A Mobile Application for a User-Generated Collection of Landmarks

Marius Wolfensberger and Kai-Florian Richter(✉)

Department of Geography, University of Zurich - Irchel,
Winterthurerstrasse 190, 8057 Zurich, Switzerland
{marius.wolfensberger,kai-florian.richter}@geo.uzh.ch

Abstract. Landmarks are crucial elements in how people understand and communicate about space. In wayfinding they provide references that are preferred and easier to follow than distances or street names alone. Thus, the inclusion of landmarks into navigation services is a long-held goal, but its implementation has largely failed so far. To a large part this is due to significant difficulties in obtaining a sufficient data set of landmark candidates. In this paper, we introduce a mobile application, which enables a user-generated collection of landmarks. Employing a photo-based interface, the application calculates and ranks potential landmark candidates based on the current visible area and presents them to the user, who then may choose the intended one. We use OpenStreetMap as data source; the app allows tagging OSM objects as potential landmarks. Integrating users into the landmark selection process keeps data requirements low, while a simple interface lowers the burden on the users.

Keywords: Landmarks · Volunteered geographic information · User-generated content · OpenStreetMap · Location-based services

1 Introduction

Geographic landmarks are defined as "any element, which may serve as reference points" [10]. They are easily distinguishable environmental features that are unique in or in contrast with their neighborhood [15,20]. Landmarks are fundamental in how humans understand and represent their environment and how they communicate about it [19].

Current navigation services construct their guidances exclusively based on metrics (time or distance), orientation and street names [16]. However, the use 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 [2] and can easily be twisted by outside influences (traffic lights, crowded pathways) [23]. To overcome these deficiencies, landmarks should be included in routing instructions. Particularly at decision points, where a reorientation is needed, they increase the performance and efficiency of users (e.g., [11]).

Currently, there are very few commercial systems that include landmarks in their navigation instructions. The primary reason is the lack of available landmark data [5,17]. There are neither widespread possibilities to access and store

J. Gensel and M. Tomko (Eds.): W2GIS 2015, LNCS 9080, pp. 3–19, 2015.
DOI: 10.1007/978-3-319-18251-3_1

landmarks [23], nor standardized characteristics defining landmarks [5]. Previous research focused on automated methods to extract landmarks from existing data. A widespread use of these approaches was hampered by vast data requirements, uneven landmark distribution, or a focus on global landmarks [18].

We developed a mobile application that provides a tool for collecting and sharing of landmark data. This tool allows the in-situ labelling of objects as landmarks, i.e., while being in the environment and close to the landmark. Using their smartphone to take a photo of the desired landmark, users receive a list of potential landmark candidates, ranked by their probability of being the landmark the user intends to collect. In order to calculate the probability of the involved candidates, a ranking system is used, which estimates the visual and semantic suitability of the examined geographic objects. 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[1].

The next section will present relevant related work. In Section 3 we will discuss some of the challenges in enabling an in-situ landmark collection and illustrate our approach. The implemented Android app is presented in Section 4, while Section 5 shows results of a small case study we performed in order to evaluate the application. Section 6 discusses our approach in light of this case study, and Section 7 finally concludes the paper with an outlook on future work.

2 Previous Work

Several automated methods have been suggested in the past for the purpose of landmark extraction. The first method was developed by Raubal and Winter [16]. Their approach transforms the three main characteristics of landmarks proposed by Sorrows and Hirtle [21] into attributes that make these characteristics computable. Sorrows and Hirtle specified three main categories of landmarks: *Visual*, *semantic* and *structural* landmarks. *Visual* landmarks are considered landmarks due to their visual prominence. *Semantic* landmarks stick out because of their historical or functional importance. *Structural* landmarks are characterized by the importance of their location or their role in space (e.g., at intersections).

In Raubal and Winter's approach extracted attributes are compared to those of surrounding objects to decide whether something is a potential landmark. Since landmarks should be unique in their neighborhood, 'landmarkness' is a relative characteristic [13,19]. Accordingly, the identification process needs to account for nearby objects. Objects may be considered a landmark if their attributes differ significantly from those of the surrounding objects. However, this approach requires a vast amount of detailed data, which hinders its broad application [17].

Other approaches use data mining approaches for the identification of landmarks using various geographic and non-geographic data sources (e.g., [4,14,22,23]. However, such approaches often only manage to detect the most famous,

[1] www.openstreetmap.org

'touristic' landmarks, but fail to pick up local landmarks, such as the small corner store in a residential neighborhood.

Duckham et al. [5] compared the category information of so-called points of interest (POI) with their surroundings in order to obtain their "landmarkness". Individual POIs were ranked by the general landmark suitability of their category (e.g., a church being generally more suited than a lawyer's office), and the uniqueness in their area. Category information is significantly more available than detailed data about an individual object's shape, color, or size. Thus, the amount of required data is greatly reduced. Nevertheless, this method still suffers from an unequal distribution of geographic (POI) information [17]. An additional limitation is that the employed heuristics may simply go wrong. Certain objects may be highly unsuitable landmarks despite their category being generally well suited. For example, while typical churches are highly suitable landmarks, as they are large, recognizable and semantically as well as architecturally distinct from their surroundings, some churches, for instance a small church-room inside an airport, cannot be considered salient [5].

To face the aforementioned difficulties arising from automated landmark identification, Richter and Winter [17,18] suggest applying principles and methods of "Volunteered Geographic Information" (VGI) in the collection process of landmark information. VGI is a form of user-generated content, specifically targeted at the acquisition of geographic information [7]. The goal is to provide a method allowing a straightforward way to collect and share landmark information. In such an approach, users perform the identification of what 'sticks out from the background', i.e., implicitly or explicitly filter objects with respect to their 'landmarkness'. Consequently, the lack of sufficient existing geographic data disappears. However, it is replaced by a need for a simple mechanism for collecting landmarks because otherwise it will be impossible to attract a sufficient number of users. Some previous work by Richter and Winter started off in this direction [6,18], but did not (yet) run on mobile devices and fell short in terms of usability. We believe that the mobile application presented in this paper solves these issues to a large extent.

3 In-Situ Collection of Landmark Candidates

Our aim is to provide a tool for the manual selection of landmarks while being in the environment close to a landmark. Since one of our requirements is a simple, easy-to-use interface, we opt for a photo-based collection procedure. This way, we create a kind of 'point&click' interface. Users take a photo of the geographic object they intend to mark as landmark. In that moment the application registers the user's position (via GPS) and heading (via the inbuilt compass sensor). With this sensor information the application calculates the geographic area visible to the user and retrieves the associated geographic data. This data is ranked according to the likelihood of being the intended object selected. In a final step, users have to confirm (or reject) any of the suggested landmark candidates.

It is important to note that we refrain from any content-based image retrieval approaches in our application. The resulting photo is saved to the phone, but not scanned or analyzed in any other form for possible landmark candidates. Our extraction method is based on location sensor data, and not on image content. Taking a photo is only used as a trigger to collect this sensor data, and because it offers an easy interface that for the user closely links the real-world geographic object with the selection process on the smartphone. Also, for the time being we restrict landmark selection to geographic objects in built environments.

Involving users in the process of landmark identification further has the advantage of creating a dataset directly based on human cognition. The application's main task is to automatically compute useful suggestions, so that the user can confirm the intended landmark. An important factor is to ensure a high performance, i.e., low latency between taking a photograph and confirming the selected object. This process provides several challenges; in this paper we focus on the following implementation aspects: Dealing with sensor inaccuracies in determining the visible area; extracting possible landmark candidates out of the objects in that area; quantification of the candidates' 'landmarkness' attributes; ranking of the remaining candidates by their suitability as landmarks.

3.1 Determining the Visible Area from Location Sensors

The first step is to determine the geographic objects visible to the user. As stated above, we obtain a user's position from the in-built GPS sensor and the heading (viewing direction) from the compass. Combining this sensor information allows for computing a field of view which includes all visible objects. Figure 1 shows an example of such a field of view, indicated by the triangle. Estimating the visible area is severely affected by a mobile device's sensor inaccuracies.

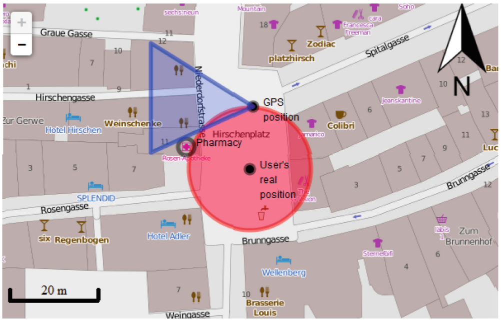

Fig. 1. Miscalculated field of view (triangle) due to GPS inaccuracies leads to missing the pharmacy. The circle indicates an inaccuracy radius of 15 meters (map source (c) OpenStreetMap users; CC BY-SA).

Smartphones use low-cost hardware parts. Consequently, their sensors have rather large inaccuracies [1]. GPS has an accuracy of 5 to 10 meters depending on satellite visibility, which is also achieved by smartphones. However, there is large variation in this accuracy. Even in wide streets inaccuracy can reach up to 15 meters, and much more in narrow lanes with tall surrounding buildings [12].

The compass of handheld devices typically has a mean error of 10 to 30 degrees (in either direction) while the device is moving [3], which differs from device to device. Also, the user is not meant to move while capturing a landmark. Therefore, we performed our own compass test in order to calibrate our application. The test used a mirror compass with a magnetic needle as ground truth (Recta DP-2) and two different cell phones (HTC One and Samsung Galaxy S II). Data was collected while keeping the phones stationary. The smartphone compass showed a mean error of approximately 16 degrees to the mirror compass.

These inaccuracies need to be accounted for when calculating the field of view and, accordingly, the visible objects. In Figure 1, the pharmacy is not detected due to a GPS inaccuracy of 15 meters. Although the compass returns an accurate result, instead the nearby restaurant will be shown as a landmark candidate. Similar errors can occur if the compass returns inaccurate results. Errors, such as this, cannot be fully avoided, but we implemented several strategies to decrease the influence of GPS and compass inaccuracies (discussed in Section 4.3).

3.2 Extracting Possible Landmark Candidates

The geographic data in the computed visible area needs to be scanned for possible landmark candidates. As the number of candidates can become very large if the viewing distance is not restricted, we focus on objects near to the users. We limit these candidates to either point entities or polygons of building size or smaller, since the capturing of paths or other linear features and of large areas is rather difficult to achieve with a mobile phone camera. The minimum demands of an entity to be a potential landmark candidate are to have a location, at least one tag and the possibility to derive a name out of the object's metadata. The name must either be directly available or other attributes, for example, category information or the address, must allow an appropriate naming. This ensures that the landmark can be referred to and users are able to recognize the suggested landmarks.

Furthermore, we restrict the categories of point data considered as landmark candidates. Categories, which frequently occur in clusters, such as pedestrian crossings or traffic signals, are excluded since it is difficult to reliably assign suggested candidates to the correct real world object. Some point objects are discarded as they only appear as part of larger units, such as building entrances. Vegetation is also excluded, based on the fact that trees and other plants are often subject to rapid change and, thus, are rather unreliable landmarks [24]. All objects in the visible area that do not meet these requirements are discarded. The remaining objects represent the set of potential landmark candidates.

3.3 Quantification of Landmarkness Attributes

In order to calculate the most probable candidate, a quantification of the given metadata and sensor information needs to be performed. For this reason, measurable attributes are allocated to the landmark characteristics of Sorrows and Hirtle (see Section 2). These attributes need to be available for a large part of objects in order to include and compare as many candidates as possible. Figure 2 shows characteristics determining a landmark's saliency. Characteristics that can be determined for most geographic entities are highlighted. These can be quantified and, thus, compared between different entities.

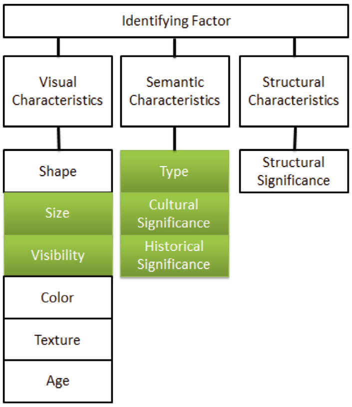

Fig. 2. The identifying characteristics defining landmark salience. Colored boxes show the characteristics which can be derived for most of the available data (after [18]).

In contrast to other approaches, which use similar characteristics to quantify their landmark candidates (e.g., [16]), our method incorporates also sensor data of a mobile device. The visual characteristics 'visibility' is calculated from sensor values. Characteristics not accounted for in our approach are either only rarely available (e.g., *color*) or not available at all (*3D shape*). In selecting the involved characteristics we keep the computational effort for quantification low to ensure a high level of performance. Consequently, 'structural significance' is not taken into account. 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 transportation network. However, since the presented method relies on direct user input, we assume that users (implicitly or explicitly) account for structural significance in their selection of landmark candidates.

The selected characteristics are calculated using the following data. Apart from area and visible range, which are only available for polygons, the selected underlying data can be derived for all geographic entities that include at least category information.

– Size: Area (only available for polygons);
– Visibility: Distance and azimuth deviation to the user, visible range (the angle range in which an object is visible to the user - only for polygons);
– Type: Tags describing the function / category of an object (e.g., *amenity, leisure, shop*);
– Cultural / historical significance: Number of tags, background information (object's own website / Wikipedia article[2]), frequency of the category in surrounding area.

3.4 Ranking of Landmark Candidates

Geographical objects have differing suitability to act as landmarks. Therefore, a ranking system is introduced employing an entity's metadata and visibility to find the most appropriate landmark candidate. Based on the categorization of Sorrows and Hirtle [21], Raubal and Winter [16] developed a measure for determining the salience of a specific object:

$$s_{vis} \cdot w_{vis} + s_{sem} \cdot w_{sem} + s_{str} \cdot w_{str} \tag{1}$$

s stands for the salience measure and w is a weighting factor. The indices *vis*, *sem* and *str* describe visual, semantic and structural salience, respectively. As just discussed, structural characteristics are not taken into consideration in this approach. Thus, s_{str} is dropped from Equation 1. The remaining parameters s_{vis} and s_{sem} are calculated using the attributes listed in Section 3.3.

Calculating the Ranking Factor for Visual Characteristics
In order to derive the factor s_{vis} describing the visual salience, the size of an object (X_{size}), the visible range (X_{vis_range}), the azimuth (X_{az}), and distance to the user ($X_{distance}$) are accounted for. i refers to the current object and min and max to the respective minimum or maximum value for all objects. A normalization into the range [0,1] is performed for each parameter. Divisions by zero are handled in all cases with the return of the value 0.

Size (only polygons)

$$X_{size} = \frac{X^{i}_{size} - X^{min}_{size}}{X^{max}_{size} - X^{min}_{size}} \tag{2}$$

Visible range (only polygons)

$$X_{vis\_range} = \frac{X^{i}_{vis\_range} - X^{min}_{vis\_range}}{X^{max}_{vis\_range} - X^{min}_{vis\_range}} \tag{3}$$

The formulas for X_{az} and $X_{distance}$ are squared in order to prioritize nearness and compliance with the azimuth:

[2] www.wikipedia.com

Azimuth

$$X_{az} = \Big(\frac{\pi - ((2\pi + X_{az}^{sensor} - X_{az}^{i}) \mod 2\pi)}{\pi}\Big)^2 \tag{4}$$

X_{az}^{sensor} stands for the sensor's azimuth value. If the azimuth of the object is 180 degrees 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. In the case of a polygon, the most outside edges of the entity are used as reference points to calculate the azimuth deviation.

Distance

$$X_{distance} = \Big(\frac{X_{user\_max} - X_{dist}^{i}}{X_{user\_max}}\Big)^2 \tag{5}$$

The maximum viewing distance X_{user_max} is set in the application as a parameter. This gives the following equations for the calculation of the factor s_{vis}. For polygons:

$$\begin{aligned} s_{vis} = {} & X_{size} \cdot w_{size} + X_{az} \cdot w_{az_{poly}} \\ & + X_{distance} \cdot w_{distance_{poly}} \\ & + X_{vis_{range}} \cdot w_{vis_{range}} \end{aligned} \tag{6}$$

And for points:

$$s_{vis} = X_{az} \cdot w_{az_{point}} + X_{distance} \cdot w_{distance_{point}} \tag{7}$$

Calculating the Ranking Factor for Semantic Characteristics
The factor s_{sem} defining the semantic characteristics is calculated in a similar way to s_{vis}. The involved parameters X_{type} (type) and X_{signif} (significance) are also 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 of that category. This is following Duckham et al.'s [5] approach to using categories in determining landmark candidates (see Section 2). We assigned each of the most frequent categories in our data a suitability factor in the range of [1,10]. Any entity of a category that is not assigned a value to will receive a factor of 1.

$$X_{type} = \frac{X_{type}^{i} - X_{type}^{min}}{X_{type}^{max} - X_{type}^{min}} \tag{8}$$

X_{type}^{min} is the lowest and X_{type}^{max} the highest category value in the visible area.

Significance
Cultural and historical significance are combined in a factor X_{signif}, since no clear distinction can be made between these two factors without checking other sources than the metadata. As stated in Section 3.3, this parameter captures the number of tags X_{tag}, the frequency of the category X_{freq} and potential background information, such as a website $\phi_{website}$ or a Wikipedia article ϕ_{wiki}.

$$\begin{aligned} X_{signif} = & \left(1 - \frac{X^{i}_{freq} - X^{min}_{freq}}{X^{max}_{freq} - X^{min}_{freq}}\right) \cdot w_{freq} \\ & + \frac{X^{i}_{tag} - X^{min}_{tag}}{X^{max}_{tag} - X^{min}_{tag}} \cdot w_{tag} \\ & + \phi_{wiki} + \phi_{website} \end{aligned} \tag{9}$$

Accordingly, s_{sem} is calculated as:

$$s_{sem} = X_{type} \cdot w_{type} + X_{signif} \cdot w_{signif} \tag{10}$$

4 Implementation

For our application, Android was chosen as the development platform due to its high market share (around 80% in 2014; [9]) and the openness of its system. As with most Android applications ours is implemented in Java.

4.1 Geographic Data

OpenStreetMap (OSM) is used as underlying geographic data. OSM allows for a world-wide unrestricted access to geographic data [8]. The associated geographic data in the calculated visible area is downloaded using the Overpass API[3].

The conceptual data model of OSM consists of three basic geometric components: *Nodes*, *ways* and *relations*.[4] Nodes represent specific coordinate points as standalone entities or as part of a more complex geometry. Ways consist of at least two nodes and represent polylines. If the first node is equal to the last one (closed ways), they represent polygons describing the geometry of areas, for example, of buildings. Relations are used to define logical or geographic relationships between elements, for example, a building with an inner and an outer geometry. All these elements can be described in more detail by using tags. A tag is a key/value pair, describing one feature of a specific element (e.g., stating that a particular polygonal entity represents a hotel). As there is no established OSM landmark tag, we use the tag "uzh_landmark" to label the collected landmarks.

[3] www.overpass-api.de
[4] http://wiki.openstreetmap.org/wiki/Elements; retrieved 04.06.2014

4.2 Collecting a New Landmark

On start-up the application ensures that there is a GPS signal and the device has Internet connection. Once the user takes a photograph and acknowledges that this is indeed the photo they wanted to take, the visible area is calculated and OSM data is downloaded. The data is filtered for potential landmark candidates, which are then ranked according to their suitability using the formulas presented in Section 3.4. Figure 3 shows a screenshot of how this ranking is presented to the users. In the top right corner is the photo previously taken by the user. Next to it is the top-ranked object (name, category and distance to the user) listed as the primary suggestion of a landmark candidate. Beneath are four additional suggestions (the next four objects in the ranking). The application also shows a map of the user's location and viewing direction.

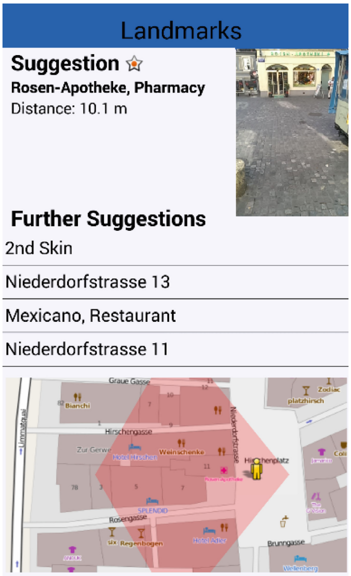

Fig. 3. The interface for the list of suggested landmarks (map source (c) OpenStreetMap users; CC BY-SA)

4.3 Dealing with Sensor Inaccuracies

As discussed in Section 3.1, the application needs to deal with sensor inaccuracies, which may lead to suggesting unintended landmark candidates. Several strategies are used to prevent such errors. The GPS position and its accuracy as well as the calculated visible area are shown to the users (see Figure 3), so they are able to check the sensor performance. The size of the visible area is adapted to the current GPS accuracy to avoid missing any objects that may otherwise fall outside this area. GPS inaccuracy may not exceed 15 meters; otherwise the application refuses to take a photo. Tests during development showed that inaccuracies greater than 15 meters often resulted in unreliable performance.

The viewing angle is fixed at 120 degrees. This value provides good results in circumventing sensor inaccuracies without losing possible objects. The average compass inaccuracy showed a mean error of around 16 degrees (Section 3.1). Therefore, any azimuth deviation of an object to the provided compass value smaller than 16 degrees in either direction is still considered as in front of the user. Finally, five potential landmark candidates are suggested to the users. This increases the chance that the intended landmark is included in the results.

4.4 Weighting Parameters

Our current implementation uses the parameter values listed in Table 1. These values were determined empirically; they show good performance in the environments we ran our tests in. Accordingly, these values are not necessarily of general validity, and changing some of them slightly will likely not have any major impact. However, especially the weights w_{vis} and w_{sem} can significantly change the outcomes as our evaluation has shown (see Section 5). Overemphasizing the visual characteristics leads to unlikely results as distance to and size of objects become overriding factors. Overemphasizing semantic characteristics basically ignores sensor feedback and, thus, may miss out on landmark candidates.

Table 1. The weighting factor values used in the application

Weighting factor	Parameter	Weight
General Factors:		
Visual weight	w_{vis}	3
Semantic weight	w_{sem}	2
Individual visual Characteristics:		
Parameters for polygons:		
Size	w_{size}	2
Visible range	w_{vis_range}	2
Azimuth deviation	$w_{az_{poly}}$	10
Distance to user	$w_{distance_{poly}}$	10
Parameters for nodes:		
Azimuth deviation	$w_{az_{point}}$	12
Distance to user	$w_{distance_{point}}$	12

Weighting factor	Parameter	Weight
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

5 Case Study

We performed a small case study as a proof of concept and to get a feel for the performance of our landmark collection application. This study has two parts. First, we marked different geographic objects as landmarks, seeing how often the intended objects show up in the list of suggestions. This test was run in an urban and a more rural (small town) environment. Second, we had a naive user collect (the same) landmarks to get some first impressions of usability.

5.1 Experimental Setup

We tested our application in an area in the inner city of Zurich, Switzerland, and the small town Zumikon. The geographic data of the chosen areas shows great variety in 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. During this test, 30 landmarks were collected with the application; 20 in Zurich and 10 in Zumikon.

We did not predefine objects to mark, but selected them in-situ to cover a range of different landmarks. Candidates included prominent and less prominent geographic objects, located in regions with and without surrounding buildings, low and high density of candidates, and identical categories next to each other. In the first part of the test each object was captured either by considering only visual characteristics, only semantic characteristics, or both combined. Additionally every landmark was captured from a near (5-15m) and a far distance (15-35m). In every setting, each object was captured twice to reduce randomness of the results. The maximum viewing distance was set to 50 meters throughout the entire test. The test was performed with a HTC One smartphone.

In the second part of the study, a naive user, who did not know the application beforehand and has no background in geographic information or computer science, was asked to capture the same landmarks in the Zurich area as selected by us. In this test, we chose the optimal settings for the application, namely both rankings activated and capturing landmarks from a near distance, and used the same smartphone as before. This test with the naive user was performed in order to see whether people unfamiliar with the application achieve similar results, and whether there are any obvious usability issues that we had previously missed.

5.2 Results

For the inner city area in Zurich 838 landmark candidates were counted on an area of 0.351km^2. Zumikon has 24 of such candidates in an area of 0.224km^2.[5]

Table 2 lists the results of the study. It shows the number of missed and found landmarks, and the according success rate in finding the desired landmark. 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 either have no representation in OSM, cannot be found by the application, or offer no possibility to derive a name from the metadata. A subsequent check showed that the missed landmarks were caused exclusively by non-existing OSM data for the desired geographic object. Therefore, missed landmarks are not included in the ranking results, as they would not explain the performance of the ranking system.

The second part of the table shows the 'hit rate' of the application, i.e., how often (in %) landmarks ended up on a particular ranking position. The average positions in Table 2 suggest that the best detection is achieved when both rankings are activated and landmarks are captured from a near distance

[5] Based on OSM data from 03.08.2014.

Table 2. Overall results of the study

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
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

with a mean ranking position of 1.6 (Zurich) and 1.1 (Zumikon), respectively. By increasing the distance to the landmark to 15 to 35 meters, this value deteriorates to 2.5 (1.25). The semantic ranking provides an average position of 3.625 (1.5) in near distance and 3.65 (1.75) from the far distance. The decline caused by the increased distance is considerably smaller when only using this ranking. The visual ranking has the highest decline in position between near and far distance with an average of 2.525 (1.25) for close distance and 3.7 (2.1) from far distance

In Zurich, nearly two-thirds (62.5%) of all near distance attempts using both rankings were placed on the first position, for the far distance this decreased to nearly half the attempts (47.5%). 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% and 75%).

With both rankings activated, only very few landmark capture attempts result in rankings below 5th place. With only semantic ranking, from both distances around a quarter of attempts ended below 5th place; for visual ranking this happened in 5% of the cases for the near distance (22.5% for the far distance). In Zumikon no intended object was ever ranked below 5th place since the number of landmark candidates in any visible area was small to begin with.

The average measurement deviation states the average difference in ranking between the two captures for each object. It gives an indication of the robustness of the results. In Zurich, the semantic ranking has the highest stability through lesser dependence on exact sensor data, whereas the visual ranking has the highest instability in the position (which is still only about 1). Accordingly, the combined ranking is in-between these two. Zumikon shows a similar pattern, although the difference between the different deviations is much smaller.

Results for the second part of the study–the naive user test–are very similar to those achieved in the first part with both rankings activated and capturing from a near distance. On average, the intended object has a ranking position of 1.675. In 65% of the cases, the object ended up on the first place of the ranking; only in one case it was ranked below fifth place.

6 Discussion

Overall, our research shows that collecting landmark candidates using principles and methods of user-generated content is a feasible approach. The smartphone application works reliably, achieves a very good hit rate, and presents results within a few seconds. The naive user test showed no significant problems with the interface. Thus, a widespread use of the application seems possible in principle.

For the purpose of identifying the intended object, we introduced a ranking system similar to the one introduced by Raubal and Winter [16]. The ranking is composed of visual and semantic characteristics. The purpose of the ranking is to distinguish between salient and non-salient (or less salient) objects in the field of view and to determine the object most probably intended by the user. The results of the strictly visual approach show that it is not possible to achieve stable results by relying only on sensor data. Hence, semantic characteristics need to be integrated to measure the 'landmarkness' of geographic objects.

Previous automated methods were hampered, among others, by the required amount of data to determine landmarks [17]. A main advantage of a user-generated approach is that the main part of the selection process is done by users through the manual confirmation of results. The application only suggests likely landmark candidates and does not prescribe them. This allows for a substantial reduction of data requirements.

The results of the case study demonstrate good performance in finding the intended landmarks. The accuracy of the integrated smartphone sensors seems to satisfy the demands. Restricting GPS accuracy to at least 15 meters and the need for a user to wait for a stable position ensured a viable sensor performance in most cases. This may not always be feasible in every situation, though. Other strategies, such as setting the viewing angle to 120 degrees, also contribute to improving the results. Without these strategies the success rate would drastically deteriorate, especially in situations with reduced satellite visibility.

Using OSM as data source turns out to be a limiting factor. This is especially apparent in the rural area, where the amount of available landmark candidates decreased drastically from 838 to 24 in a comparably sized area. This leads to an increased amount of missed landmarks, and asks for additional strategies that would, for instance, allow for adding missing objects on the fly.

Finally, objects may be considered to be landmarks due to many reasons, for example, their color or their age. And what makes an object salient often does not include the whole object, but instead an eye-catching feature of the object. In the current application, we only store that a specific object is a landmark, without specifying why or which parts make it a landmark. The submission of more detailed characteristics would allow the computation of more complex references to objects, such as “the blue building on the corner with the striking shop window”. However, saving all associated information (landmark tag, justification, characteristics) on OSM would significantly increase the amount of stored data and, thus, the number of needed tags. Data volume would even further increase if photo-related information is stored as well, for example, the time of the recording, the azimuth angle to the landmark, and the location where the picture was taken. This might clutter the OSM data and be irritating to OSM users not interested in landmarks. As an alternative solution it would be possible to upload all gathered results on a dedicated publicly available server to offer a platform for the sharing of all landmark related information.

7 Conclusions

Compiling a set of landmark candidates that is of high enough quality to be uniformly useful has largely failed so far due to a mix of high demands on detailed geographic data and a lack of suitable base data. We believe that methods of user-generated content would alleviate these issues to a large extent. This paper presented first steps in that direction, namely a mobile application that allows for the in-situ collection of landmark candidates. Our work has shown the feasibility of such an approach and offered important insights into requirements and challenges that need to be dealt with to ensure a reliable data collection. However, more large-scale studies are needed to properly evaluate the usability and the scalability of the application.

Compared to automated methods to identify landmarks, the presented approach provides several advantages: First of all, it substantially reduces the amount of required data as a large part of the data filtering is done through direct input of the users. Second, in principle the application allows the landmark tagging of arbitrary geographic objects, large or small, world famous or only salient at a particular street corner. The level of detail is only limited by the completeness of the underlying OpenStreetMap data. However, this is also a major challenge for our approach. The results strongly depend on the quality of OSM data in a given area. OSM has known deficiencies of coverage in rural areas [8]. The low data density in these regions negatively affects the rate of identified landmarks and, thus, the usability of the application in such areas. This may be tackled by introducing further user-generated data collection methods.

The greatest challenge of any user-generated approach, however, is to find locals willing to contribute to such a project. To find enough users, people must be informed about and believe in the added value of landmarks and their possible uses. In addition, some mechanisms for quality checks need to be implemented, as user-generated content does not guarantee that the data is actually useful. Data correction and feedback mechanisms may be introduced towards this end.

Acknowledgments. We would like to thank Stephan Winter and Masha Ghasemi for their conceptual input in earlier stages of this project, as well as Kjartan Bjorset for his implementation of a first (map-based) prototype.

References

1. Bauer, C.: On the (in-) accuracy of GPS measures of smartphones: A study of running tracking applications. Studies **3**, 9 (2013)
2. 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, MobileHCI 2006, pp. 81–88. ACM, New York (2006)
3. Blum, J.R., Greencorn, D.G., Cooperstock, J.R.: Smartphone sensor reliability for augmented reality applications. In: Zheng, K., Li, M., Jiang, H. (eds.) MobiQuitous 2012. LNICST, vol. 120, pp. 127–138. Springer, Heidelberg (2013)
4. Crandall, D.J., Backstrom, L., Huttenlocher, D., Kleinberg, J.: Mapping the world's photos. In: Proceedings of the 18th International Conference on World Wide Web, pp. 761–770. ACM, New York (2009)
5. Duckham, M., Winter, S., Robinson, M.: Including landmarks in routing instructions. Journal of Location Based Services **4**(1), 28–52 (2010)
6. Ghasemi, M., Richter, K.F., Winter, S.: Landmarks in OSM, State of the Map 2011 (2011)
7. Goodchild, M.F.: Citizens as sensors: The world of volunteered geography. GeoJournal **69**(4), 211–221 (2007)
8. Haklay, M.: How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B **37**(4), 682 (2010)
9. Lomas, N.: Android still growing market share by winning first time smartphone users. http://techcrunch.com/2014/05/06/android-still-growing-market-share-by-winning-first-time-smartphoneusers/ (retrieved on: July 23, 2014)
10. Lynch, K.: The Image of the City. The MIT Press, Cambridge (1960)
11. 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)
12. 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, pp. 209–218. Hannover, Germany (2006)
13. Nothegger, C., Winter, S., Raubal, M.: Selection of salient features for route directions. Spatial Cognition & Computation **4**(2), 113–136 (2004)
14. Papadopoulos, S., Zigkolis, C., Kompatsiaris, Y., Vakali, A.: Cluster-based landmark and event detection on tagged photo collections. IEEE Multimedia **18**(1), 52–63 (2010)
15. Presson, C.C., Montello, D.R.: Points of reference in spatial cognition: Stalking the elusive landmark. British Journal of Developmental Psychology **6**, 378–381 (1988)
16. Raubal, M., Winter, S.: Enriching wayfinding instructions with local landmarks. In: Egenhofer, M., Mark, D.M. (eds.) GIScience 2002. LNCS, vol. 2478, pp. 243–259. Springer, Heidelberg (2002)
17. Richter, K.F.: Prospects and challenges of landmarks in navigation services. In: Raubal, M., Mark, D.M., Frank, A.U. (eds.) Cognitive and Linguistic Aspects of Geographic Space - New Perspectives on Geographic Information Research. Lecture Notes in Geoinformation and Cartography, pp. 83–97. Springer, Heidelberg (2013)
18. Richter, K.F., Winter, S.: Harvesting user-generated content for semantic spatial information: the case of landmarks in OpenStreetMap. In: Proceedings of the Spatial Sciences & Surveying Biennial Conference, Wellington, NZ (2011)

19. Richter, K.F., Winter, S.: Landmarks – GIScience for Intelligent Services. Springer, Cham (2014)
20. Siegel, A.W., White, S.H.: The development of spatial representations of large-scale environments. In: Reese, H. (ed.) Advances in Child Development and Behaviour, pp. 9–55. Academic Press, New York (1975)
21. Sorrows, M.E., Hirtle, S.C.: The nature of landmarks for real and electronic spaces. In: Freksa, C., Mark, D.M. (eds.) COSIT 1999. LNCS, vol. 1661, pp. 37–50. Springer, Heidelberg (1999)
22. Tezuka, T., Tanaka, K.: Landmark extraction: a web mining approach. In: Cohn, A.G., Mark, D.M. (eds.) COSIT 2005. LNCS, vol. 3693, pp. 379–396. Springer, Heidelberg (2005)
23. 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, pp. 90–104 (2004)
24. 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, pp. 278–285. ACM, New York (1999)