (see Section 2.1.1), "landmarkness" is a relative characteristic [Nothegger et al., 2004]. This implies that nearby features also need to be included in the process of the identification [Richter and Winter, 2011]. Duckham et al. [2010] compared the category information of so-called points of interest (POI) with their surroundings in order to obtain their "landmarkness". These POIs were ranked by the overall suitability of their category and the uniqueness in their area. This provides the advantage, that category data are significantly more

---

[1] `www.flickr.com`

widespread than detailed information about shape, color etc. In this manner, the amount of required data was greatly reduced. Nevertheless, this method suffered from an unequal distribution of information to extract appropriate landmark candidates [Richter, 2013]. Another problem raised by the exclusive focusing on category data becomes apparent in the example shown in Figure 2.2. Both objects are described as ”drinking water” in OpenStreetMap (for more information on OpenStreetMap see Section 2.3.1). While the first fountain (a) is situated on a decision point and provides salient visual properties through its size and ornaments, the other one (b) is small and of no structural relevance. Thus, objects can have the same category information and can either be very suitable landmark candidates or not at all. To automatically detect such discrepancies, the inclusion of more data than only category information is required.

(a) Fountain near Grossmünster (b) Fountain at Hirschenplatz

Figure 2.2: Comparison of two fountains in Zurich with the same category information in OpenStreetMap (”Drinking water”)

To face the aforementioned difficulties arising from automated landmark identification, Richter [2013] suggests, applying the concept of ”Volunteered Geographic Information” to induce the manual collection of landmarks. VGI is a form of user-generated content, specifically targeted at the acquisition of geographic information (For further information on VGI see Section 2.3) [Goodchild, 2007]. The goal is to provide a method which allows a straightforward way to collect and share landmark information. Ideally, this approach should be implemented on current smartphones to become of widespread use [Richter and Winter, 2011]. This would allow using the built-in GPS sensor and the compass to determine the currently visible objects. By involving the user in the selection procedure, this method may overcome the lack of sufficient data as a large part of the data filtering can be performed manually. Furthermore, such a system might lead to a more evenly distribution of landmark information, since it would allow the selective acquisition of landmarks where they are required.

2.2 Related Projects

Several projects already used the combination of location- and compass-data from smartphone sensors to provide and collect geographical information. Three of them are discussed in the following section, to offer an overview about the variety of possibilities through the usage of these sensors.

The mobile application *IPointer*[1] allows to retrieve relevant information about objects in the user's environment by pointing at them. Through the usage of the location data and the integrated compass, the application tries to find the correct object and looks for related content. The presented facts mainly consist of addresses, phone numbers and promotional information.

The *MapIT* system [Frommberger et al., 2013] allows the mapping of small areas through the use of a smartphone. The case study presented a development project involving the acquisition of geographical information about fish ponds to ensure a proper protein supply for villagers. By taking a photo of the intended area with a smartphone, users are able to draw the contours of the pond on the screen. The resulting data is subsequently uploaded to a GeoServer. This system enables mapping on a level which otherwise is not accessible for example by aerial photographs. Therefore, *MapIT* provides a good example of how geographical data can be collected even by people with zero-experience in such tasks.

The *TRIPOD* project [Jones et al., 2009] offers a method which automatically generates captions and metadata for photos based on their location and, if existing, azimuth information. Underlying data are, among others, the Corine landcover dataset and OpenStreetMap data [Purves et al., 2010]. The location is used to determine the landscape type or nearby features which are potentially present in the image. The azimuth for its part helps to narrow down the number of visible objects [Jones et al., 2009]. The automatically generated captions are thereby mimicking natural language image descriptions [Edwardes and Purves, 2007].

2.3 Volunteered Geographic Information and OpenStreetMap

Until recent developments, the collection of geographical data and mapping was conducted by experts. In the last years however, large numbers of private citizens par-

---

[1] `www.ipointer.com`

ticipated in mapping tasks throughout the internet. These people have little to no expertise in this subject and their actions are based on a voluntarily basis [Goodchild, 2007]. Goodchild [2007] calls this phenomenon "Volunteered Geographic Information". VGI is defined as the "voluntary act of creating geographic information" [Ather, 2009] and is based on the more general concept of Crowdsourcing. Crowdsourcing describes the solving of tasks by a great number of voluntary users without the necessity of any special knowledge [Surowiecki, 2005].

The proliferation of Web 2.0 has led to several web-services implementing VGI [Ather, 2009]. With the help of systems such as GPS trackers people are able to collect geographical data with adequate accuracy and to provide this information on the web [Goodchild, 2007]. An example is FixMyStreet[1], where users can report damages or problems related to streets [Richter and Winter, 2011]. A further success story of VGI is OpenStreetMap[2] [Haklay and Weber, 2008]. OSM offers freely available geographical data all over the world collected by voluntary participants.

2.3.1 OpenStreetMap

OpenStreetMap was founded in 2004 by Steve Coast at the University College London. It allows users to create, edit and access free to use map data [OSM, 2014a]. Any volunteer can add or change existing data to improve the service [OSM, 2014a]. As of August 2014, the user base of OSM has surpassed 1,700,000 registered users [OSM, 2014d].

The key intention of this project is to enable unrestricted admission to geographical information all over the world [Haklay, 2010]. The data is mainly collected by volunteers and can be edited online directly on the OSM web page (see Figure 2.3) or via third party software. The standard workflow of collecting data is to trace the outlines of buildings etc. from the available aerial imagery or by using GPS devices to collect data in the field [Mooney et al., 2010]. In addition to user-generated data, various organizations and governments have donated their data to OpenStreetMap [Haklay and Weber, 2008].

2.3.2 Data Quality in OpenStreetMap

There have been continuing discussions about the data quality of OSM [Ather, 2009]. Systems relying on user-generated content can suffer from lack of accuracy of the generated data, loss of control or vandalism [Flanagin and Metzger, 2008]. Hence, there is no guarantee that the provided information complies with certain quality standards [Richter, 2013]. The greatest strength of user-generated content, the concept that anyone can participate, may also be a burden.

---

[1] `www.fixmystreet.com`

[2] `www.openstreetmap.org`

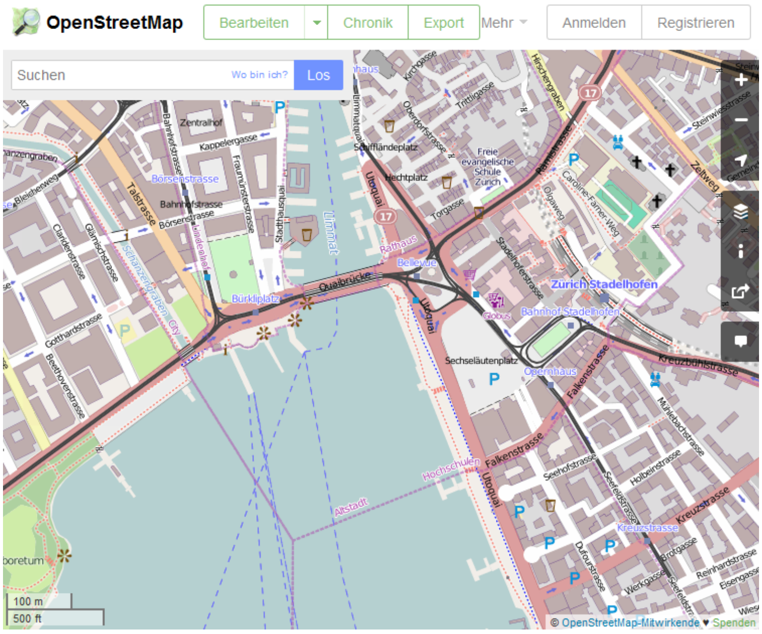


Figure 2.3: The main page of OpenStreetMap showing the inner City of Zurich [OSM, 2014]

Research has been carried out to systematically review the positional accuracy and the data quality of OSM data (e.g. [Ather, 2009; Haklay, 2010; Mooney et al., 2010]). They all shared similar conclusions after comparing these datasets with official records: Whereas completeness and precision varies, OSM accomplished to create geographical information comparable to authoritative data in at least more densely populated regions of the Western world [Richter, 2013]. A lack of coverage is stated in rural and less developed areas leading to an inconsistency in terms of quality [Haklay, 2010]. Haklay [2010] raised the question whether there is a critical value in decrease of quality and data density where the information is no longer useful for cartographic output and analysis. However, despite the identified weaknesses, the advantage of the free access to a vast amount of geographical data is a major benefit compared to chargeable data sources [Haklay, 2010].

2.3.3 Data Primitives in OpenStreetMap

All data in OSM are stored as XML objects. XML is a common language for the interchange of data over the internet [Nurseitov et al., 2009]. The conceptual model consists of three basic geometric components: *Nodes*, *ways* and *relations*. The properties of these elements are further described by using so-called *tags* [OSM, 2011]:

Node:
A node represents a specific point. All nodes in OSM include an id and a location defined by WGS84 coordinates (the attributes lat / lon). Nodes can describe standalone features such as fountains or cash machines etc. or may be a part of a more complex

geometry (way / relation) to describe vertices of polygons. Nodes are the only geometric primitives in OSM which actually include an absolute geographic position in the form of coordinates. Code-Fragment 2.1 shows an example of how nodes are stored in OSM. This particular instance represents a waste basket. Besides the id and the location, the listed attributes describe, among others, the time of the last update (timestamp), the user responsible for the recent change (user) as well as the number of revisions in total (version).

```
<node id="638966072" visible="true" version="2" changeset="5089193" timestamp="2010-06-27T12:41:05Z" user="ThePacki" uid="666" lat="47.3734111" lon="8.5438382">
 <tag k="amenity" v="waste_basket"/>
</node>
```

Code-Fragment 2.1: Example of a node in OSM [OSM, 2014]

Way:
A way consists of at least two nodes to form a polyline for example roads or streets. As illustrated in Code-Fragment 2.2, ways contain references to the nodes of which they consist, stored in the form of "<nd ref="26853074"/>". If the first node is equal to the last one (Closed ways), they represent polygons describing the geometry of areas for example of buildings.

```
 <way id="51414044" visible="true" version="2" changeset="9797423" timestamp="2011-11-11T15:51:01Z" user="ThePacki" uid="666">
 <nd ref="26853074"/>
 <nd ref="656327524"/>
 <tag k="highway" v="pedestrian"/>
 <tag k="name" v="Spitalgasse"/>
 <tag k="oneway" v="yes"/>
</way>
```

Code-Fragment 2.2: Example of a way in OSM [OSM, 2014]

Relation:
Relations are used to define logical or geographical relationships between elements [OSM, 2014c]. They can have several purposes: A relation can for example consist of several polygons to describe a multipolygon, e.g. a building with an inner and an outer geometry. Relations can also express route relations such as a bus route including all stops and paths. They contain references to ways or nodes representing their members. Members are represented as "<member type="way" ref="4415453" role="from"/>". Code-Fragment 2.3 shows an example of a relation describing a turn restriction.

```
  <relation id="301705">
   <member type="way" ref="4415453" role="from"/>
```

```
    <member type="node" ref="44850081" role="via"/>
    <member type="way" ref="11943595" role="to"/>
    <tag k="restriction" v="no_left_turn"/>
    <tag k="type" v="restriction"/>
  </relation>
```

Code-Fragment 2.3: Example of a relation in OSM [OSM, 2014]

Tag:
All elements can be described in more detail by using tags. A tag is a key/value pair. For example the key "name" constitutes the name of an arbitrary element, whereas the value contains the associated name. There is no fixed list of tags, thus every registered user can create own tags. However, it is recommended to use existing tags, unless there is none for the desired purpose [OSM, 2014e].

2.3.4 Points of Interest in OpenStreetMap

Points of interest are particular locations on the earth, which are either interesting or useful for a specific task [Ghasemi, 2011]. For example, restaurants are considered as typical POIs, as people like to be informed about possible places to eat. Although POIs sometimes serve as alternatives for landmarks in wayfinding services, there are major differences between them [Nothegger et al., 2004]: POIs are primarily intended to inform users about nearby features such as offered services (e.g. cash machines), attractive locations (e.g. tourist sites) or to raise the user's awareness, for example in the case of a nearby speed camera. In contrast, landmarks are in the first place defined by their distinctiveness. Another difference lies in the purpose of the landmarks: POIs are predominantly used to provide information. Landmarks on the other hand are an instrument to guide the intended way. The usage of POIs as substitute for landmarks is therefore not suggested without further consideration of them. Nevertheless, depending on their salience, POIs can be ideally suited as landmarks in routing instructions.

Although the name "Point of interest" indicates that they are restricted to point data, OpenStreetMap follows a different approach: Every element in OSM should only be mapped once according to the "One feature, one OSM element"-practice [OSM, 2014b]. As a consequence, POIs in OSM are also mapped as areas, if their shape is relevant [OSM, 2013, 2014b]. Thus, if one wishes to extract landmark information from OSM datasets, all three basic geometric components (Node, Way, Area) have to be included.

2.3.5 Landmarks in OpenStreetMap

While searching the OSM wiki for landmarks, one discovers two usages about the term "landmark" [Ghasemi, 2011]:

- "seamark:type=landmark"
  This tag is specified for the usage in combination with nautical landmarks. The

definition and use of the term "landmark" in this thesis, however, differs strongly from the conventional usage of this tag. The focus in this work lies on complementing routing instructions with landmarks. The stated tag is intended to describe reference points at sea. Thus, the semantic is not transferable to navigation by land.

- naptan:landmark = *
  NaPTAN are official datasets for bus stops in the United Kingdom which have been offered to the OSM Project [OSM, 2010]. Landmarks in this case are distinct intersections or buildings next to bus stops in the UK. Thus, this usage is very specific to UK and does not offer a holistic solution to describe landmarks.

Landmarks according to the definition of this work have not been introduced officially into OSM. Ghasemi [2011] proposed to introduce a new tag in OSM, allowing a quick identification of landmarks. Nevertheless, as of today, there is no uniform landmark-tag accepted by the community. This, however, is a requirement to enable a widespread access to the stored landmark information.

2.3.6 The Overpass API

The Overpass API is a read-only language to query data stored on the OpenStreetMap servers. Different to the main OSM API, which is mainly used for editing data, the Overpass API is specifically optimized to provide data selected by tags, location etc. [Olbricht, 2011]. The queries are performed on the main OSM server, which means it allows access to the live data on the servers. The Overpass query "node ["amenity"="cafe"];out;" for example, returns an XML file including all nodes categorised as "cafe". An important feature of Overpass is the possibility to query ways and relations recursively, thus enabling to download all associated nodes to derive their geometry.

3 Developing the Application

In order to answer the research questions stated in Section 1.2, a mobile application is required to enable the process of landmark-collecting on smartphones. Thus, the following section illustrates the planning and developing of this application. The goal is to provide a tool allowing users to mark objects in their environment as landmarks. Furthermore, this application should permit to share this information on OpenStreetMap.

The identification of suitable landmark candidates is to be achieved through the usage of integrated smartphone sensors. As the involved sensors are limited in their accuracy, several methods are introduced focusing on an improvement of the correctness of the results. This requires knowing the approximate degree of this inaccuracy. Thus, the discrepancy of the sensor is estimated in this section by consulting suitable literature and by performing field tests.

3.1 Application Overview

In the process of this work, a mobile application is developed allowing the user to label objects as landmarks. The intended functionality is as follows: By taking a picture of the desired landmark the user receives a list of candidates, ranked by their possibility of being the intended object. The suggested candidates need to be manually confirmed by the user in order to save them. Furthermore, the application enables the sharing of the gathered data on OpenStreetMap.

The necessary system requirements to use the application are to have a mobile internet connection, an integrated back camera, a built-in GPS receiver and a compass. The geographical data is obtained from OSM. Possible landmarks are limited to point data or polygons representing buildings since the capturing of paths and larger areas is rather difficult to achieve with a camera. Detailed information about this and further constraints is given in Section 3.3.1.

Numerous existing methods for the purpose of extracting landmarks are mainly oriented towards larger scales (see for example Papadopoulos et al. [2010]). In contrast, the planned application allows the tagging of landmarks on a scale only limited by the level of detail of the underlying OSM data. Android was chosen as the development platform due to its high market share (around 80% in 2014 [Lomas, 2014]) and the openness of its system. Applications on Android are primarily coded in the Java programming language. This also applies to the application developed in the frame of this thesis.

The built-in compass and GPS sensors of smartphones are used to determine the area in front of the camera. The combination of these two sensors allows to substantially narrow down the number of potential landmarks to a small amount by computing a representative field of view (FOV) of the user. The application looks for suitable landmarks in this selection. In order to circumvent possible errors in determining the FOV due to sensors inaccuracies, ways of providing more reliable results are needed (see Section 3.2.3).

Not all objects are equally well suited as landmarks. Thus, a ranking is introduced determining the object's "landmarkness". By involving visual- and semantic characteristics, this system should prevent that less suited objects, for instance a common waste basket, are considered more important than for example a museum behind it (see Section 3.4).

3.2 Mobile Sensors

As previously discussed, the application uses integrated smartphone sensors to calculate the field of view of the user. Smartphones use low-cost hardware parts, hence the sensors are afflicted with certain inaccuracies [Bauer, 2013]. For the planned application, it is essential to know this discrepancy in order to suggest the desired landmark. This section presents the used sensors, their limitations and discusses problems involved in connection with their usage.

3.2.1 GPS Sensor

To achieve the specified goals, the most accurate positioning method available is needed. A method is considered accurate if there is a high level of agreement of the provided data with the actual true value [NCSU, 2004]. GPS is therefore the only viable option, as alternatives such as the Wi-Fi-based positioning system (WPS) do not reach comparable accuracy and coverage [von Watzdorf and Michahelles, 2010]. The public version of GPS reaches a maximum accuracy of up to 5 to 10 meters [Bauer, 2013]. This can also be achieved by low-priced receivers as common in smartphones. The usage of GPS, however, involves two challenges regarding the planned application:

First of all, the GPS needs time to activate the sensor, to locate all available satellites and to compute the best location-fix. During this process (hereinafter referred to as GPS latency), the provided location varies considerably and can be highly inaccurate. The time of this latency is thereby notably higher compared to other available positioning methods [Brimicombe and Li, 2009]. Secondly, the GPS accuracy can vary strongly, for example in the presence of large buildings [Brimicombe and Li, 2009]. This is called the "urban canyon effect". Modsching et al. [2006] measured an inaccuracy of up to 15 meters in the presence of surrounding houses even in wide streets.

Regarding the planned application, a low GPS accuracy can lead to significant problems as shown in Figure 3.1. In this scenario, the program is not able to detect the indicated target (the pharmacy) due to a GPS deviation of 15 meters. Although the compass returns an accurate result, the nearby restaurant will be shown as a suitable landmark candidate. Errors like this cannot be fully avoided. Section 3.2.3 explains several approaches to downsize the number of wrong results due to GPS inaccuracies.

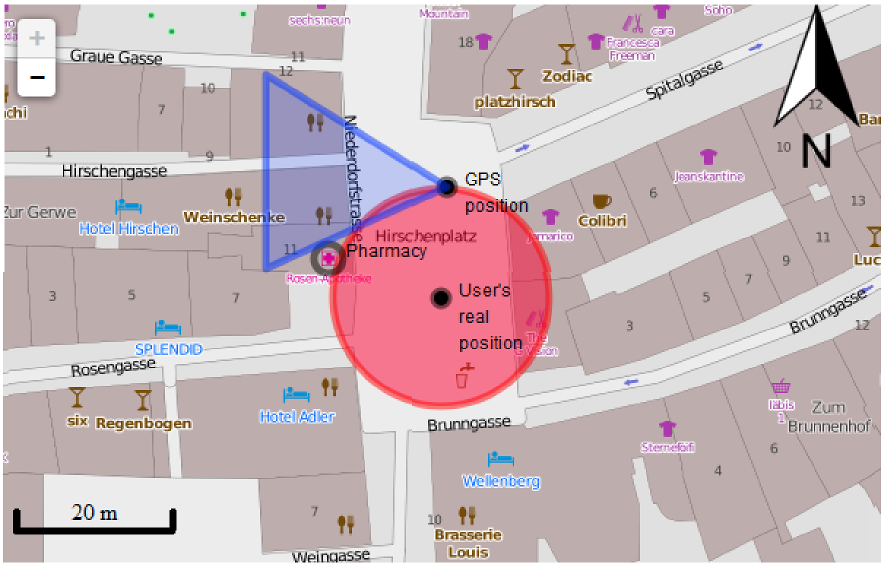


Figure 3.1: Missed landmark due to a miscalculated field of view (blue) because of GPS inaccuracies - Red circle: Inaccuracy radius of 15 meters [Map source: OSM, 2014]

For every location update, the Android API provides an associated accuracy value. The Google API defines this value as ”the radius of 68% confidence” i.e., there is a 68 % chance, that the true location is inside the circle spanned by the given accuracy radius [Google, 2014b]. While testing in the field with an HTC One, the accuracy of the GPS reached values around 4 - 10 meters in open areas and 10 - 30 meters in the presence of surrounding buildings. The latency time was around 15 - 40 seconds depending on the satellite visibility. A key finding of the field tests was that the real inaccuracy occasionally differed significantly from the presented value. In these cases, it was necessary to wait for a few seconds until the provided location was within the given radius of inaccuracy.

3.2.2 Compass Sensor

Only few studies have been conducted to measure the reliability of the integrated compass of handheld devices. One study was performed by Blum et al. [2013]. They tested 3 different smartphone devices and determined a mean compass error of +/- 10 - 30° while walking along a predefined path.

The cited study only measured the compass accuracy whilst moving. People, however, normally do not walk while taking a picture. Moreover, the accuracy of the sensors as well as the calibration of the sensor differs from phone to phone [Blum et al., 2013]. It is therefore necessary to perform an additional test for a viable mean compass error value for the use in the application.

The Android API offers two different types of compass data [Google, 2014d]:

- **TYPE_ORIENTATION:**
  This constant delivers compass data of the cell phone, which has already been processed by the system. The provided value is considered out-of-date, since problems occurred concerning devices with landscape-mode (tablets) as their default orientation [Morrill, 2010].
- **TYPE_ACCELEROMETER and TYPE_MAGNETIC_FIELD:**
  These constants are the current orientation sensor values. They deliver the raw data of the accelerometer and the magnetic field from the sensors. In contrast to the orientation constant, these values provide the three-dimensional position of the cell phone.

The planned application is not intended for the usage with tablets since they often do not possess a GPS sensor, a back camera or a mobile internet connection. Thus, it can be considered as non-critical to use the deprecated constant.

In order to obtain a reliable value for a mean error of the smartphone compass, a test was performed involving a mirror compass with a magnetic needle as ground truth (Model: Recta DP-2) and two different cell phones (HTC One and Samsung Galaxy S II). Common mirror compasses have an accuracy of +/- 1 - 3° [Kahl, 1991]. Additionally, the magnetic field declination in Switzerland is around + 2° [Maus et al., 2010]. Thus, a relatively high accuracy with a mean error of approximately 4° can be expected. This value is significantly superior to the mean compass error from smartphones of +/- 10 - 30° stated in Blum et al. [2013]. Due to this fact and the higher transparency based on its simpler construction, the mirror compass is used as a reference value in this test.

The testing sequence was as follows:

1. Ensuring that there are no outside influences to the compass such as metallic items.
2. Pointing at a distant object with the mirror compass to derive a reference value.
3. Taking 5 pictures from the same object in the same angle with both smartphones.

This process involves two potential sources of inaccuracies: Firstly, possible errors while reading the mirror compass and secondly inconsistencies in the position of the smartphone.

The acceleration sensor was recalibrated previous to the test to ensure the correctness of the sensor reference. To simulate a realistic process of taking a photo, the experiment was performed without changing the location during the recording. Hence, the accelerometer plays a minor role in these results. This may increase the overall mean error of the sensor.

Table 3.1 shows the results of this test. The outputs of the cell phones are compared to the compass value to retrieve a mean error of the sensor.

Compass	HTC One O[1]	HTC One MA[2]	Samsung Galaxy S II O	Samsung Galaxy S II MA
332°	309°	300°	352°	349°
	321°	311°	355°	1°
	313°	330°	353°	6°
	318°	299°	352°	352°
	318°	305°	355°	2°
23°	28°	42°	23°	3°
	30°	45°	20°	5°
	31°	39°	34°	43°
	28°	35°	34°	33°
	27°	28°	36°	37°
168°	174°	179°	134°	159°
	172°	172°	133°	183°
	170°	165°	132°	143°
	171°	175°	134°	151°
	172°	187°	134°	145°
Mean error:	8.6°	15.5°	21.2°	20.1°
Std. deviation:	2.35°	8.18°	2.24°	5.35°

Table 3.1: Comparison of values derived from smartphone compasses to the corresponding values of a compass using a magnetic needle

Following conclusions were obtained during this test:

1. Both methods have inaccuracies on both smartphones. The average mean error was approximately +/- 16°.
2. The newer phone (HTC One) has a better sensor with an average mean error of around +/- 12°.
3. When the pitch (y-axis) exceeded 90°, the magnetic field / accelerometer constants returned their results partially shifted by 180°. These cases are not included in Table 3.1, as they distort the calculation of the mean error.
4. The standard deviation of the orientation constant is lower compared to the magnetic / accelerometer constants, therefore providing a more stable result.

The orientation constant has a lower mean error on the newer phone and a slightly higher inaccuracy on the older one and provides more stable results. However, as this

---

[1] Orientation constant

[2] Magnetic field- and accelerometer constant

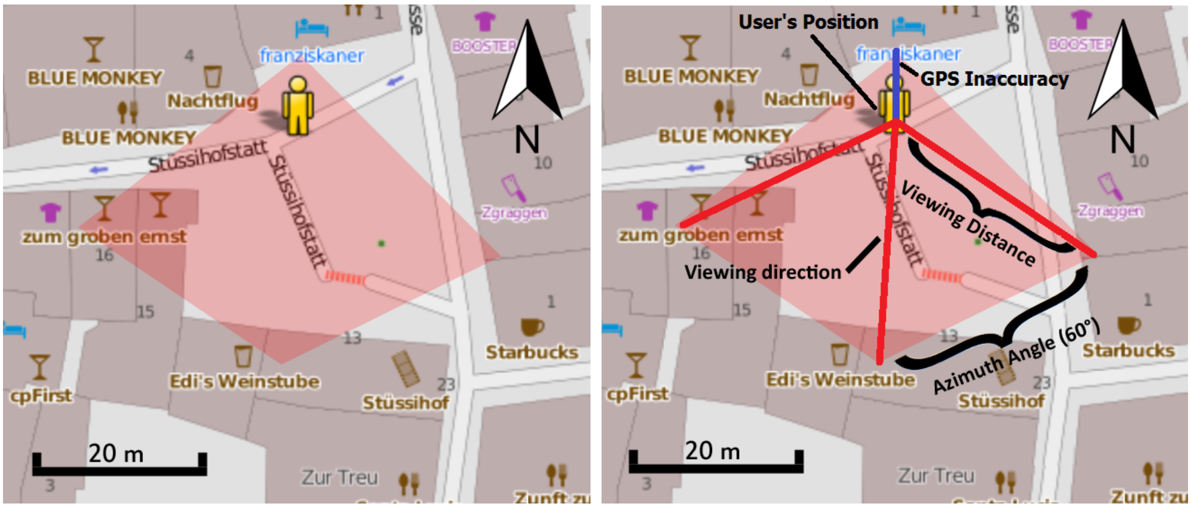
constant is deprecated, it is possible that future phones will no longer support this value. Therefore, the application uses the average of both constants and assumes a mean error of +/- 16°. This would allow a quick adjustment of the application if the orientation constant can no longer be used in future Android versions. However, if the values differ more than 30° from each other, only the orientation constant is used due to the shifting by 180° in certain cases (see Point 3 of the conclusions above).

3.2.3 Dealing with Sensor Uncertainties

As discussed in the previous sections, the application must deal with certain inaccuracies of the used sensors, which can prevent the program from suggesting the proper landmark candidate. This section introduces several strategies to overcome this problem:

Visualization of the Extracted Sensor Results

As shown in Figure 3.2, the FOV (congruent with the area from which the OSM data is downloaded) as well as the location of the user is displayed after taking the picture. Thus, the user can assess the sensor performance (orientation and location) while the result is displayed. Moreover, a map shows the currently provided position before taking a picture.



(a) The field of view without legend (b) The field of view with legend

Figure 3.2: Representation of the field of view (red) in the application. The adjustable viewing distance is set to 50 meters in the example. [Map source: OSM, 2014]

Minimum Accuracy for GPS

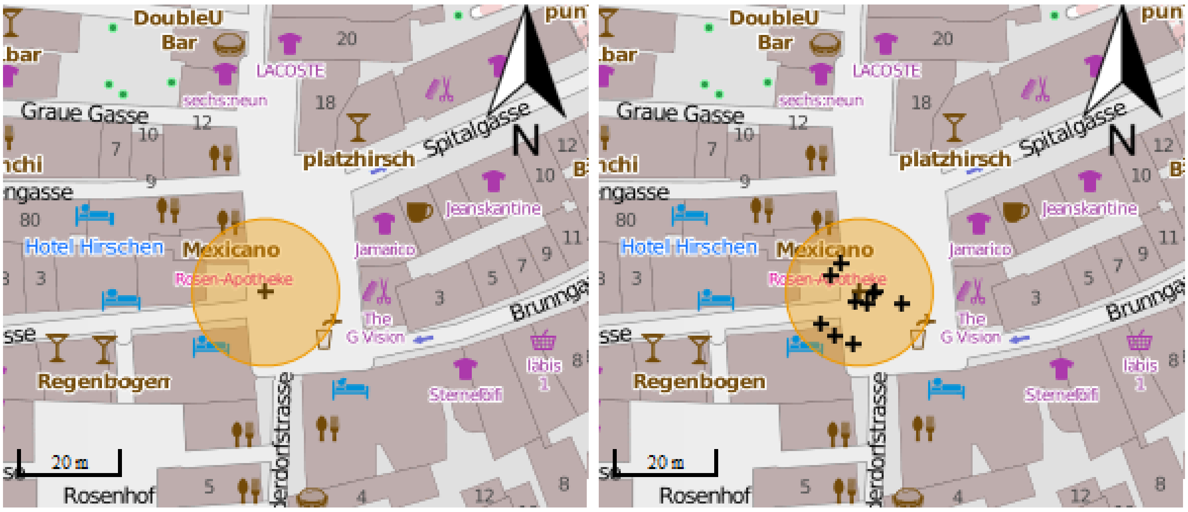
The current accuracy of the GPS is displayed before collecting a landmark. If this value exceeds 15 meters, the program refuses to take a picture. During field tests, an accuracy radius greater than 15 meters caused highly unreliable results. In addition, values below this threshold indicated in most cases a fixed GPS location. The provided GPS accuracy is included in the FOV, indicated by the blue line in Figure 3.2b.

Increase the Number of Suggested Landmark-Candidates

Provided that there are a sufficient number of objects, the user receives up to 5 suggestions of landmark candidates after taking a picture, ranked by their probability of being the photographed landmark. This increases the chance, that the intended landmark is included in the results.

Location Sampling

When calculating the distance or the azimuth from the phone to an arbitrary object the GPS inaccuracy is included. This is implemented by using a location sampling (see Figure 3.3). The distance as well as the bearing is calculated from 10 randomly chosen points inside the "confidence-circle" (a circle with the radius of the given GPS inaccuracy). The application calculates the average distance and azimuth out of this 10 points to process the data. It is intended that this should provide more stable results with less outliers.



(a) Without location sampling

(b) With location sampling - Black dots: Sampled locations

Figure 3.3: Location sampling with a GPS accuracy of 15 meters - Yellow circle: Inaccuracy radius of 15 meters [Map source: OSM, 2014]

Including Azimuth Deviation

The sensor review in Section 3.2.2 showed a mean compass error of around 16° according to the tested devices. Therefore, the azimuth of an object is allowed to have a deviation of 16° to the provided compass value in either direction and will still be considered as in front of the user.

Adjusting the Viewing Angle

The viewing angle is fixed at 120° in order to avoid missing landmarks (see Figure 3.1). In tests this value provided a good balance between missed objects and including too many elements.

3.3 Extracting and Storing Landmark Information with OpenStreetMap

As stated in Section 2.3.3, all data on the OpenStreetMap servers are stored in the form of XML objects. One challenge of the planned application is to extract the raw OSM data and to convert it to landmark-candidates. This section describes this process and its underlying elements and explains in detail how landmark information is derived from OSM objects.

3.3.1 Potential Landmarks in the Application

Landmarks exist in many ways and forms. However, this approach intends to label landmarks through the usage of the camera. It is therefore difficult or even impossible to capture larger areas or paths such as roads and rivers, as a camera is only able to cover specific parts of their geometry. Most algorithms calculating routes in a network with present landmarks rely on point-like landmarks. By sticking to point objects, the algorithms can reduce the relation between the navigator and the landmark to distance and orientation [Caduff and Timpf, 2005]. Nevertheless, as discussed in Section 2.3.4, several POIs and thus potentially suited landmarks as well as other salient objects are only stored as polygons in OSM. The large part of these cases is represented by buildings. As a consequence, the application also allows to label polygons with the "building"-tag as landmarks next to point objects. Thus, potential landmark-candidates are limited to standalone point data and buildings.

Not all point categories are included in the results: Categories which frequently occur in clusters at certain places such as pedestrian crossings or traffic signals are excluded since the user would not be able to clearly assign the suggested candidates to the correct real world object. Some point objects are discarded since they only appear as part of larger units such as building entrances. Vegetation is also excluded, based on the fact that trees etc. are often subject to rapid change and are thus often unreliable landmarks [Vinson, 1999]. These omissions do not apply if the objects include a specific "name"-tag, as this indicates a certain importance of the location.

Code-Fragment 3.1 shows examples of typical landmark-candidates from OSM. The objects all contain a "name"-tag and a category. The node- and way example use the keyword "amenity" for the purpose of describing their type. The "amenity"-tag covers a number of community facilities (e.g. restaurants, theatres, bar etc.) [OSM, 2012a]. The presented relation uses the keyword "tourism" to describe its category, indicating that this object is of particular interest for tourists [OSM, 2012b]. Several other tags included in these examples are utilized in the application, for instance those referring to the address ("addr:") or include links to additional information ("website", "wikipedia"). The exact purpose of these tags within the scope of the developed software is described in the next section.

```
<node id="317774611" lat="47.3697432" lon="8.5446131">
  <tag k="addr:city" v="Zuerich"/>
  <tag k="addr:country" v="CH"/>
  <tag k="addr:housenumber" v="14"/>
  <tag k="addr:postcode" v="8001"/>
  <tag k="addr:street" v="Kirchgasse"/>
  <tag k="amenity" v="restaurant"/>
  <tag k="cuisine" v="bistro"/>
  <tag k="name" v="Karl Der Grosse"/>
  <tag k="website" v="http://www.stadt-zuerich.ch/karldergrosse"/>
  <tag k="wheelchair" v="no"/>
</node>
<way id="292437788">
  <nd ref="2860465381"/>
  <nd ref="2959935291"/>
  <nd ref="2959935294"/>
  <nd ref="2959935297"/>
  <nd ref="2860465358"/>
  <nd ref="2860465355"/>
  <nd ref="2860465361"/>
  <nd ref="2860465381"/>
  <tag k="amenity" v="place_of_worship"/>
  <tag k="building" v="yes"/>
  <tag k="denomination" v="protestant"/>
  <tag k="name" v="Grossmuensterkapelle"/>
  <tag k="religion" v="christian"/>
  <tag k="start_date" v="1859"/>
</way>
<relation id="3768127">
  <member type="way" ref="147633773" role="outer"/>
  <member type="way" ref="284013752" role="inner"/>
  <tag k="addr:country" v="CH"/>
  <tag k="addr:housenumber" v="1"/>
  <tag k="addr:street" v="Heimplatz"/>
  <tag k="building" v="yes"/>
  <tag k="name" v="Kunsthaus"/>
  <tag k="tourism" v="gallery"/>
  <tag k="type" v="multipolygon"/>
  <tag k="website" v="http://www.kunsthaus.ch/"/>
  <tag k="wheelchair" v="yes"/>
  <tag k="wikipedia" v="de:Kunsthaus Zuerich"/>
</relation>
```

Code-Fragment 3.1: Example of typical landmark-candidates in OSM [OSM, 2014]

3.3.2 Transforming OpenStreetMap Data into Landmark-Candidates

Before the OSM objects are converted into landmark-candidates, the objects in the FOV not meeting the minimum requirements are sorted out. The minimum demands of a landmark candidate in the application are to include **a location**, at least **one tag** and the possibility to derive **a name** out of the object's metadata. This means, that the name must either be available as a ”name”-tag or otherwise other attributes, for example category information or the address, must allow an appropriate naming (e.g. railway=”bus_station” is named ”Bus Station”). This ensures that the landmark can be referred to and the user is able to recognise the suggested landmarks. However, the consequence of this naming method is, that if there is no available ”name”-tag or address, objects with categories unknown to the application are automatically ignored. Ways and relations also must include the ”building”-tag due to the limitation to point data and buildings (see Section 3.3.1). All objects in the FOV which do not meet these minimum requirements are sorted out. The remaining objects represent the involved landmark candidates.

In order to calculate the most probable candidate, a quantification of the given OSM metadata and sensor information has to be performed. For this reason, attributes must be derived from the available data to measure ”landmarkness”. The included OSM properties need to be available for a large part of objects in order to include as many candidates as possible. To ensure a connection of the attributes to the term ”landmark”, they are assigned to the characteristics of Burnett et al. [2001] (see Section 2.1.1). The resulting allocation of the involved OSM- and sensor data is presented in Table 3.2. Apart from area and visible range, which are only available for ways and relations and the background information, the enumerated underlying data can be derived from all OSM objects that include at least present category information. An important aspect in defining these characteristics was, to keep the computational effort to derive them within reasonable limits to ensure a high level of performance. As a consequence thereof, the characteristic ”Usefulness of Location” is not calculated by the application itself but is determined by the user. A direct measurement would be rather challenging, as this would require involving factors such as the structural use, the accessibility and the role of the object within the network.

Characteristic	**Underlying Data**
Permanence	-Average permanency of a typical instance of this category
Visibility	-Distance to the user -Azimuth deviation to the camera -Area of the object -Visible range: The angle range in which an object is visible to the user -Average salience of a typical instance of this category
Uniqueness	-Frequency of a specific category (number of objects with the same category in the surrounding area, general rarity of occurrences of this specific category) -Estimated significance of the object (background information) -Number of tags of the object
Usefulness of Location	-Determined by the user
Brevity	-Given by the constraint of only including landmarks that offer the possibility for an appropriate naming (see above)

Table 3.2: The characteristics from Burnett et al. [2001] described by OSM- and sensor data

To allow a comparison between various values for the ”landmarkness” of different objects, the listed attributes from Table 3.2 are translated into the three main characteristics of landmarks [Sorrows and Hirtle, 1999] (see Section 2.1). Figure 3.4 shows a modified version of the previously shown Figure 2.1. The green fields show the available identifying factors. Uncoloured characteristics are either only occasionally available (e.g. *Color*), not available at all (*Shape* in 3-Dimensions) or too complex to derive from OSM data (*Structural Significance*).

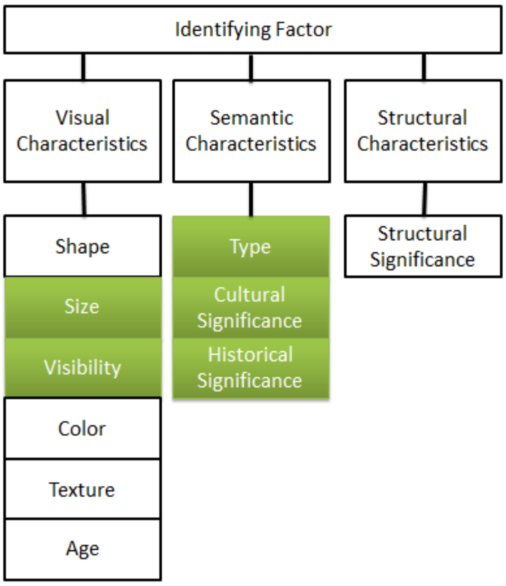


Figure 3.4: The identifying factors defining the saliency of a landmark. Green boxes show the characteristics which are derived from OSM data in the application [Original by Richter and Winter, 2011]

The specified data from Table 3.2 is allocated as follows:

Size: Area (only available for ways / relations)

Visibility: Distance and azimuth to the user, visible range (only ways / relations)

Type: Tags describing the function / category of an object (e.g. *amenity, leisure, shop*)

Cultural / Historical Significance: Number of tags, background information (own website / Wikipedia article[1]), frequency of the category

Contrary to other approaches considering the extraction of landmarks which also include the characteristics of Sorrows and Hirtle [1999], the presented method integrates the sensor data of the mobile phones. The visual characteristics are thereby determining

---

[1] `www.wikipedia.com`

the user's intention and the visibility of an object, whereas the semantic characteristics are measuring how suitable a certain object is as a landmark according to its tags. Section 3.4 describes the process of comparing different objects in the FOV in regard to their "landmarkness'.

3.3.3 Storing Landmark Information on the OpenStreetMap Servers

The application includes a possibility to store the gathered information on the OSM servers. The new tag **"uzh_landmark=true"** is added to label the affected objects since there is no official landmark tag available (see Section 2.3.5). The abbreviation "uzh" is thereby indicating the origin of the tag (Universität Zürich). As pictures cannot be stored directly on the OSM servers, the user will not be able to upload the photographs taken in the final product. This could be added in a later version, combined with further information describing the landmark (reasons for being a landmark, salient features etc.).

3.4 Ranking of Landmarks

Geographical objects have differing suitability to act as landmarks. Therefore, this thesis introduces a ranking system based on the object's metadata and visibility to find the most appropriate landmark.

Based on the three main categories from Sorrows and Hirtle (see Section 2.1 & 3.3.2), Raubal and Winter [2002] developed a measure for determining the saliency of a specific landmark:

$$s_{vis} \cdot w_{vis} + s_{sem} \cdot w_{sem} + s_{str} \cdot w_{str} \tag{3.1}$$

s stands for the salience measure and w is an individual weighting factor (weighting factors are further explained in Section 3.4.3). The parameters vis, sem and str describe visual, semantic (cognitive) and structural saliency [Richter, 2007].

As described in Section 3.3.2, structural characteristics are not taken into consideration in the frame of this application. However, as this method of collecting landmarks is user-generated, one can assume that users will automatically choose structurally adequate landmarks (e.g. the corner of a building complex, a central spot etc.). Leaving out structural characteristics in Formula 3.1, following equation is defined to calculate the overall ranking of a landmark candidate:

$$\psi_{ranking\_value} = s_{vis} \cdot w_{vis} + s_{sem} \cdot w_{sem} \tag{3.2}$$

The remaining parameters s_{vis} and s_{sem} are calculated using the attributes explained in Section 3.3.2. The semantic part of the equation can be compared to automatic landmark extraction algorithms, as it is only reliant on the metadata. In contrast, the visual ranking is based on the position of the user and the size of the object.

3.4.1 Calculating the Ranking Factor for Visual Characteristics

In order to derive the factor s_{vis} describing the visual suitability, the formulas within this section are used. The involved parameters are X_{size} (Size), X_{vis_range} (Visible Range), $X_{azimuth}$ (Azimuth) and $X_{distance}$ (Distance). i refers to the current object and min / max to the respective minimum / maximum value of all objects. A normalization is performed for each parameter, to bring all values separately into the range [0,1]. Divisions by zero are handled in all cases with the return of the value 0.

Size (Only Ways / Relations):

$$X_{size} = \frac{X_{i\_size} - X_{min\_size}}{X_{max\_size} - X_{min\_size}} \tag{3.3}$$

The Visible Range (Only Ways / Relations)

$$X_{vis\_range} = \frac{X_{i\_range} - X_{min\_range}}{X_{max\_range} - X_{min\_range}} \tag{3.4}$$

The formulas for $X_{azimuth}$ and $X_{distance}$ are squared. This leads to a higher prioritization of nearness and compliance with the azimuth:

The Azimuth

$$X_{azimuth} = \left(\frac{180^\circ - ((360^\circ + X_{sensor\_azim} - X_{i\_azim}) \mod 360)}{180^\circ}\right)^2 \tag{3.5}$$

X_{sensor_azim} stands for the sensor's azimuth value. If the azimuth of the object is 180° in the opposite direction, the candidate receives the value 0. If the azimuth value is equal to the sensor's azimuth it obtains the value 1. This formula also considers the allowed azimuth deviation of +/- 16° derived in Section 3.2.3. As a consequence, objects within the range of 16° to the azimuth are also ranked with 1. In the case of a polygon, the most outside edges of the object are used as reference points to calculate the azimuth deviation. If X_{sensor_azim} is between these points or within the range of 16° to the average value of both edges, the normalized value $X_{azimuth}$ is equal to 1.

Distance

$$X_{distance} = \left(\frac{X_{user\_max} - X_{i\_dist}}{X_{user\_max}}\right)^2 \tag{3.6}$$

The maximum viewing distance X_{user_max} is manually defined by the user.

Resulting Formula

These factors result in following equations for the calculation of the factor s_{vis}:

For ways / relations:

$$s_{vis} = X_{size} \cdot w_{size} + X_{azimuth} \cdot w_{azimuth\_way} + X_{distance} \cdot w_{distance\_way} + X_{vis\_range} \cdot w_{vis\_range} \tag{3.7}$$

And for nodes:

$$s_{vis} = X_{azimuth} \cdot w_{azimuth\_node} + X_{distance} \cdot w_{distance\_node} \tag{3.8}$$

3.4.2 Calculating the Ranking Factor for Semantic Characteristics

The factor s_{sem} defining the value of the semantic characteristics is calculated in a similar way. The involved parameters X_{type} (Type) and X_{signif} (Significance) are as well normalized to the range [0,1].

Type

In order to calculate the value X_{type}, a weighting factor for each category is defined describing the "landmarkness" of a typical representative:

$$X_{type} = \frac{X_{i\_type} - X_{min\_type}}{X_{max\_type} - X_{min\_type}} \tag{3.9}$$

X_{min_type} is the lowest and X_{max_type} the highest value category value in the visible field. The category weighting factors will be explained in further detail in Section 3.4.4. Candidates without a category or a category unknown to the application are ranked with $X_{type} = 0$.

Significance

Cultural significance and historical significance are combined to a factor X_{signif}, since no clear distinction can be made between these two factors without checking other sources than the OSM metadata. As stated in Section 3.3.2, this parameter consists of the number of tags X_{tag_n}, the frequency of the category X_{freq} and present background information such as an own website $\phi_{website}$ or an available Wikipedia article ϕ_{wiki}. This results in following equation:

$$X_{signif} = \left(1 - \frac{X_{i\_freq} - X_{min\_freq}}{X_{max\_freq} - X_{min\_freq}}\right) \cdot w_{freq} + \frac{X_{i\_tag\_n} - X_{min\_tag\_n}}{X_{max\_tag\_n} - X_{min\_tag\_n}} \cdot w_{tag\_n} + \phi_{website} + \phi_{wiki} \tag{3.10}$$

Resulting Formula

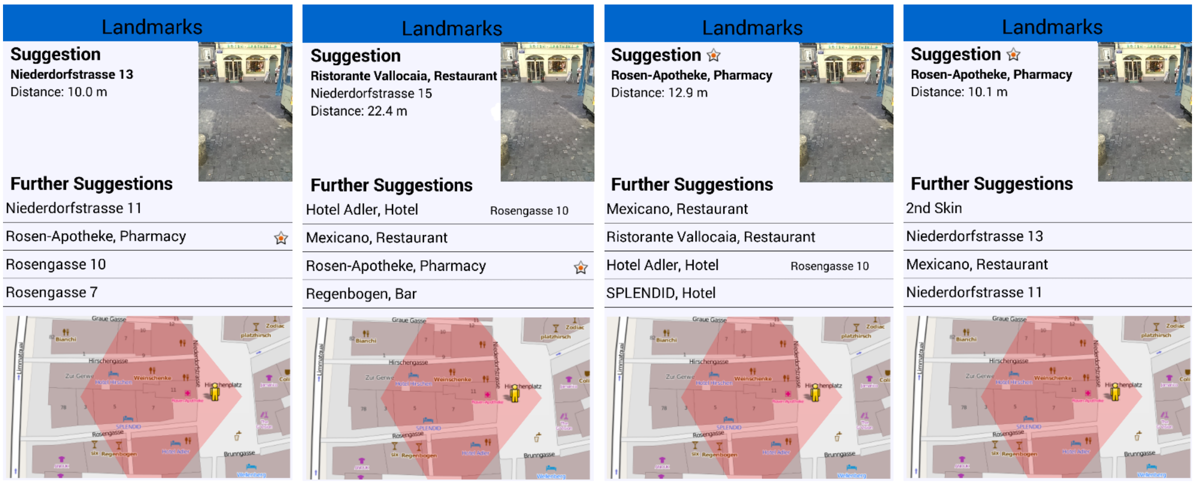
The final formula for the factor s_{sem} is:

$$s_{sem} = X_{type} \cdot w_{type} + X_{signif} \cdot w_{signif} \tag{3.11}$$

3.4.3 Determining the Weighting Parameters

The individual weighting factors were obtained empirically, based on the accuracy of the sensors and the estimated importance of the individual attributes. This means, that these factors are of no general validity, but have performed well in the test areas. Therefore, not all of them are discussed here in detail.

The most influential factors to determine were w_{vis} and w_{sem}. Figure 3.5 shows four different results with alternating values for w_{vis} and w_{sem}. Other factors were kept constant and correspond to the values in the final version. All runs were simulated with the exact same location (distance differences to same objects stem from the location sampling - see Section 3.2.3). The intended landmark is the "Rosen-Apotheke", a pharmacy in the centre of Zurich, directly in front of the user.


Landmarks
Suggestion
Niederdorfstrasse 13
Distance: 10.0 m
Further Suggestions
Niederdorfstrasse 11
Rosen-Apotheke, Pharmacy ☆
Rosengasse 10
Rosengasse 7

Landmarks
Suggestion
Ristorante Vallocaia, Restaurant
Niederdorfstrasse 15
Distance: 22.4 m
Further Suggestions
Hotel Adler, Hotel Rosengasse 10
Mexicano, Restaurant
Rosen-Apotheke, Pharmacy ☆
Regenbogen, Bar

Landmarks
Suggestion ☆
Rosen-Apotheke, Pharmacy
Distance: 12.9 m
Further Suggestions
Mexicano, Restaurant
Ristorante Vallocaia, Restaurant
Hotel Adler, Hotel Rosengasse 10
SPLENDID, Hotel

Landmarks
Suggestion ☆
Rosen-Apotheke, Pharmacy
Distance: 10.1 m
Further Suggestions
2nd Skin
Niederdorfstrasse 13
Mexicano, Restaurant
Niederdorfstrasse 11

(a) Only w_{vis} (b) Only w_{sem} (c) $w_{vis} = 2$, $w_{sem} = 3$ (d) $w_{vis} = 3$, $w_{sem} = 2$

Figure 3.5: Four different value pairs for the parameters w_{vis} and w_{sem} [Map source: OSM, 2014]

The visual ranking places the desired landmark on the third place. Instead, the two buildings "Niederdorfstrasse 11" and "Niederdorfstrasse 13" are ranked on the first two positions as they are nearer to the user, even though the pharmacy is placed in one of these buildings. The semantic ranking lists the desired landmark on the fourth place. Hence, the parameters w_{vis} and w_{sem} are required to create a balance between sensor data and semantic importance. In the third picture the parameters are set to $w_{vis} = 2$, $w_{sem} = 3$. This returns the expected result. However, the two buildings directly in front of the user ("Niederdorfstrasse 11" and "Niederdorfstrasse 13") are excluded from the results, as they do not contain a category and are therefore semantically low rated. With $w_{vis} = 3$ and $w_{sem} = 2$, the user obtains the correct result and also receives the buildings in the list, in case one of them was the desired landmark.

As in the example given above, buildings typically do not contain category information in OSM. Businesses are often situated only in a part of a building. Hence, OSM uses two separate objects to represent the situation: One polygon for the building and a node for the offered service. For this reason, category information is generally only included in the nodes. The typical category tag "amenity" for example has over 100'000 entries in Switzerland whereof only about 11'000 entries are used in combination with the "building"-key [Swiss Taginfo, 2014a, (Status: 25.07.2014)]. In this case, it is standard practice in human route communication to refer only to the establishment in a building instead of the building itself, for example "next to the hairdresser / restaurant" [May et al., 2003]. Hence, the recognition value and saliency of an object are often exclusively based on its provided services. As a consequence thereof, the factor for categories w_{type} is highly prioritized. In the presence of suitable nodes, buildings without a category are thus usually only considered if there is a high level of agreement with the visual characteristics (as in Figure 3.5d).

Weighting factor	**Parameter**	**Weight**
General Factors:		
Visual weight	w_{vis}	3
Semantic weight	w_{sem}	2
Individual visual Characteristics:		
Parameters for ways / relations:		
Size	w_{size_way}	2
Visible range	w_{vis_range}	2
Azimuth deviation	$w_{bearing\_way}$	10
Distance	$w_{distance\_way}$	10
Parameters for nodes:		
Azimuth deviation	$w_{bearing\_node}$	12
Distance	$w_{distance\_node}$	12
Individual semantic Characteristics:		
General semantic Factors:		
Type weight	w_{type}	8
Significance weight	w_{signif}	6
Individual parameters for significance:		
Frequency	w_{freq}	1
Tag Number	w_{tag_n}	1
Bonus for website	$\phi_{website}$	0.3
Bonus for Wikipedia article	ϕ_{wiki}	0.6

Table 3.3: The final weighting factor values in the application

Table 3.3 shows the weighting factors applied in the final application. The sum of all individual visual characteristics for ways / relations and nodes is kept equal. The reason behind the low rating of size and visible range is motivated by the goal, that this application should also be able to capture landmarks on a small scale as mentioned in Section

3.1 and that these attributes only appear on polygons. Furthermore, it can be assumed that there is direct correspondence between these two factors since both attributes are related to the dimension of an object. The factor w_{freq}, describing the frequency of a particular category, does not play a major role, as it provides no information about the saliency of individual objects. Therefore, this value is rated rather low.

3.4.4 Weighting Factors of Categories

Duckham et al. [2010] (see Section 2.1.2) weighted the categories accordingly to how suitable a typical instance is as a landmark. The planned application is following a similar approach: Table 3.2 showed that category data is involved to quantify the three factors permanence, uniqueness and visibility. Thus, a weighting factor for each category / category-class with the range [1,10] is defined, determining the average compliance with these factors. Thereby weight = 1 describes the least significant categories and weight = 10 most suitable landmarks.

This approach, however, has a noteworthy limitation, which was already discussed in Section 2.1.2: For example, while typical churches are highly suitable landmarks, as they are large, recognisable and semantically as well as architecturally distinct from their surroundings, other churches, for instance a church inside an airport, are not considered salient [Duckham et al., 2010]. Yet, in contrast to Duckham et al. [2010], this limitation can be overcome by including the user in the process of acquiring landmarks. Nonetheless, these parameters must be understood as average values of the suitability of typical instances of this category and are not valid for all representatives.

Appendix A shows the resulting list covering most of the relevant categories from Switzerland. Unless otherwise stated, an object with unknown or non-existent category information automatically gets one point. Some objects with the same key, for instance "shop", earn an equal amount of points. This is due to the fact, that shops are normally distinctive due to their display window, their brand or store space and not primarily because of the goods that they sell.

3.5 Application Design

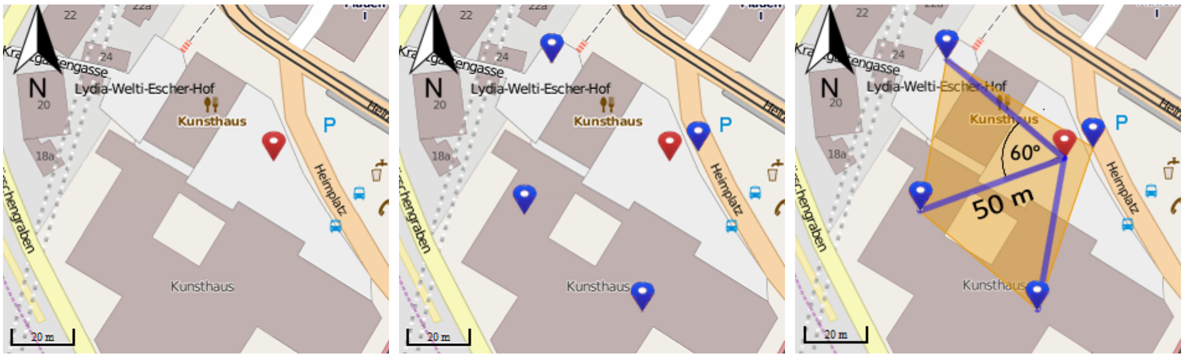
An Android application consists of several activities. An activity is a component of an application, where the user can perform a particular action (e.g. dial a phone, view a map etc.). Each activity has a window where the user interface is drawn [Google, 2014a]. Every application runs its main activity on start-up. After starting the application, the main activity first assures, that there is an available GPS signal and that the cell phone is connected to the internet. If this is not the case, the application requests the user to activate these services. If the user declines this request, the program refuses to collect any landmarks. After the start-up, the time to first fix (TTFF) of the GPS is awaited to download the statistics on category data of the surrounding area. Afterwards the user is able to select one of the available activities. The main activity lets the user choose between the four activities: "Collect a new landmark", "Show all landmarks", "Upload my landmarks" and "Settings".

The GPS is set to refresh its position with a frequency of 1 Hz (1 update per second). This might lead to high power consumption. Yet, the location data must be as current as possible to provide viable results.

3.5.1 Activity: Collect a New Landmark

The activity "Collect a New Landmark" includes every step of labelling an object as a landmark, except for the upload of the information. In order to collect a new landmark, the user has to take a picture of the desired object. Therefore, "Collect a new Landmark" first starts a new camera-activity. In the moment of capturing the landmark, the camera is triggering the method "onShutter()" [Google, 2014c]. The application uses this method to save the azimuth- and location data as near as possible to the moment when the photo was captured.

If the user accepts the picture, a query is sent to OSM using the Overpass API (see Section 2.3.6) and the picture is saved on the phone. As shown in Figure 3.6, a polygon based on the provided location, viewing direction and GPS inaccuracy is created. The three polygon edges in front of the user are in a predefined distance (in this case 50 m) to the location of the phone. The external angles are inclined in an angle of 60° to the line of vision. As explained in Section 3.2.3, the point behind the user represents the distance of the GPS inaccuracy. The query returns all nodes and buildings from this area as an XML file. Due to the limitations of the Overpass API (e.g. no OR-Operator if selecting tags), the client (the phone) has to take over a part of the selection process. In order to do so, the application parses through the received XML file to create objects with their associated properties (location, tags, etc.) and filters out unsuitable objects. The remaining candidates are ranked according to their attributes as described in Section 3.4. The entire process of collecting a landmark is illustrated in Figure 3.7.



(a) User location (red mark)

(b) User location with edges of data polygon (blue marks)

(c) With labels and coloured FOV (orange)

Figure 3.6: Constructing a FOV with a viewing distance of 50 m [Map source: OSM, 2014]

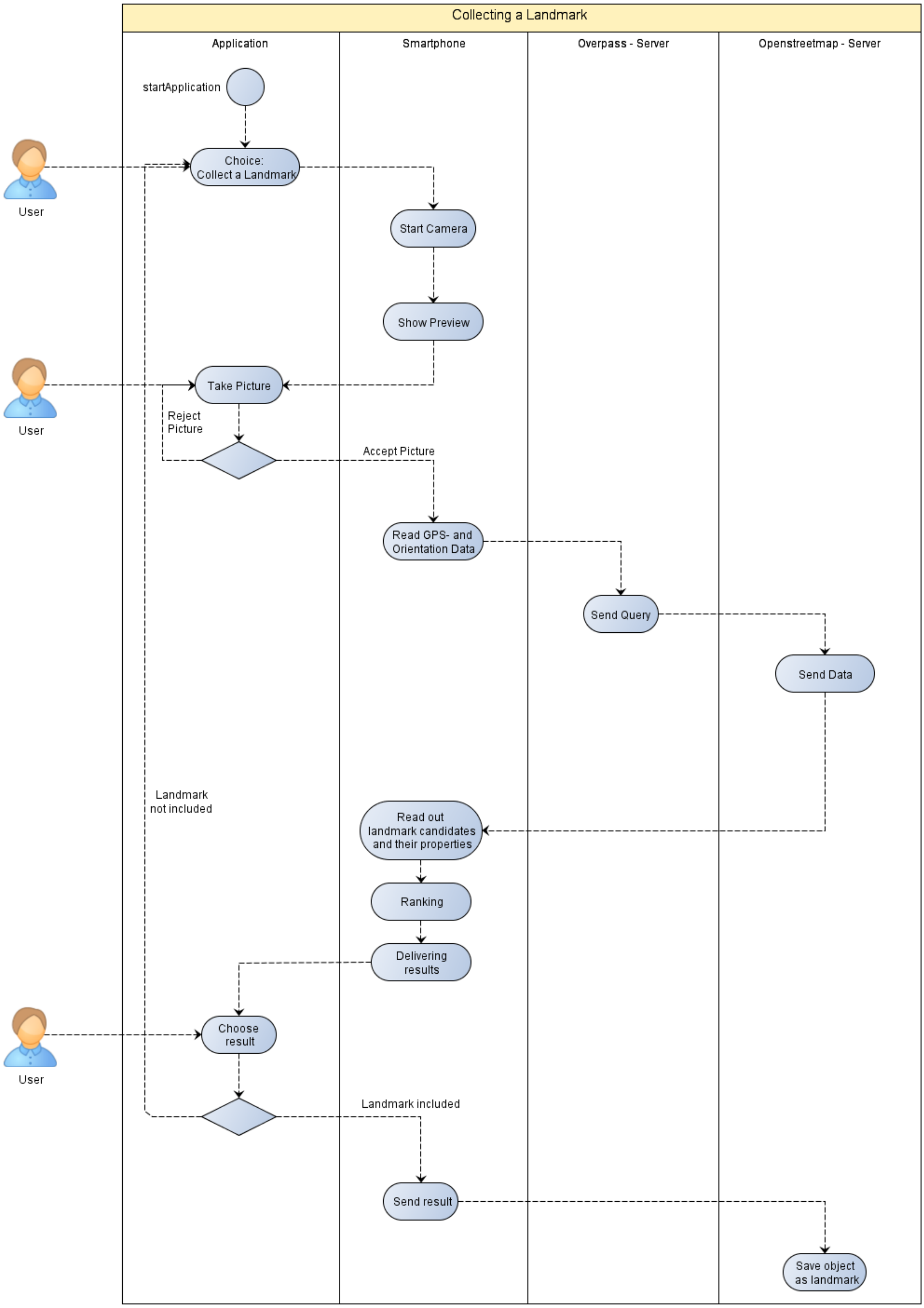
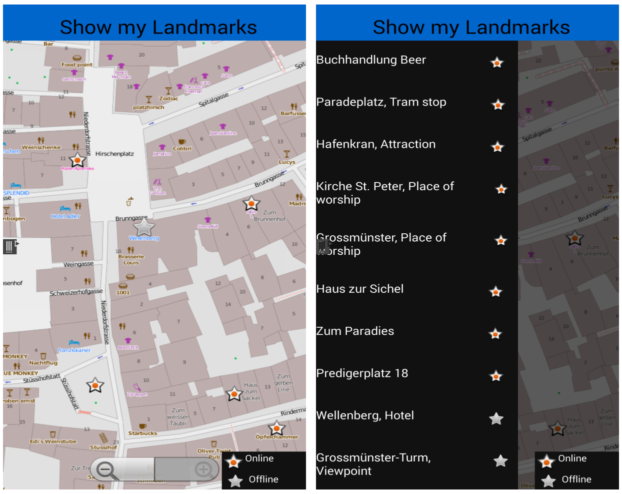


Figure 3.7: Process model of the sequence of collecting a new landmark

The result is presented subsequently in a new activity. This activity displays the five highest ranked landmarks and lets the user choose one of them. The interface of this activity is shown above in Figure 3.5. If the user accepts one of the suggested candidates, it is saved to a local SQLite database; otherwise the application returns to the main menu.

3.5.2 Activity: Show All Landmarks

The "Show All Landmarks"-activity (Figure 3.8) allows the user to browse through all previously gathered landmarks. The activity includes a map showing all landmarks as well as a list of them to pull-out. With this list, the user can zoom to the desired landmarks or delete landmarks which have not been uploaded yet. The program therefore distinguishes between offline and online landmarks. Whereas the online landmarks are stored on the OSM server, the offline landmarks are saved locally.



Buchhandlung Beer
Paradeplatz, Tram stop
Hafenkran, Attraction
Kirche St. Peter, Place of worship
Grossmünster, Place of worship
Haus zur Sichel
Zum Paradies
Predigerplatz 18
Wellenberg, Hotel
Grossmünster-Turm, Viewpoint

(a) The map showing all collected landmarks (b) The list with all the landmarks

Figure 3.8: The "Show all Landmarks"-activity [Map source: OSM, 2014]

3.5.3 Activity: Upload My Landmarks

The "Upload My Landmarks"-activity allows users to upload the locally stored landmarks to the OSM server by using the OSM API v0.6[1]. Subsequently, the landmarks are deleted from the local database. In order to perform the upload, the user needs an OSM account. The application uses the OAuth standard for authorization. This enables third-party applications to access user-functions of a server without sharing any user credentials with the application itself. This activity has adopted a part of the code for the authentication on the OSM servers from the "Open-Landmarks"-project[2].

---

[1] `wiki.openstreetmap.org/wiki/API_v0.6`

[2] `www.openlandmarks.net`

3.5.4 Activity: Settings

In the "Settings"-activity, the user can choose the viewing distance and whether the data should be uploaded to the OSM development server or to the main server. The development server has the restriction that only landmarks in form of a node can be uploaded, as the references of ways and relations do not exist on this server. Therefore, the usage of this server is not recommended in order to save the results and should only be used for testing purposes. The settings also include an option to deactivate the visual or the semantic ranking of landmarks to perform the field test (see Section 4.1).

3.5.5 Calculations

The application uses several calculations to compute the required attributes. Two of them are important mentioning in the methods as they may have an effect on the outcome:

The Haversine Formula for Distance Calculation:
The Haversine formula is a common method to calculate the shortest distance d over the earth's surface [Sinnott, 1984]. This calculation has the advantage of being well suited to determine distances over a short range and that the coordinates do not have to be converted into a Cartesian system [Chamberlain, 1996]. The distance is calculated as following [Veness, 2014]:

$$\begin{aligned} a &= sin^2(\Delta\phi_{1,2}/2) + cos(\phi_1) \cdot cos(\phi_2) \cdot sin^2(\Delta\lambda_{1,2}/2) \\ c &= 2 \cdot atan2(\sqrt{a}, \sqrt{1-a}) \\ d &= R \cdot c \end{aligned} \tag{3.12}$$

where ϕ is the latitude, λ the longitude and R the earth's radius (6371 km). All distances in the application are calculated with this formula.

Calculating Areas for Polygons:
The application abstains from using complex formulas to calculate the area of the objects in return for an improved efficiency. Instead, the area is estimated by multiplying the distance differences between the maximum and minimum values of the object's coordinates i.e., its minimum bounding box. Typically, this leads to an overestimation of the area. However, this can be viewed as non-critical as size is not treated as a significant parameter since only ways and relations possess an area. Additionally, the factor of the overestimation can be expected to remain more or less constant within the field of view.

4 Evaluating the Application

Based on the research questions in Section 1.2, the present thesis aims to show whether smartphones provide a suitable instrumentation to collect landmarks. For this purpose, the application needs a certain level of performance in finding the intended landmark. This calls for an analysis of the application's success rate.

During the evaluation, several landmarks with various characteristics in different contexts are captured while using differing settings for the ranking. Additionally, a user study is conducted to ensure that persons unfamiliar with the application obtain similar results.

4.1 Experimental Setup

The application is evaluated by performing a field test. During this test, two regions are examined: One region represents an urban environment (Zurich) and the other shows a predominantly rural character (Zumikon). The chosen regions exhibit greatly varying conditions in data density and thus in the number of possible landmark candidates. By investigating these areas, conclusions can be drawn about the influence of data density on collecting landmarks.

In order to perform this evaluation, landmarks are captured by using the application. The position in the ranking of the desired objects (see Section 3.4) is logged during this test. While position 1 corresponds to the highest ranked landmark, numbers above 5 are considered as "insufficient" as they are not displayed in the final implementation (see Section 3.2.3). The landmarks are not defined beforehand. Instead they are selected on sight during the process to match the criteria listed below. The specified factors regarding different types of landmarks should ensure a consideration of a diverse range of OSM objects. The examination of the listed scenarios enables to test a wide variety of circumstances and thus to draw conclusions about the performance in different situations. With this procedure, one can determine for which situations and types of landmarks the application is well-suited and in which cases difficulties may occur. The evaluation is therefore highly qualitative, as it does not provide conclusions about the overall population of landmarks.

The different types of landmarks should include:

- Prominent landmarks such as tourist sites, churches etc.
- Less prominent, but nonetheless salient landmarks for example a fountain, a remarkable shop.

A number of different scenarios are examined:

- Areas with and without surrounding buildings

- Less prominent landmarks next to popular landmarks, for example a fountain next to a church
- Low and high POI densities
- Identical category types next to each other (e.g. several shops)

To determine the influence of the ranking, each landmark is captured either by considering visual or semantic characteristics or both characteristics combined. By taking separate rankings into account, conclusions can be drawn about the influence of each characteristic in order to find the correct landmark. Each landmark is captured from a near (5 - 15 meters) and a far (15 - 35 meters) distance. This results in a total of 6 pictures per object (3 ranking variations · 2 locations). Due to the high GPS refresh frequency of 1 second and the positioning inaccuracy, the provided location nay vary between the recordings. Thus, each combination is captured twice to reduce the randomness of the results. Table 4.1 shows the resulting protocol for one landmark as an example. The column "position" states the result of the ranking.

Landmark	Visual Characteristics	Semantic Characteristics	Distance	Position	Description
Rosen Apotheke	yes	yes	near	1	relatively open area, high POI density (restaurants, bars etc.), less prominent landmark
	yes	yes	near	1	
	no	yes	near	3	
	no	yes	near	3	
	yes	no	near	2	
	yes	no	near	5	
	yes	yes	far	2	
	yes	yes	far	2	
	no	yes	far	4	
	no	yes	far	4	
	yes	no	far	4	
	yes	no	far	2	

Table 4.1: An example test protocol for the landmark "Rosen-Apotheke"

The customizable viewing distance is kept constant at 50 meters during the experiment. A "HTC One" phone will be used for the study. Before collecting a landmark, a stable GPS position is awaited. Additionally, the position accuracy needs to be below 15 meters (see Section 3.2.3). The OpenStreetMap data of the desired landmarks is not inspected prior to the test. This ensures that landmarks are also considered even if they are not present in OSM.

After the landmarks are collected, a user study will be conducted with one participant in order to find out whether people unfamiliar with the application achieve similar results. The objective of the subject will be to collect already accumulated landmarks while both rankings are activated. It is therefore important to ensure, that the landmark-tag is not uploaded in advance to the user study as the number of tags is part of the ranking (see Section 3.4). The participant will be instructed to wait for a stable GPS position and to take the pictures from a near distance (approximately 10 meters). The test subject will use the same smartphone as was used in the evaluation before. By complying with these rules the results should be comparable to the near distance recordings with both rankings activated. Possible differences may stem from varying GPS and compass accuracies, different locations while taking the pictures or divergent camera angles. The performance of this study will be restricted to the urban region.

4.2 Results of the Study

This section shows the captured landmarks and statistics about their ranking positions during the tests. The first part presents the captured objects and provides justifications why they were selected as landmarks. Secondly, overall statistics are listed and described to analyse the impact of different ranking settings.

Furthermore, individual examples are analysed to compare results under various external influences. These insights serve as a basis for the discussion, to describe the quality of the different ranking approaches and to determine the value of this work.

4.2.1 Captured Landmarks

Figure 4.1 (Legend shown in Table 4.2) shows a map of the accumulated landmarks in Zurich. In total, 20 landmarks were captured in Zurich. Whereas several of them are obvious landmarks due to their prominence (e.g. No. 4, 7 and 18), some are less apparent and only salient in their respective contexts (e.g. No. 9, 12 and 19). Several other candidates are present within that region. However, the objective of this study was not to cover all landmarks in the map extent, but to find examples of salient features to cover the criteria mentioned in Section 4.1.

Figure 4.2 (Legend in Table 4.3) shows the collected landmarks in the rural region Zumikon. In this area, 10 landmarks were accumulated in total. Most of them are located in the village centre as there were large numbers of missed landmarks outside the village core. Hence, a cluster of landmarks is visible in the centre of the village.

The two map extracts of Figures 4.1 and 4.2 provide an impression of the differences regarding the data density in both regions. In quantitative terms, the map extent in Zumikon has 24 objects which meet the defined minimum requirements of a landmark candidate (see Section 3.3.2) on an area of 0.224 km^2. In contrast, Zurich presents 838 of such objects on an area of 0.351 km^2 (Status: 03.08.2014). Thus, the chosen landmarks in Zumikon were strongly influenced by the available data.

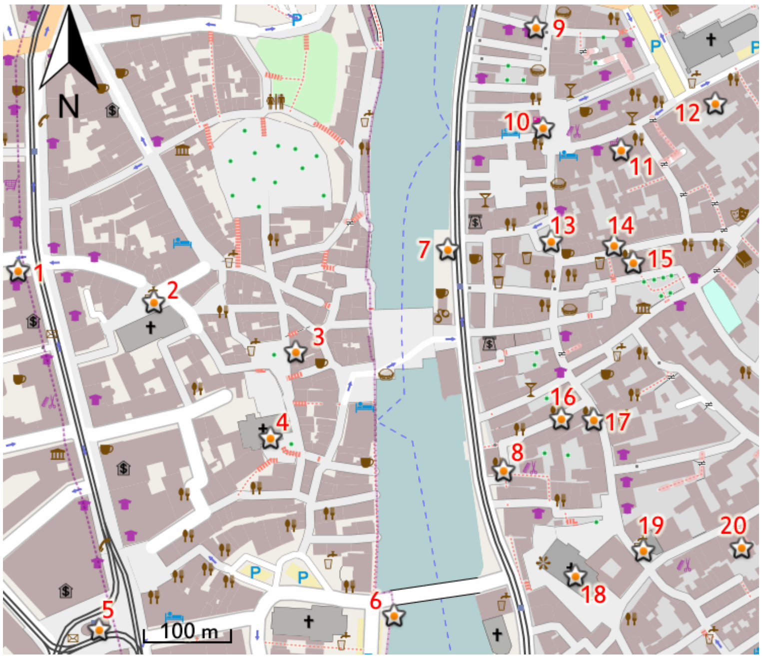


Figure 4.1: Captured landmarks in Zurich [Map source: OSM 2014]

Nr.	**Type in OSM**	**Name of the Landmark**	**Reason to mark it as a landmark**
1	Artwork	Pavillon-Skulptur	Remarkable sculpture at prominent street
2	Drinking Water	-	Richly decorated fountain
3	Book Shop	Beer Buchhandlung	Eye-catching shop-sign
4	Place of Worship	Kirche St. Peter	Famous old church
5	Tram Stop	Paradeplatz	Famous place, centrally located
6	Tourist-Attraction	Hans Waldmann	Prominent statue, tourist-destination
7	Tourist-Attraction	Hafenkran	Widely debated and prominent attraction
8	Restaurant	Haus zum Rüden	Very old prominent restaurant
9	Telephone	-	One of few remaining phone booths in Zurich City
10	Pharmacy	Rosen-Apotheke	Only pharmacy in its surrounding area
11	Convenience-Store	Läbis 1	Well-established local store
12	Building	Predigergasse 18	Only blue house in street
13	Fountain	-	Prominently positioned fountain with sculpture
14	Building	Haus zur Sichel	Golden sickle in front of building
15	Restaurant	Öpfelchammer	Locally famous restaurant with eye-catching sign
16	Restaurant	Kaffee Schoffel	Meeting place, eye-catching showcase
17	Restaurant	Bodega	Well identifiable, famous restaurant
18	Place of Worship	Grossmünster	Most prominent landmark in Zurich
19	Drinking Water	-	Large fountain, easily detectable
20	Building	Haus zum Paradies	Very eye-catching oriel window

Table 4.2: Legend for the captured landmarks in Zurich

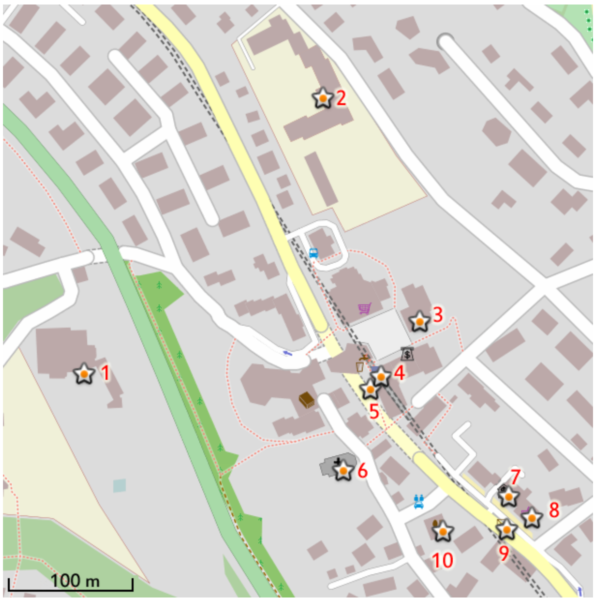


Figure 4.2: Captured landmarks in Zumikon [Map source: OSM 2014]

Nr.	**Type in OSM**	**Name of the Landmark**	**Reason to mark it as a landmark**
1	Building	Hallenbad Zumikon	Only indoor swimming pool in Zumikon
2	School	Schulhaus Farlifang	Locally known school
3	Townhall	Gemeindehaus	Prominent building in the centre of the village
4	Station	Bahnhof Zumikon	Main station of Zumikon
5	Playground	Kleiner Spielplatz	Well visible playground elements
6	Place of Worship	Kirche Zumikon	Only church in Zumikon, prominent landmark
7	Bank	Zürcher Kantonalbank	Clearly visible bank, relatively large building
8	Convenience-Store	Volg	Well-known local store
9	Post Box	-	Well visible post box
10	Restaurant	Gasthof Rössli	Locally known restaurant

Table 4.3: Legend for the captured landmarks in Zumikon

Figure 4.3 shows statistics about the tags from all 30 collected landmarks. This provides insights about the influence of the selection process and the frequency of certain tags occurring on landmark candidates. Such information may help to refine the semantic ranking in a future version of the application. "Name" was the most common tag, followed by the category information "amenity". Other frequent tags were related to the address, indicated a website or provided information about the accessibility of the objects with the key "wheelchair".

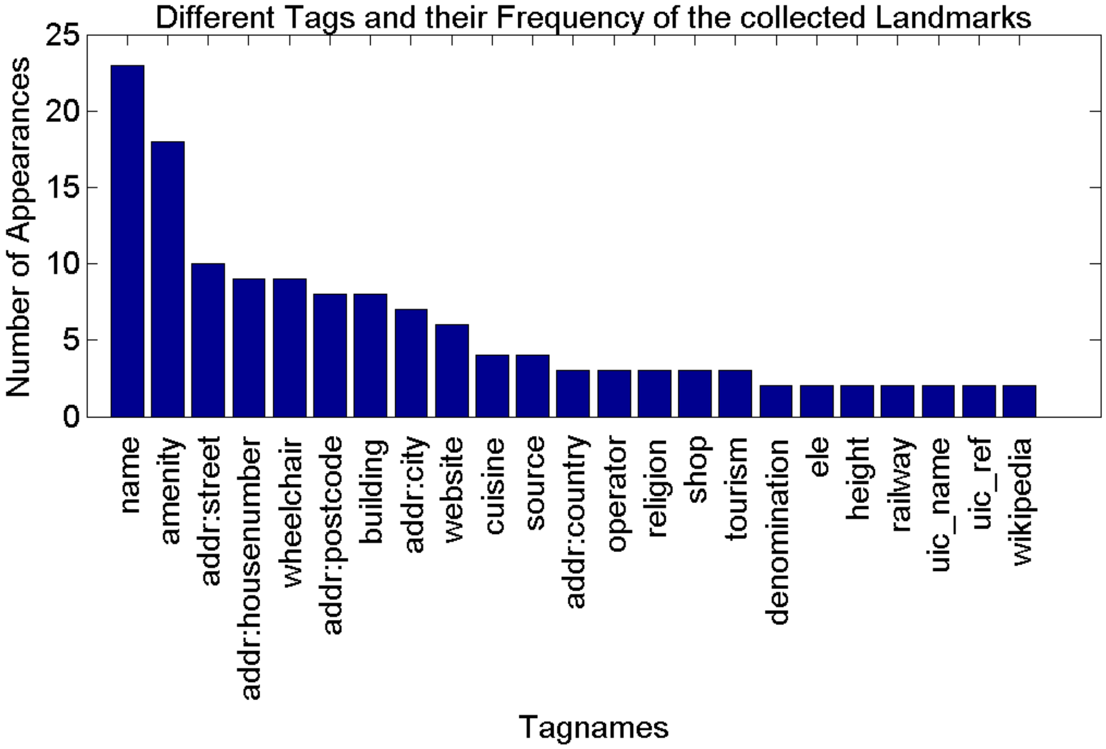


Figure 4.3: The frequency of all tags from the collected landmarks with more than one appearance [Data source: Swiss Taginfo, 2014b]

4.2.2 Overall Results

This section presents the overall results. The statistical evaluation is confined to descriptive statistics as no statements can be obtained about the total population of landmarks and no hypothesis was defined in advance. The raw data of all recordings can be found in Appendix B.

Table 4.4 lists the overall results of the study. First, the numbers of missed and found landmarks are listed. Based on these values the success rate in finding the desired landmark is determined. This rate is significantly higher in Zurich with a ratio of 87% (20 found, 3 missed) against 45 % (10 found, 12 missed) in Zumikon. Missed landmarks generally either have no representation in OpenStreetmap, cannot be found by the application or offer no possibility to derive a name out of the metadata. However, a subsequent clarification indicated, that the missed landmarks were caused exclusively by non-existing OSM data. Missed landmarks are therefore not included in the ranking results, as they do not explain the performance of the ranking system.

The second part of the table shows the percentages, indicating how many hits were on a particular ranking position. The first place from a near position in Zurich for example shows the values 62.5% with both rankings activated, 35.0% for the semantic ranking and 37.5% for the visual approach. Therefore, a reading example is as follows:

> *62.5 % of all indicated landmarks in Zurich were ranked on the first position while both rankings were activated and the distance to the landmark was approximately 5 to 15 meters.*

Furthermore, the statistical measurements arithmetic mean, mode and the average measurement deviation are listed in the table. The average measurement deviation specifies the average deviation between the two pictures, which were captured with the exact same settings and location. Thus, this provides a criterion for a comparative assessment of the robustness of the different approaches.

Statistic	**Zurich**			**Zumikon**		
Found Landmarks	20			10		
Missed Landmarks	3			12		
Sucess Rate (%)	87			45		
Position Near (5-15m)	Both R.	Semantic R.	Visual R.	Both R.	Semantic R.	Visual R.
1. Place (%)	62.5	35.0	37.5	90.0	70.0	75.0
2. Place (%)	17.5	12.5	27.5	10.0	10.0	25.0
3. Place (%)	17.5	22.5	15.0	0.0	20.0	0.0
4. Place (%)	2.5	0.0	10.0	0.0	0.0	0.0
5. Place (%)	0.0	7.5	5.0	0.0	0.0	0.0
≥6. Place (%)	0.0	22.5	5.0	0.0	0.0	0.0
Average Position	1.6	3.625	2.525	1.1	1.5	1.25
Average Measurement Deviation	0.6	0.25	1.15	0.2	0	0.1
Mode	1	1	1	1	1	1
Position Far (15-35m)	Both R.	Semantic R.	Visual R.	Both R.	Semantic R.	Visual R.
1. Place (%)	47.5	40.0	25.0	80.0	55.0	60.0
2. Place (%)	12.5	7.5	30.0	15.0	25.0	15.0
3. Place (%)	17.5	12.5	7.5	5.0	10.0	0.0
4. Place (%)	7.5	12.5	15.0	0.0	10.0	5.0
5. Place (%)	7.5	2.5	0.0	0.0	0.0	20.0
≥6. Place (%)	7.5	25.0	22.5	0.0	0.0	0.0
Average Position	2.5	3.65	3.7	1.25	1.75	2.1
Average Measurement Deviation	0.6	0.2	1.1	0.1	0.1	0.4
Mode	1	1	2	1	1	1

Table 4.4: Overall results of the study

The average positions in Table 4.4 suggest that the best detection was achieved when both rankings were activated with a mean ranking position of 1.6 (Zurich) or 1.1 (Zumikon) respectively from a distance of 5 to 15 meters. By increasing the distance to the landmark to 15 - 35 meters, this value has deteriorated to 2.5 / 1.25. The semantic ranking provided an average value of 3.625 / 1.5 in near distance and 3.65 / 1.75 from a farther distance. The decline caused by the increased distance is therefore considerably smaller while using this ranking. The visual ranking had the highest decline in the ranking position between near and far with an average of 2.525 / 1.25 for close positions and 3.7 / 2.1 while being further away.

In Zurich, nearly two-thirds (62.5%) of all near distance attempts using both rankings were placed on the first position, whereas the separate rankings both had about one third "direct hits" (35% and 37.5% respectively). In Zumikon, this difference is much smaller (90 % against 70 % / 75 %), as the rankings all have relatively high values. While using both rankings further away, nearly half (47.5%) of all captured landmarks in Zurich landed on the top position.

While both rankings were activated, only few attempts were insufficient i.e., below the fifth place, and those occurred only from a far distance. The semantic ranking had around a quarter of such outliers in Zurich from both distances, the visual ranking 5% and 22.5% respectively. In Zumikon, no intended landmarks were placed below the 5th position, regardless of the ranking. However, in this area the number of candidates in the FOV rarely exceeded five.

The average measurement deviation gives an indication about robustness of the results. In Zurich, the semantic ranking has the highest stability through lesser dependence on exact sensor data, whereas the visual ranking had the highest instability in the position. The deviation with both rankings is in an intermediate position between the two separate approaches. Zumikon shows a similar pattern, although the difference between the different deviations is much smaller. Every ranking / distance variation places the mode on the first position with one exception indicating a central tendency towards the first place.

Subsequently, two box plots (Figures 4.4 and 4.5) are used to illustrate the results. These are included to improve the understanding of the data and to show possible trends. These boxplots present a similar pattern: In Zurich, both rankings combined from a near position returned the best results. The performance of both rankings from a far position is comparable to the visual ranking from a close distance. Outliers were quite rare and occurred primarily in the visual ranking from a far position. The biggest range between the upper and lower quartile was found in the semantic ranking (4 respectively 4.5 positions). In Zumikon, the range in the semantic ranking is considerably smaller (1 position). In this region, the combined ranking also showed the best performance. Despite the visual ranking, the positions are scarcely affected by the distance differences. The visual ranking from far distance had the worst overall performance. Yet, all obtained results are in the "sufficient" region (between 1-5).

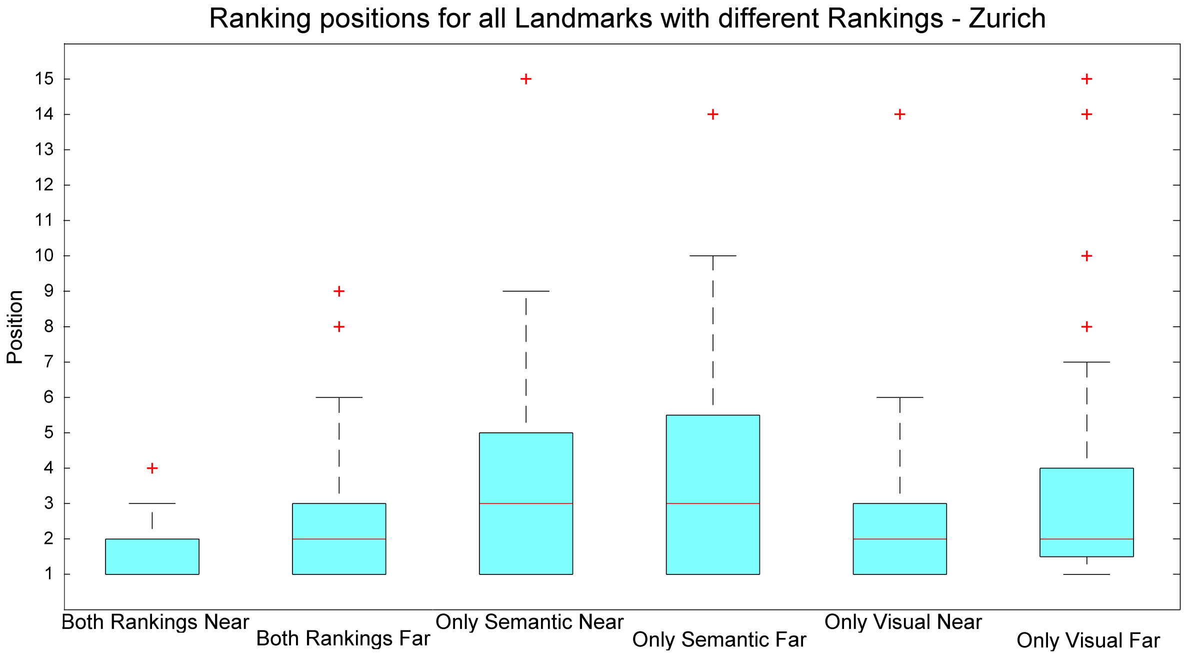


Figure 4.4: Boxplot of all results of the various rankings for Zurich

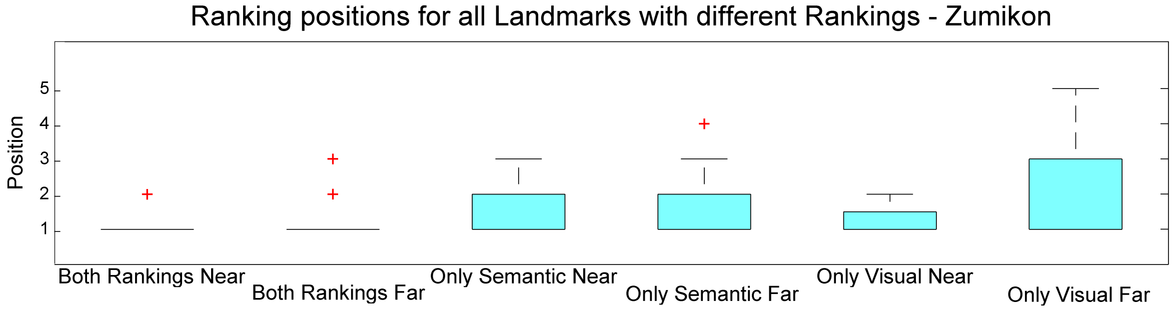
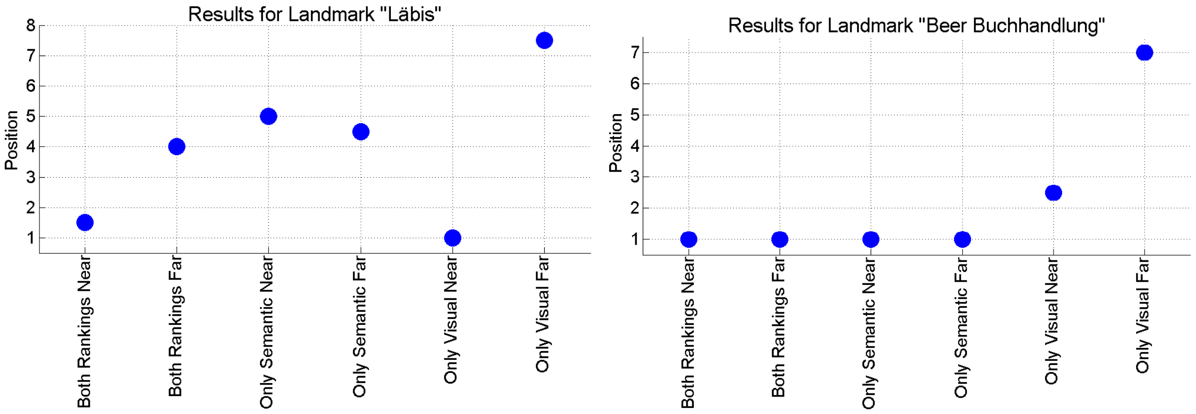


Figure 4.5: Boxplot of all results of the various rankings for Zumikon

4.2.3 Individual Results

In this section, individual results are presented. This provides an impression of how various circumstances influence the outcome. Three particular differences are highlighted: Difference in the surroundings, GPS accuracy and the popularity of the landmark. A list of all individual results can be found in Appendix B.

Difference in Surroundings:
Figure 4.6 shows the results for two shops, one bookshop ("Beer Buchhandlung") and one convenience-store ("Läbis"). Since both are categorised as shops, they receive the same category value (5). Additionally, the visual ranking and the amount of potential



(a) Landmark "Läbis" (b) Landmark "Beer Buchhandlung"

Figure 4.6: Example results for differences in POI density and metadata

landmarks in the field of view (Läbis: 28 candidates, Beer Buchhandlung: 33 candidates) are comparable. Yet, with both rankings activated, the bookshop was ranked on the first position for the near and far position, whereas the convenience-store had an average ranking of 1.5 or 4 respectively.

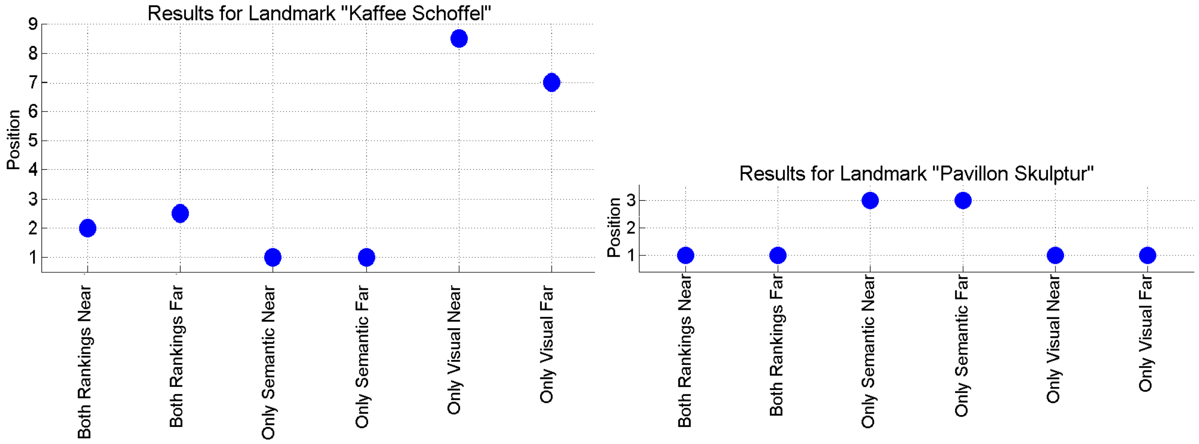
Läbis	
Key	Value
name	läbis 1
shop	convenience

Buchhandlung Beer	
Key	Value
addr:housenumber	10
addr:street	St.Peterhofstatt
name	Buchhandlung Beer
shop	books
website	http://www.buch-beer.ch

Table 4.5: OSM metadata of two landmarks with differences in their surroundings [OSM, 2014]

Regarding the metadata (Table 4.5), two differences are apparent: The listed website of the bookshop and the number of tags (2 vs. 5). Both parameters are integrated in the ranking, which means that the bookshop receives a higher semantic ranking. However, the main reason for the noticeable difference is the surrounding area where the landmarks are situated. The application is designed to provide objects with category information a higher priority. From the 28 candidates in the field of view of "Läbis", 5 are points of interest according to the definition of this study (see Section 2.3.4), thereby providing guaranteed category information. Other included objects are buildings. Moreover, these 5 objects are very close to each other. Inside the FOV of the bookshop there are only 2 POIs, additionally located further apart from each other. These circumstances ease the process of prioritizing the bookstore for the application.

Difference in GPS Accuracy:



(a) Landmark ”Kaffee Schoffel”

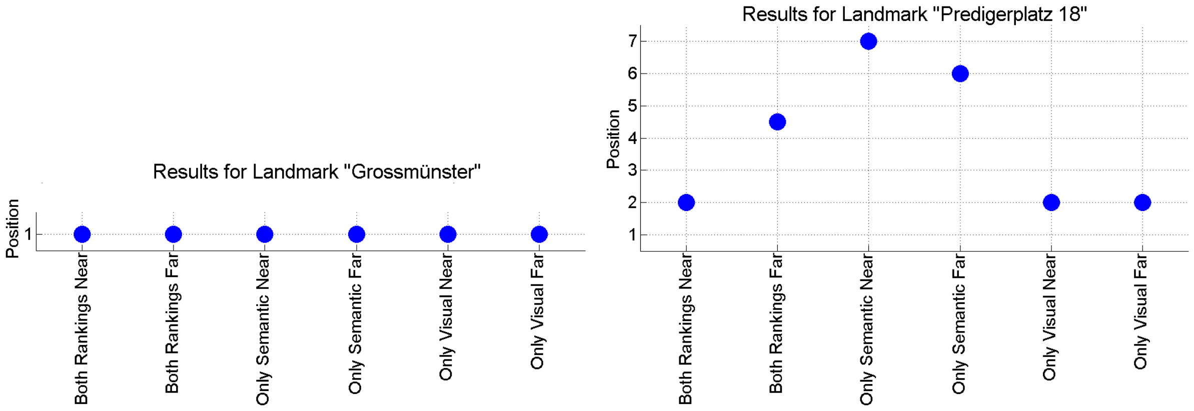
(b) Landmark ”Pavillon Skultpur”

Figure 4.7: Example results for differences in the GPS accuracy

The comparison from Figure 4.7 shows effects of different GPS accuracies. The ”Kaffee Schoffel” (a coffee shop) was captured with a poor GPS signal, while the accuracy in the case of the ”Pavillon Skulptur” (a sculpture) was almost ideal. The coffee shop is located on a very narrow street and surrounded by a lot of tall buildings, hence the GPS signal is highly limited. This has a strong effect on the visual ranking: Although the photo was taken directly in front of the landmark, the visual ranking provides a very low value for the near distance with an average rating of 8.5. With a value of 7, the far distance is also low ranked. Solely because of the high semantic ranking (position 1), the coffee shop manages to be at the 2nd / 2.5th position with both rankings activated. The opposite is the case for the sculpture: The street is wide enough to enable a precise GPS signal. From both distances, the visual ranking suggests the intended landmark. This also leads to the first position using the combined approach despite of the 3rd place in the semantic ranking.

Difference in the Popularity of the Landmarks:

Figure 4.8 shows differences in the detection of popular and non-popular landmarks. The landmark ”Grossmünster” with its distinctive two towers is the emblem of the city of Zurich, therefore it is one of the most popular landmarks in the nearby area. In combination with a low data density and a good GPS signal, this landmark is ranked on the first position in every tested variation. Regarding the metadata (Table 4.6), the object is easily detectable as a landmark through its tags since ”place_of_worship” is one of the highest ranked categories in the application. Additionally, there is also a listed Wikipedia article and a website. This is in sharp contrast to the other example ”Predigergasse 18” (a building), which was only chosen as a landmark due to its unique color. The listed tags do not allow any conclusions to select this object as a landmark,



(a) Landmark "Grossmünster" (b) Landmark "Predigerplatz 18"

Figure 4.8: Example results for differences in the popularity of the landmark

which makes it difficult for the semantic ranking to detect it as such. Thus, the only possibility for the application to rank the building on a top place is a clear accordance with the visual ranking. As a consequence, the combined approach shows this landmark only on the 2nd / 4.5th position.

Grossmünster	
Key	Value
amenity	place_of_worship
building	yes
denomination	protestant
height	62
name	Grossmünster
religion	christian
website	www.grossmuenster.ch
wheelchair	limited
wikipedia	de:Grossmünster

Predigerplatz 18	
Key	Value
addr:city	Zürich
addr:country	CH
addr:housenumber	18
addr:postcode	8001
addr:street	Predigerplatz
building	yes
source	Stadt Zürich Open Government Data; Geomatik + Vermessung; Tiefbau- und Entsorgungsdepartement

Table 4.6: OSM metadata of two landmarks with different popularity [OSM, 2014]

4.3 Results of the User Study

This section presents the results of the user study. As the participant was instructed to take the pictures in a distance of about 10 meters to the landmark with both rankings permanently activated, the results are comparable to the near distance results. The compliance with the prescribed distance was evaluated by looking at the captured photos. Compared to the main study, differences regarding the recording position and the camera angles have been observed.

Statistic	Both Rankings Near	User Study
1. Place (%)	62.5	65.0
2. Place (%)	17.5	15.0
3. Place (%)	17.5	15.0
4. Place (%)	2.5	2.5
5. Place (%)	0.0	0.0
≥6. Place (%)	0.0	2.5
Average Position	1.6	1.675
Average Measurement Deviation	0.6	0.65
Mode	1	1

Table 4.7: Results of "Both Rankings Near" compared to the user study

Table 4.7 shows the obtained results. With the exception of one outlier (Läbis 1) in the user study, the overall results are very similar. The average measurement deviation of the user study is influenced by one outlier, therefore slightly higher. The mode remains unchanged. The distribution of the ranking positions is practically identical. The same pattern is evident in the boxplot shown in Figure 4.9. While looking at the individual results (Appendix B), it can be seen that the specific positions differ slightly from each other.

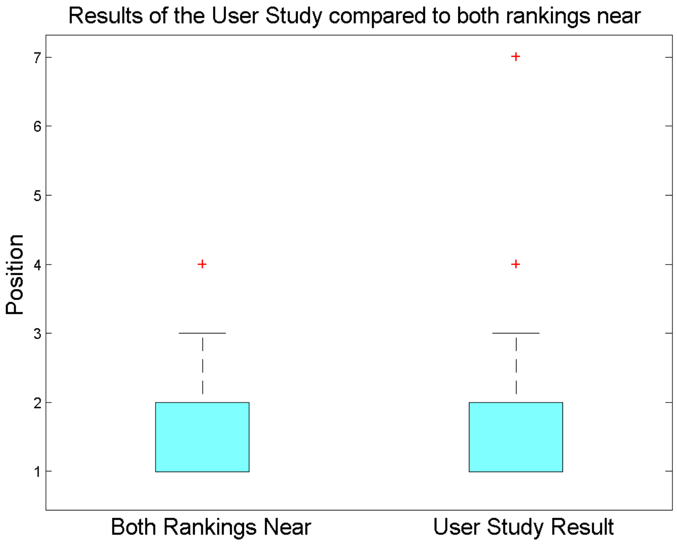


Figure 4.9: Boxplot of the results of "Both Rankings Near" compared to the user study

5 Discussion

In this section, the developed application and the results of its evaluation are discussed. The obtained results allow conclusions about the value of the introduced ranking system. Furthermore, benefits and shortcomings of the application as well as problems that emerged during its developmental process are discussed. Finally, the research questions from Section 1.2 are answered.

5.1 Investigation of the Study Results

During the evaluation of the study, 30 landmarks were collected using the application. 20 thereof were gathered in the city centre of Zurich, representing an urban area with high data coverage. 10 landmarks were obtained in Zumikon to examine the application's performance in a rural area. As stated in Section 4.1, landmarks were chosen to meet certain criteria and not to provide an overall coverage of the study area. These criteria included the consideration of prominent and less prominent landmarks under varying circumstances regarding for example the POI density or the quality of the GPS signal. Due to this selective approach, no conclusions can be drawn about the application's performance in handling the overall population of landmarks. Comparisons between the success rates of the different rankings and the numbers of found and missed landmarks are however meaningful.

A key finding of the study is that the usability of the developed software depends to a large extent on the density and detailedness of the provided data. Low information density led to several missed landmarks in the case of Zumikon. There, the amount of missed landmarks even exceeded the number of found landmarks. As stated in Section 2.3.2, Haklay [2010] brought up the idea of a critical value in terms of the coverage, where the provided OSM data is no longer useful for cartographic analysis. It is therefore assumed, that in the case of Zumikon, the data density is not sufficient to provide a fundamental basis of landmark candidates. Discussions regarding the overall performance of the application are therefore limited to the results obtained in Zurich.

In the town of Zurich, the amount of missed landmarks was negligibly small. The challenge here was to find the intended landmark despite sensor inaccuracies and large numbers of objects. By combining semantic and visual characteristics, the application was able to detect the investigated landmark in 62.5 % (near) / 47.5 % (far) of all cases as the most probable object and in 100 % (near) / 85 % (far) as one of the 4 most plausible candidates. With respect to the identified inaccuracies and constraints of smartphone sensors, the achieved success rate is considerably high, notably compared to the tested separate ranking approaches. As the desired landmarks were detected in almost all cases, this rate should be sufficient to provide a basis for practical purposes.

While maintaining the same location, the visual ranking showed comparatively strong fluctuations. The reasons therefore were the unstable GPS positioning, imprecise azimuth data and slightly different camera angles. On the other hand, the semantic ranking was very stable as the objects in the field of view usually stayed the same. The study revealed that the approaches with separate rankings could not compete with a combined method. The amalgamation of both rankings, visual and semantic, led to a distinct improvement of the results, thereby identifying the weaknesses of approaches limited to only one characteristic. The necessity of including different rankings is mainly caused by sensor inaccuracies. Yet, even with perfect GPS- and compass data, a semantic ranking must be included to measure the "landmarkness" of the intended objects. Otherwise, there would be no assessment of the suitability whether the desired object provides a suitable landmark.

A distance of 10 - 15 meters seemed to be the best position to capture a landmark in respect to the sensors. Standing closer to the object could cause the landmark to appear behind the provided position. By involving the GPS inaccuracy in the FOV, the application still managed to include the landmark in these situations. Nevertheless, the calculated azimuth ranking was particularly low in these cases, as the object appeared to be in the opposite direction. With larger distances the compass data was more reliable since the influence of positioning errors decreased. Yet, too large distances to the landmark led to a prioritization of nearer objects. In the majority of cases, a distance of 10 - 15 meters also was the most "natural" way of taking a photo. The adequate distance to the landmark depended, however, also on the size of the object.

The individual results (see Section 4.2.3) revealed three challenging situations: Low GPS accuracies, high data density and lacking popularity of the landmark. The GPS accuracy had a strong influence on the visual ranking, thus making it difficult to find the intended landmark. High data density was challenging for both rankings as more objects had to be considered as possible candidates. Low popularity influenced the semantic ranking as the metadata did not point out the saliency of the object. As a consequence of these constellations, popular landmarks were identifiable without any further problems. Less prominent landmarks had to rely mainly on accurate visual data.

The user study confirmed the previously gathered results and showed that the application can be handled by persons unfamiliar with its functioning. The achieved success rates were very similar compared to the rates before. Yet, individual differences in the case of certain landmarks indicated a certain degree of randomness of the results. The user study was limited to one test subject, thus not allowing to establish statements of general validity.

5.2 Discussion of the Ranking System

This thesis introduced a ranking system based on a formula developed by Raubal and Winter [2002] for the purpose of calculating an object's salience (See Section 3.4). The ranking is composed of visual and semantic characteristics. On the one hand, the intention of this system is to distinguish between essential and unimportant objects in the FOV and on the other hand to compensate for possible inaccuracies of the sensors. As discussed in Section 5.1 the results of the strictly visual approach imply that one is not able to achieve satisfactory results over longer periods and altering conditions by relying mainly on sensor data. Hence, the purpose of integrating semantic factors is to provide stability in the outcomes and to measure the "landmarkness" of the examined objects. An overestimation of semantic properties, however, may hamper chances of finding the intended landmark. By assigning appropriate individual weights to the identified characteristic factors (see Table 3.3) the different rankings were balanced as well as possible.

The values of the weighting factors and categories are only based on tests, personal experience and assumptions and are therefore not scientifically justified. All these weights have constant values, thus implying that they are valid under a wide range of operating conditions. Nevertheless, the importance of the individual factors may differ depending on the context: "Size" for example is a very low rated factor as it did not seem to be very influential in the tested areas to determine the intended object. Yet, for instance in regions with major differences regarding the building size, "Size" may stand out as a salient property.

There are attributes in the ranking where a connection with "landmarkness" is not beyond doubt: A missing website for example does not mean that there is none available. Additionally, the number of tags do not necessarily coincide with an object's saliency. However, since the whole application is exclusively based on OSM data, it is necessary to consider as much available information as possible. Nevertheless, the correlation between certain factors and an object being a landmark has not been broadly tested. Bearing this in mind is important, since the semantic ranking is theoretically capable of preventing correct results. This might occur when an intended landmark has a lower semantic rating compared to its surroundings and if at the same time the visual factors are underestimated or inaccurate. Yet, by increasing the number of suggested candidates to 5, only three missed attempts occurred (Landmark not in the top 5) from a total of 80 while both rankings were used (See Table 4.4 / Appendix B).

One critical aspect is the fact that during the study the application has only been tested with a single smartphone model (HTC One). Other devices may have differing "optimal" weighting values depending on the accuracy of the sensors. Additionally, newer generation of mobile phones will further improve the sensors. This may reduce the importance of the semantic ranking.

Despite its associated challenges and critical aspects, the combined ranking has significantly improved the results. As a consequence, this may provide a promising way to develop common standards. Nevertheless, to optimize the results, the weighting factors should be adjusted to device- and situational contexts. Such adjustments could include for example an increase of the visual weighting value depending on the estimated GPS reliability. Therefore, an adaption to situational dependent characteristics such as the POI density or the sensor accuracy is proposed.

5.3 Further Benefits and Shortcomings of the Selected Approach

Through the development of this application, the present thesis has shown that a manual collection of landmarks on smartphones is not only a theoretical possibility, but a feasible approach. Requirements are a stable mobile internet connection, preferably UMTS, an integrated camera, a GPS receiver with an active signal and a built-in compass. Depending on the connection speed and the amount of downloaded data, the application manages to present the results within a few seconds. The observed success rate of finding the intended landmark allows a widespread use, thus enabling a user-generated collection of landmarks.

In order to collect a landmark, the user is required to use the camera of the phone. Hence, a database of pictures with their associated landmarks could be established. In contrast to popular photosharing platforms such as Flickr, the pictures are taken and uploaded with the motivation of sharing landmark information and thereby enhancing the quality of navigational systems. Photographs on Flickr are made for a number of different reasons for example touristic purposes, art etc. This hinders a widespread usage in direction giving [Bell et al., 2009]. A photo database only based on pictures with navigational background would theoretically allow a seamless integration in routing instructions. Nevertheless, for such an implementation, the rules for uploaded pictures must be clearly communicated. Additionally, several landmarks would require multiple pictures captured from various angles if persons are approaching from different directions.

This thesis introduced several strategies which contributed to enhance the detection of the intended landmark (see Section 3.2.3). The visualization of location-related data on a map before and after collecting a landmark allowed an examination of the sensor's performance. Depending on the user's level of map reading and available reference points in the environment, the user is able to check whether the accuracy is sufficient to accomplish the current task. The minimum accuracy for the GPS provided an effective method to prevent the user to collect landmarks with insufficient positioning accuracy. In regard to the obtained results, increasing the number of suggested landmarks to five nearly doubled the chance of including the landmark in the results. The integration of the mean compass error helped to reduce the inaccuracy provided by this sensor. The location sampling was introduced to provide more stable results, yet has not been sufficiently tested to confirm its additional value.

Previous automated methods were hampered, amongst others reasons, by the requested amount of data to generate landmarks [Richter, 2013]. A main advantage of a user-generated approach is that the main part of the selection process is done by the user with the manual confirmation of the result. The included algorithms only suggest landmarks and do not define them unlike automated methods. This enables a substantial reduction of the data requirements.

The only data source for this project is OpenStreetMap. The application is therefore strongly reliant on the quality, currentness and completeness of OSM data. The inclusion of other sources, for example search engine results from services such as Flickr or Google, could lower this dependency and may help to improve the obtained results. The integration of such sources, however, might cause some difficulties: The enumerated search engines rely on a viable naming of the objects. For a large part of all candidates the derived name is limited to their category or is ambiguously formulated, thus does not allow to define a clearly distinguishable search query. Furthermore, it is uncertain whether these data sources would actively improve the results or only confirm the current ranking.

One further problem arising from the restriction to OSM data was missing objects. Due to imprecise finger input and possible duplicates, the application only allows the conversion of existing objects to landmarks instead of creating new ones. This was especially apparent in the rural area, where the amount of available landmark candidates decreased drastically from 838 to 24 on a comparable size of the area (see Section 4.2.1). Part of a solution would be, to allow users to create objects directly on-site if they are not present. Using this method in areas with low data coverage such as in Zumikon, one would be able to overcome the problem of non-existing objects in OSM. Nevertheless, this would require an exact placing on mobile screens. Particularly in regard to more complex geometries, this could turn out rather difficult.

The created application limits landmark-candidates to objects with a given name, a known category or buildings with an address. Filtered objects include line-shaped objects for example streets, rivers, polygons without a "building"-tag representing objects such as lakes or hillsides and point entities where the application wasn't able to derive a name. On the one hand, the omission of point entities happened on purpose, for example in the case of single trees (see Section 3.3.1). On the other hand, point objects without a "name"-tag and with an unknown category are also automatically omitted. OSM data includes a vast diversity of different key/value pairs. The naming method needs to include as many keys as possible and therefore must be continuously enhanced in order to avoid losing candidates. During the evaluation of the application no such case was encountered. Yet, this must be kept in mind if providing the application to a larger user base. The filtering of polygons without a "building"-tag constituted no problem in the selected sites, since a large part of them represents natural objects. In some environments, however, this omission could lead to a rather low density of landmarks due to a lack of man-made objects. Further enhancements could include such objects.

An unused potential lies in the content of the image. The application renounces completely from any content-based image retrieval approaches (CBIR). CBIR could theoretically be used to detect objects in the image. For example, if a fountain is detected as the main object of the picture, the application could automatically limit the results to matching hits. However, CBIR is not yet developed enough to detect image content from such a high variety of objects [Purves et al., 2008]. A limited integration could nevertheless provide added value. Another unused possibility is the estimation of visibility with a line of sight analysis. This was refrained as there is no existing height information in OSM data and the inaccuracy from the GPS would probably falsify the results. As a trade-off, the user is allowed to change the viewing distance at all times.

In the current state, only the fact, that the object is a landmark is stored. It is not possible for the user neither to justify the choice nor to describe any special characteristics of the landmark. Nevertheless, objects can be a landmark due to many reasons, for example their color, their age etc. The salient property of a landmark often does not include the whole object itself but instead an eye-catching feature thereof, such as a shop sign. The storing of these characteristics would allow the computation of more complex expressions when referring to an object such as "the blue building on the corner with the striking shop window". If one wants to save all associated information (landmark-tag, justification, characteristics) on OSM, this would significantly increase the amount of stored data and thus the number of needed tags. This amount would become even larger if photo-related information should be saved, for example, the time of the recording, the azimuth angle to the landmark and the location where the picture was taken. This might be disturbing to parts of the users who are not interested in landmark information. The pictures in the application are only locally stored. This has to do with the fact that this project aims to store all its findings entirely on OSM. OSM offers no possibility to save any images. As an alternative solution, it would make sense, to upload all gathered results on an own publicly available server to offer a platform for the sharing of all landmark related information.

The greatest challenge of user-generated approaches, however, is to find local experts willing to contribute to a project [Richter and Winter, 2011]. In order to find enough users, the people must be informed about the added values of landmarks and their possible usage. In addition, user-generated content does not guarantee any usefulness of the provided information [Richter, 2013]. It is unclear, whether existing reputation mechanisms of OSM are sufficient to overcome this shortcoming.

5.4 Encountered Problems During the Process

Next to the already mentioned problems related to the GPS accuracy and missing OSM objects, several challenging situations and difficulties were encountered during the development and the testing of the application:

Further GPS Related Problems:

As stated in Section 3.2.1, the GPS needs some time before the provided location is accurate enough for an appropriate usage in this application (GPS latency). Capturing the intended landmark may be impossible during this period as there is a strong possibility of not including the landmark in the FOV due to lacking accuracy. Every position update provides an associated accuracy value given in meters describing the radius in which the real position is assumed to be with a certainty of 68% [Google, 2014b] (see Section 3.2.1). To prevent the user from taking pictures before the location is fixed, a minimum GPS accuracy radius of 15 meters is enforced. Yet, during the tests, the given accuracy radius was not always provided in a reliable manner, especially during the time before a fixed location was reached. As a consequence, it is possible to collect landmarks with a real inaccuracy significantly higher than 15 meters. Thus, more sophisticated approaches are needed in order to decide whether the provided accuracy is adequate. A promising approach would be for example, to lock the camera if there was a certain change in the GPS position during the last few seconds to ensure that the location is fixed.

Another problem regarding the GPS latency is that long waiting times could tax the user's patience. During the testing, the time until the first position fix was between 15 and 30 seconds, depending on the satellite "visibility". Such a recurrent waiting period may cause frustration for the users.

Areas as Points:

All map objects, regardless of their properties, have a certain area extent in the real world. In OSM, it may occur that even objects with a considerable large area are stored as points. To capture these objects, the point must be included in the FOV. Depending on the object's size, this point can be missed when photographing the building. Figure 5.1 shows such an example for the Bahnhof Zumikon. The signed entrance to the underground station lies in a distance of around 20 meters to its OSM point representation. If the user takes a picture right in front of the station entry, the FOV may not include the OSM position of the landmark. In the study results this situation led to an unexpected low rate of the visual ranking (See Appendix B - Bahnhof Zumikon). Nevertheless, in other cases the application may not be able find the intended landmark at all.

Different Tags for Same Purposes:

Every registered user can edit OSM data and add own tags. Besides its benefits, this philosophy leads to an inconsistency of data: Several users created different key/value combinations for the exact same purpose. Train stations in Switzerland for example, are represented as "public_transport=station", "railway=station" or "station=yes" [Swiss Taginfo, 2014c]. The amount of different combinations is uncountable and it is nearly impossible to cover all existing combinations. This may hinder a more widespread distribution of the application in its current state, as only common key/value combinations in Switzerland were considered.

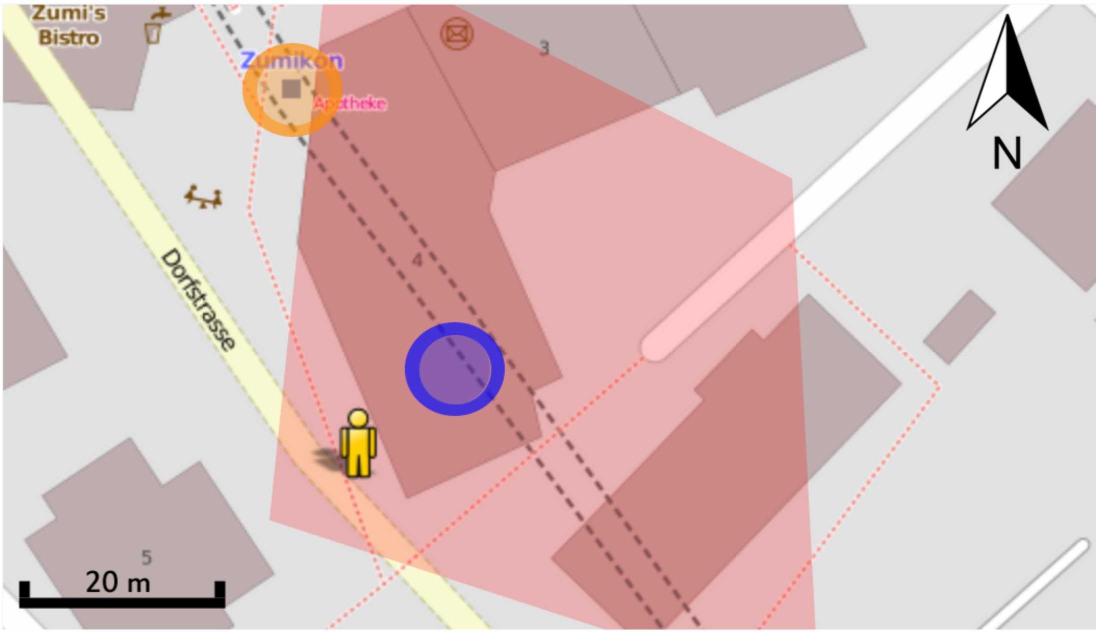


Figure 5.1: The main station in Zumikon. An example for a large construction stored as point data in OSM making it difficult to recognize the object for the application - Orange circle: The location of the station in OSM; Blue circle: The entrance of the station; Red: The field of view [Map source: OSM, 2014]

Inclusion of Different Geometric Types:
The mixture of point and polygon data led to several challenges regarding the ranking. In quantifying the visual characteristics of point objects, the application exclusively relies on sensor related parameters, whereas polygons also integrate size and visible range of the candidates. In the situation, where a polygon includes a point object, for example a building containing a shop, it is generally estimated that the polygon (the building) is closer to the user (See Section 3.4.3). Through its smaller distance, the polygon receives a higher distance ranking thus being preferred in the visual ranking. The lower weighting value of the distance for ways (see Table 3.3) attempts to compensate this problem. Nevertheless, depending on the size of the building and the position of the user, this can lead to unintended ranking positions.

Currentness of Data:
During the tests, some outdated data was encountered (not during the study). If such an object is used as a landmark in a routing instruction, this could do more harm than good by confusing the user. Therefore, landmark data must be kept up to date if used in navigational systems.

5.5 Addressing the Research Questions

Hereafter, the research questions stated in Section 1.2 are answered, summarising the value and contribution of the present thesis.

Research Question 1

> *Is a camera-based interface an appropriate way to collect landmarks with the help of smartphone sensors?*

The ubiquitous presence of smartphones led to a widespread availability of sensors yielding the user's location and viewing direction. Thus, the described application provides an easily accessible method for a user-generated collection of landmarks. The use of the built-in camera ensures the focusing of the desired object. In this way, the sensor data can be extracted as close as possible to the moment of taking the picture. Additionally, this approach enables a photo-based collection of landmarks and thus a further use of these pictures in routing instructions.

The present thesis has shown that the technical background for a user-generated landmark collection is given. By using OSM to provide the underlying data, the application reaches relatively high rates in finding the intended landmarks. Yet, the thesis has mainly focused on the technical feasibility of this approach. Therefore, questions concerning the viability in combination with a larger user base cannot be answered finally without providing the application to a wider audience. Furthermore, it remains an open question, whether users would accept this approach and how they would respond to its shortcomings such as missed landmarks and GPS latency time.

Research Question 2

> *Are smartphone sensors accurate enough to identify the desired landmarks?*

The success rate of the application was carried out on the basis of an evaluation outdoors. The obtained values demonstrated a satisfactory result in finding the intended landmarks, thus indicating an acceptable accuracy level of the sensor technology. However, the limitation to a GPS accuracy of 15 meters and the requirement to wait for a stable position ensured a viable sensor performance in most cases. Other strategies such as setting the viewing angle to 120° also contributed to an improvement of the results. Without these restrictions and strategies the success rate would have drastically deteriorated especially in situations with reduced satellite-visibility.

Although the sensors were accurate enough to always include the intended objects in the FOV and to achieve relatively high success rates, an improvement of them would simplify the process of detecting the desired landmarks. This would probably also allow a further improvement of the results and an easing of the restrictions.

Research Question 3

> *Is the data density of OpenStreetMap high enough to provide sufficient landmark candidates for a user-generated landmark collection?*

In the frame of this thesis, the application has been tested in Zurich and Zumikon to estimate the influence of data density on the correct suggestion of landmarks. The corresponding tests were based on the assumption that the chosen regions are valid representatives of a typical urban and a predominantly rural area respectively. The density of OSM data varied considerably between these locations. Although there are fewer objects to be mapped in the rural area by nature, the completeness and the quality of the data could not match the dataset in Zurich. This is explained by the presence of more active mappers in urban areas [Zielstra and Zipf, 2010].

The lack of completeness of data had a direct influence on the number of missed landmarks: In Zurich only few objects could not be found, whereas the amount of missed landmarks in Zumikon exceeded the total number of found ones. Whether the data density of OSM is high enough for the application, is therefore strongly dependent on the investigated region.

Research Question 4

> *Does OpenStreetMap provide a valid method to store and access landmark information?*

Due to the stated lack of an official landmark-tag in OSM, the present thesis has introduced the key/value pair "uzh_landmark=true". Nevertheless, this is considered as a minimum solution. Neither is a justification involved nor does this method enable to store further landmark specific data like the captured picture, special characteristics etc. For a sophisticated referring to landmarks in navigational systems, more tags are needed. Complex references such as the example from Raubal and Winter [2002] from Section 2.1.1 *("... until Haas building, a dark building of architectural significance containing a (signed) Zara shop at the right")* would require a great number of different tags. In addition, further research is required to find an efficient conceptual model of storing and accessing this data.

One specific characteristic of this application is the exclusive querying and storing of the data on OSM. Yet, the saving of the derived data on a separate server would provide a more flexible solution. This would allow an unrestrained storing of all desired properties such as day/night visibility, salient features etc. of the landmarks.

5.6 Reached Goals

The objectives achieved in this work can be summarised as follows:

- Development of a camera-based application for Android phones to capture landmarks and to upload the results to OpenStreetMap.
- Introduction of a ranking system to determine the intended landmark based on smartphone sensors and OpenStreetMap metadata.

- Evaluation of the success rate of the application and thereby achieving rates paving the way for a user-generated collection of landmarks.

With the development of a mobile application for a user-generated collection of landmarks, the present thesis has shown the feasibility of such an approach. The results provided by this study have enabled a better understanding of how a widespread manual landmark collection can be implemented. With the inclusion of visual and semantic properties, a relevant ranking criterion has been found, which contributes to gather well-suited landmarks for navigation.

6 Conclusion and Outlook

In this Section the reached goals and the outcomes of this work are discussed. On that basis, it is concluded whether a camera-based approach combined with integrated smartphone sensors is a valuable approach to create a collection of landmarks. Furthermore, an outlook is given on possible future work on this subject.

6.1 Findings and Conclusion

This thesis has presented a method for a user-generated collection of landmarks in form of a mobile application on Android. Landmarks are a main feature request by users of navigational systems to be included in automatically generated routing instructions [May et al., 2003]. Nevertheless, they have not been established in widespread commercial systems until today as no widespread approach exists to extract or store them [Richter and Winter, 2011]. Previous research focused on automated methods to gather landmarks from existing data. A widespread usage of these approaches was hampered by vast data requirements, uneven landmark distribution or a focusing on global landmarks [Richter and Winter, 2011].

The worldwide dissemination of low-cost hardware parts has led to a widespread usage of ”intelligent” handheld devices. These phones have integrated sensors to derive the position of the phone and its viewing direction. The developed application makes use of these sensors to compute the visible field of the user. The underlying data is provided by OpenStreetMap. To tag objects as landmarks with the developed application, it is necessary to take a corresponding picture. By obtaining the GPS- and compass data in the moment of triggering the camera, the application defines a field of view. The objects inside this area are ranked by their possibility of being the intended landmark. This is achieved by an estimation of the object’s visual and semantic landmark properties. The user is then able to ”convert” the suggested candidates to landmarks by confirming them with manual input. Accumulated landmarks are subsequently tagged on OSM with the key ”uzh_landmark”.

By means of the presented application, one is able to inspect and label potential landmarks simultaneously. Nevertheless, the used sensors involve a certain degree of inaccuracy. Several strategies were introduced in this thesis to overcome these uncertainties such as a minimum accuracy for the GPS. In a further step, the success rate of determining the correct landmark was tested in an urban and a rural area. On account of the results obtained, it was established that a combination of both visual and semantic characteristics notably improves the results compared to using separate approaches. A strictly visual approach was afflicted with sensor inaccuracies and an approach limited to semantic characteristics suffered from the lack of recognizing the user’s intention.

In light of the aforementioned weaknesses of automated methods to collect landmarks, the presented approach provides several advantages: First of all, the introduction of a user-generated process of landmark collecting enables to substantially reduce the amount of required data as a large part of the data filtering is done by direct input of the user. Secondly, the application allows in principal the selection of arbitrary small objects and therefore the usage in local levels. However, the application only enables the tagging of already existing objects. Therefore, the results are solely dependent on the completeness and quality of OSM data. OSM has known deficiencies of coverage in rural areas [Haklay, 2010]. The low data density in these regions negatively affected the rate of found landmarks and thus the usability of the application in such areas. Hence, the presented method is not able to fully overcome the shortcoming of an uneven landmark distribution.

Particularly in urban regions with typically high coverage of data, the application was able to achieve a satisfying success rate in finding the intended landmark. This statement is underlined by a conducted user study, where the goal was to collect the landmarks already accumulated. Therefore, the findings of this thesis and the developed application contribute to the establishment of a standard practice for obtaining and distributing landmark information.

6.2 Future Research and Outlook

Based on the outcomes of this work, five remaining questions were brought to attention, which should be addressed in further research:

Uniform Definition of Landmarks

Duckham et al. [2010] stated that until now there are no agreed characteristics to define landmarks. Further studies need to assert a uniform definition, which also can be passed to non-expert users of such an application.

Motivation of Users

Future studies should examine, whether mappers are motivated to go around and collect landmarks in the provided way. There is the possibility that the given method is not seen as rewarding enough for a widespread usage, since the involved users may not see any directional benefits of their efforts. Additionally, this system suffers to a certain extent from sensor inaccuracies, missing landmarks and waiting times for a viable GPS signal. This may have a negative effect on the motivation of the users.

Input Parameters

The input weights of the ranking factors are not scientifically examined. Instead they are based on tests and personal assumptions. Extended user-studies might pursue the target of finding more appropriate weights for the input parameters. A possible research question might involve what users remember most about their desired landmarks, for example a distinct category or their size, etc. In addition, these studies could determine, if some of the used inputs really have a connection with an object being a landmark such as the number of tags or the frequency of a category.

Appropriate Way of Storing Landmark Data

Proper ways are needed to store and access landmarks and their corresponding attributes. This would allow the creation of descriptive routing instructions. For such a study, it has to be determined how algorithms can properly imitate people's reference to landmarks on a similar level of complexity. These methods must take into account that different landmarks need distinct attributes for their characterization. Depending on the complexity of the landmark description, a vast number of needed attributes are required. Therefore, the present thesis suggests to store this information on an own server system.

Examining Captured Landmarks

During this thesis, new landmarks were only captured by the author. However, the application paves the way for studies about the type of landmarks people are collecting. This would especially be interesting as most users are not familiar with the underlying OpenStreetMap data. Results such as the emerged tags during the investigation as seen in Figure 4.3 could give valuable insights of what people perceive about landmarks.

Closing words

Considering the possible usage of landmarks, a common method to gather them would provide large benefits. This thesis has presented an alternative approach to the identified problem of lacking information on landmarks by integrating the user in the process of gathering suitable objects. The established success rates in determining the intended landmark make this approach a viable alternative to previous automated methods. This may open the door for a widespread landmark information storage based on volunteered geographic information.

References

Ather, A.: A quality analysis of OpenStreetMap data. *ME Thesis, University College London, London, UK*, 2009.

Bauer, C.: On the (in-) accuracy of GPS measures of smartphones: A study of running tracking applications. *Studies*, 3:9, 2013.

Beeharee, A. K. & Steed, A.: A natural wayfinding exploiting photos in pedestrian navigation systems. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 81–88. ACM, 2006.

Bell, M.; Reeves, S.; Brown, B.; Sherwood, S.; MacMillan, D.; Ferguson, J. & Chalmers, M.: Eyespy: Supporting navigation through play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 123–132. ACM, 2009.

Blum, J. R.; Greencorn, D. G. & Cooperstock, J. R.: Smartphone sensor reliability for augmented reality applications. In *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 127–138. Springer, 2013.

Brimicombe, A. & Li, C.: Location-based services and geo-information engineering, Volume 21. John Wiley & Sons, 2009.

Burnett, G.: ‘Turn right at the traffic lights’: The requirement for landmarks in vehicle navigation systems. *Journal of Navigation*, 53(03):499–510, 2000.

Burnett, G.; Smith, D. & May, A.: Supporting the navigation task: Characteristics of’good’landmarks. *Contemporary ergonomics*, 2001:441–446, 2001.

Caduff, D. & Timpf, S.: The landmark spider: Representing landmark knowledge for wayfinding tasks. In *AAAI Spring Symposium: Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance*, pages 30–35, 2005.

Chamberlain, R. G.: What is the best way to calculate the great circle distance (which deliberately ignores elevation differences) between 2 points?, (Retrieved on: 22.07.2014). `http://www.faqs.org/faqs/geography/infosystems-faq/`, 1996.

Dale, R.; Geldof, S. & Prost, O.-P.: Using natural language generation in automatic route description. *Journal of Research & Practice in Information Technology*, 37(1), 2005.

Duckham, M.; Winter, S. & Robinson, M.: Including landmarks in routing instructions. *Journal of Location Based Services*, 4(1):28–52, 2010.

Edwardes, A. J. & Purves, R. S.: A theoretical grounding for semantic descriptions of place. In *Web and Wireless Geographical Information Systems*, pages 106–120. Springer, 2007.

Flanagin, A. J. & Metzger, M. J.: The credibility of volunteered geographic information. *GeoJournal*, 72(3-4):137–148, 2008.

Frommberger, L.; Schmid, F. & Cai, C.: Micro-mapping with smartphones for monitoring agricultural development. In *Proceedings of the 3rd ACM Symposium on Computing for Development*, page 46. ACM, 2013.

Ghasemi, M.: Developing a user generated method to add landmarks to OpenStreetMap. Master's thesis, The University of Melbourne, 2011.

Goodchild, M. F.: Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

Google: Android API - Activities, (Retrieved on: 20.07.2014). `http://developer.android.com/guide/components/activities.html`, 2014a.

Google: Android API - Class overview: Location, getAccuracy(), (Retrieved on: 06.07.2014). `http://developer.android.com/reference/android/location/Location.html#getAccuracy%28%29`, 2014b.

Google: Android API - Shutter callback, (Retrieved on: 20.07.2014). `http://developer.android.com/reference/android/hardware/Camera.ShutterCallback.html`, 2014c.

Google: Android API - Class overview: Sensor, (Retrieved on: 08.07.2014). `http://developer.android.com/reference/android/hardware/Sensor.html`, 2014d.

Haklay, M.: How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and planning. B, Planning & design*, 37(4):682, 2010.

Haklay, M. & Weber, P.: Openstreetmap: User-generated street maps. *Pervasive Computing, IEEE*, 7(4):12–18, 2008.

Hirtle, S. C. & Jonides, J.: Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13(3):208–217, 1985.

Jones, C. B.; Abdelmoty, A. I.; Smart, P. D.; Hall, M. M.; Quinn, J. & Twaroch, F. A.: Automated image captioning: The tripod project. *Geographic Information on the Internet*, page 85, 2009.

Kahl, W.: Navigation für Expeditionen, Touren, Törns und Reisen: Orientierung in der Wildnis. Schettler-Travel-Publikation. Schettler, 1991. ISBN: 978-3-8895-3301-2. URL: `http://books.google.ch/books?id=fwQBPQAACAAJ`.

Lomas, N.: Android still growing market share by winning first time smartphone users (Retrieved on: 23.07.2014). `http://techcrunch.com/2014/05/06/android-still-growing-market-share-by-winning-first-time-smartphone-users/`, 2014.

Lynch, K.: The image of the city, Volume 11. MIT press, 1960.

Maus, S.; Macmillan, S.; McLean, S.; Hamilton, B.; Thomson, A.; Nair, M. & Rollins, C.: The US/UK world magnetic model for 2010-2015. 2010.

May, A. J.; Ross, T.; Bayer, S. H. & Tarkiainen, M. J.: Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing*, 7(6):331–338, 2003.

Michon, P.-E. & Denis, M.: When and why are visual landmarks used in giving directions? In *Spatial information theory*, pages 292–305. Springer, 2001.

Modsching, M.; Kramer, R. & ten Hagen, K.: Field trial on GPS accuracy in a medium size city: The influence of built-up. In *3rd Workshop on Positioning, Navigation and Communication*, pages 209–218, 2006.

Mooney, P.; Corcoran, P. & Winstanley, A. C.: Towards quality metrics for OpenStreetMap. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 514–517. ACM, 2010.

Morrill, D.: Android developers blog - one screen turn deserves another, (Retrieved on: 11.07.2014). `http://android-developers.blogspot.ch/2010/09/one-screen-turn-deserves-another.html`, 2010.

NCSU: Accuracy and precision, NC State University, (Retrieved on: 21.09.2014). `http://www.ncsu.edu/labwrite/Experimental%20Design/accuracyprecision.htm`, 2004.

Nothegger, C.; Winter, S. & Raubal, M.: Selection of salient features for route directions. *Spatial cognition and computation*, 4(2):113–136, 2004.

Nurseitov, N.; Paulson, M.; Reynolds, R. & Izurieta, C.: Comparison of JSON and XML data interchange formats: A case study. *Caine*, 2009:157–162, 2009.

Olbricht, R.: Overpass API, (Retrieved on: 04.06.2014). `http://wiki.openstreetmap.org/wiki/Overpass_API`, 2011.

OSM: Naptan, (Retrieved on: 05.06.2008). `http://wiki.openstreetmap.org/wiki/NaPTAN`, 2010.

OSM: Elements, (Retrieved on: 04.06.2014). `http://wiki.openstreetmap.org/wiki/Elements`, 2011.

OSM: Key:amenity, (Retrieved on: 22.07.2014). `http://wiki.openstreetmap.org/wiki/Amenity`, 2012a.

OSM: Key:tourism, (Retrieved on: 22.07.2014). `http://wiki.openstreetmap.org/wiki/Key:tourism`, 2012b.

OSM: Points of interest, (Retrieved on: 23.08.2014). `http://wiki.openstreetmap.org/wiki/POI`, 2013.

OSM: About, (Retrieved on: 28.07.2014). `http://wiki.openstreetmap.org/wiki/About`, 2014a.

OSM: One feature, one OSM element, (Retrieved on: 23.08.2014). `http://wiki.openstreetmap.org/wiki/One_feature,_one_OSM_element`, 2014b.

OSM: Relation, (Retrieved on: 08.08.2014). `http://wiki.openstreetmap.org/wiki/Relation`, 2014c.

OSM: Stats, (Retrieved on: 08.08.2014). `http://wiki.openstreetmap.org/wiki/Stats`, 2014d.

OSM: Tags, (Retrieved on: 19.09.2014). `http://wiki.openstreetmap.org/wiki/Tag`, 2014e.

Papadopoulos, S.; Zigkolis, C.; Kompatsiaris, Y. & Vakali, A.: Cluster-based landmark and event detection on tagged photo collections. *IEEE Multimedia*, 2010.

Purves, R. S.; Edwardes, A. & Sanderson, M.: Describing the where–improving image annotation and search through geography. In *Proceedings of the workshop on Metadata Mining for Image Understanding (MMIU 2008)*, 2008.

Purves, R. S.; Edwardes, A.; Fan, X.; Hall, M. & Tomko, M.: Automatically generating keywords for georeferenced images. In *Proceedings of the GIS Research UK 18th Annual Conference GISRUK 2010*, pages 203–207, 2010.

Raubal, M. & Winter, S.: Enriching wayfinding instructions with local landmarks. Springer, 2002.

Richter, K.-F.: Context-specific route directions: Generation of cognitively motivated wayfinding instructions, Volume 314. IOS Press, 2007.

Richter, K.-F.: Prospects and challenges of landmarks in navigation services. In *Cognitive and Linguistic Aspects of Geographic Space*, pages 83–97. Springer, 2013.

Richter, K.-F. & Winter, S.: Harvesting user-generated content for semantic spatial information: The case of landmarks in OpenStreetMap. *Proceedings of the SPATIAL SCIENCES & SURVEYING BIENNIAL CONFERENCE 2011 21-25 November 2011, Wellington, New Zealand*, (November), 2011.

Sadeghian, P. & Kantardzic, M.: The new generation of automatic landmark detection systems: Challenges and guidelines. *Spatial Cognition & Computation*, 8 (3):252–287, 2008.

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

Sinnott, R. W.: Virtues of the haversine. *Sky and telescope*, 68:158, 1984.

Sorrows, M. E. & Hirtle, S. C.: The nature of landmarks for real and electronic spaces. In *Spatial information theory. Cognitive and computational foundations of geographic information science*, pages 37–50. Springer, 1999.

Surowiecki, J.: The wisdom of crowds. Random House LLC, 2005.

Swiss Taginfo: TagInfo Amenity (Retrieved on: 25.07.2014). `http://taginfo.osm.ch/keys/amenity`, 2014a.

Swiss Taginfo: TagInfo uzh_landmark (Retrieved on: 30.08.2014). `http://taginfo.osm.ch/keys/uzh_landmark`, 2014b.

Swiss Taginfo: TagInfo Station (Retrieved on: 15.08.2014). `http://taginfo.osm.ch/search?q=station`, 2014c.

Tomko, M.: Case study-assessing spatial distribution of web resources for navigation services. In *Proceedings of the 4th International Workshop on Web and Wireless Geographical Information Systems*, pages 90–104, 2004.

Tripomatic: Haas-building (Retrieved on: 07.08.2014). `http://www.tripomatic.com/Austria/Vienna/Haas-Building/`, 2014.

Veness, C.: Calculate distance, bearing and more between latitude/longitude points (Retrieved on: 22.07.2014). `http://www.movable-type.co.uk/scripts/latlong.html`, 2014.

Vinson, N. G.: Design guidelines for landmarks to support navigation in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 278–285. ACM, 1999.

von Watzdorf, S. & Michahelles, F.: Accuracy of positioning data on smartphones. In *Proceedings of the 3rd International Workshop on Location and the Web*, page 2. ACM, 2010.

Winter, S.; Tomko, M.; Elias, B. & Sester, M.: Landmark hierarchies in context. *Environment and Planning B: Planning and Design*, 35(3):381, 2008.

Zielstra, D. & Zipf, A.: A comparative study of proprietary geodata and volunteered geographic information for Germany. In *13th AGILE international conference on geographic information science*, Volume 2010, 2010.

Appendix

Appendix A: Individual Values for Category Weighting

This table shows the category values used in the semantic ranking:

amenity:		**tourism:**		**railway:**		**sport:**	**shop:**	**historic:**	**leisure:**
place_of_worship	10	museum	9	station	7	4	5	9	6
university	10	attraction	9	tram_stop	6				
public_building	9	gallery	8						
school	8	hotel	5						
theatre	8	information	5						
library	8	viewpoint	4						
artwork	7	other	2						
post_office	6								
fuel	6								
cinema	6								
police	6								
restaurant	5								
cafe	5								
bar	5								
fast_food	5								
pub	5								
pharmacy	5								
fountain	5								
bank	5								
nightclub	5								
car_wash	5								
arts_centre	5								
kindergarten	4								
stripclub	4								
ferry_terminal	4								
community_centre	4								
gallery	4								
doctor	3								
dentist	3								
car_rental	3								
bicycle_rental	3								
sauna	3								
veterinary	3								
brothel	3								
clinic	3								
marketplace	3								
drinking_water	2								
parking	2								
post_box	2								
car_sharing	2								
atm	2								
taxi	2								
bbq	2								
clock	2								
shelter	2								
shower	2								
childcare	2								
swimming_pool	2								
bench	1								
waste_basket, waste	1								
toilets	1								
recycling	1								
telephone	1								
vending_machine	1								
fire_hydrant	1								
compressed_air	1								

Appendix B: All Individual Results of the Field- and User Study

This table shows both ranking positions of the study from every ranking-landmark combination including the user study.

Landmark	Both Rankings Near	Both Rankings Far	Only Semantic Near	Only Semantic Far	Only Visual Near	Only Visual Far	User Study
Rosen-Apotheke	1 / 1	2 / 2	3 / 3	4 / 4	2 / 5	4 / 2	1 / 1
Fountain Stüssihofstatt	1 / 1	1 / 1	3 / 3	4 / 4	1 / 1	2 / 2	1 / 2
Hafenkran	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 2
St. Peter church	1 / 1	1 / 1	1 / 1	1 / 1	3 / 3	1 / 1	1 / 1
Beer Buchhandlung	1 / 1	1 / 1	1 / 1	1 / 1	3 / 2	4 / 10	1 / 1
Fountain next to Augustinerkirche	2 / 2	8 / 9	15 / 15	14 / 14	1 / 1	2 / 4	2 / 2
Pavillon Skulptur	1 / 1	1 / 1	3 / 3	3 / 3	1 / 1	1 / 1	1 / 2
Paradeplatz	1 / 1	1 / 1	1 / 1	1 / 1	2 / 3	2 / 2	1 / 1
Haus zur Sichel	3 / 2	3 / 5	3 / 3	3 / 3	2 / 2	6 / 7	1 / 1
Öpfelchammer	1 / 1	2 / 3	1 / 1	1 / 1	2 / 2	14 / 15	1 / 1
Predigerplatz 18	3 / 1	5 / 4	7 / 7	6 / 6	3 / 1	2 / 2	3 / 3
Läbis	2 / 1	3 / 5	5 / 5	4 / 5	1 / 1	7 / 8	7 / 1
Statue Hans Waldmann	1 / 1	1 / 2	2 / 2	3 / 2	1 / 1	3 / 4	1 / 1
Telefonzelle bei Rudolf Brun Brücke	3 / 2	3 / 3	7 / 7	6 / 7	4 / 2	2 / 3	1 / 1
Grossmünster	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
Brunnen hinter Grossmünster	3 / 4	6 / 4	6 / 6	6 / 6	2 / 2	2 / 3	3 / 3
Haus zum Paradies	2 / 1	3 / 4	9 / 5	9 / 10	2 / 4	4 / 2	3 / 3
Bodega	2 / 3	1 / 1	2 / 2	2 / 2	5 / 6	1 / 1	1 / 4
Kaffee Schoffel	1 / 3	2 / 3	1 / 1	1 / 1	3 / 14	8 / 6	2 / 1
Haus zum Rüden	3 / 1	1 / 1	2 / 3	1 / 1	4 / 4	4 / 2	1 / 1
Bahnhof Zumikon	1 / 1	2 / 3	2 / 2	2 / 2	2 / 1	5 / 5	-
Gemeindehaus	1 / 2	1 / 1	1 / 1	1 / 1	2 / 2	2 / 2	-
Kirche Zumikon	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	5 / 2	-
Spielplatz	1 / 1	2 / 2	3 / 3	3 / 3	1 / 1	5 / 4	-
Gasthof Rössli	1 / 1	1 / 1	1 / 1	2 / 1	1 / 1	1 / 1	-
ZKB Filiale	1 / 1	1 / 1	1 / 1	1 / 1	2 / 2	1 / 1	-
Post Box	2 / 1	1 / 1	3 / 3	4 / 4	1 / 1	1 / 1	-
Volg	1 / 1	1 / 1	1 / 1	2 / 2	1 / 1	1 / 1	-
Schulhaus Farlifang	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	-
Hallenbad Zumikon	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	-

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Zurich, 30.09.2014

Marius Wolfensberger