

Pharos: Improving Navigation Instructions on Smartwatches by Including Global Landmarks

Nina Wenig, Dirk Wenig,
Steffen Ernst, Rainer Malaka
Digital Media Lab, TZI
University of Bremen
{nwenig, dwenig, malaka}@tzi.de
steffen.ernst@uni-bremen.de

Brent Hecht
People, Space, and Algorithms
(PSA) Computing Research
Group
Northwestern University
bhecht@northwestern.edu

Johannes Schöning
Human-Computer Interaction
University of Bremen
schoening@uni-bremen.de

ABSTRACT

Landmark-based navigation systems have proven benefits relative to traditional turn-by-turn systems that use street names and distances. However, one obstacle to the implementation of landmark-based navigation systems is the complex challenge of selecting salient local landmarks at each decision point for each user. In this paper, we present *Pharos*, a novel system that extends turn-by-turn navigation instructions using a single *global landmark* (e.g. the Eiffel Tower, the Burj Khalifa, municipal TV towers) rather than multiple, hard-to-select local landmarks. We first show that our approach is feasible in a large number of cities around the world through the use of computer vision to select global landmarks. We then present the results of a study demonstrating that by including global landmarks in navigation instructions, users navigate more confidently and build a more accurate mental map of the navigated area than using turn-by-turn instructions.

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces — *input devices and strategies, interaction styles*

Author Keywords

Global Landmarks; Landmark-based Navigation; Computer Vision; Smartwatches; Pedestrian Navigation

INTRODUCTION & MOTIVATION

Research across many fields has robustly established that landmarks are an essential means by which humans navigate through their environments [11, 15, 16, 46]. People of all ages use landmarks in this fashion, and landmarks have been described as “the key to the ability to orient oneself and to navigate in an environment.” [48]. This literature has motivated researchers to develop a series of landmark-based navigation technologies that have been shown to outperform traditional



Figure 1: The Pharos navigation approach: Using Google Street View imagery, the visibility of global landmarks (here a municipal TV tower) is determined and then included in the turn-by-turn pedestrian navigation instructions for smartwatches.

turn-by-turn instructions in a number of studies, particularly in the case of pedestrian navigation [17, 43, 52, 53].

However, despite the success of these research prototypes, well-known navigation technologies largely do not incorporate landmarks. This is in part due to the substantial implementation challenges associated with automating landmark-based navigation. For instance, landmarks that are salient for each individual user must be selected [55, 58], the utility of specific landmarks for navigation is often gender- and language-dependent [55], and assessing the visibility of each landmark is difficult [38, 59].

In this paper, we introduce the *Pharos* pedestrian navigation approach, which seeks to maintain the benefits of landmark-based navigation while substantially increasing the tractability of landmark-based navigation systems. The key to the *Pharos* approach is to reduce the challenges associated with landmark-based navigation by utilizing *global landmarks* [24, 30], a class of landmarks that have not yet been considered in the navigation technology literature. The landmarks that are traditionally utilized are *local landmarks* [58] and are always located along route segments, particularly at key decision points. Global landmarks, on the other hand, are landmarks that can be seen from significant distances and can be located far from the user’s route. Typical global landmarks are tall buildings (e.g. the Eiffel Tower, the Burj Khalifa, municipal TV towers), but mountains (e.g. Corcovado mountain in Rio de Janeiro, or the when there are mountain ranges in the background as in Denver, Seattle, Bishkek, Cape Town) or whole downtown areas (e.g. the skyline of New York) can also serve as global landmarks.

In addition to addressing the limitations of local landmarks discussed above (e.g. language/gender dependency), global landmarks also have several other important properties that are beneficial with respect to navigation tasks. First, global landmarks are constant, meaning that users can update their position and orientation relative to a single landmark rather than jumping from landmark to landmark. Secondly, because the distance to the landmark is not an important component of global landmark-based navigation, global landmarks can be used as a sort of compass for orientation [49]. Finally, while landmark-enriched navigation instructions require that the device’s positioning system is able to locate the user with high accuracy for local landmarks, this is not true for global landmarks. A positioning error of a few meters might heavily influence the visibility of a local landmark (e.g. a shop or a street sign), but the direction of a distant global landmark remains the same.

This paper demonstrates that the *Pharos* approach is both (1) feasible and (2) has benefits compared to traditional turn-by-turn (i.e. no landmark) instructions. We establish feasibility by using computer vision (CV) and machine learning (ML) techniques to show that global landmark visibility is sufficiently extensive within cities and, critically, automatically detectable at scale with good accuracy. We also show that our approach works with publicly-available images and allows for pre-computation server-side, meaning that no further computation on a mobile or wearable device is needed.

We develop a prototype of the *Pharos* approach and use it to establish the effectiveness of global landmark-based instructions through a field-based user study. Focusing on the use case of pedestrian navigation using smartwatches, we find that global landmark-enriched instructions outperform state-of-the-art turn-by-turn instructions along several key metrics. Specifically, we find that while both approaches had roughly similar performance in terms of navigation speed and number of errors, as hypothesized, *Pharos* outperformed turn-by-turn with respect to navigation confidence and the accuracy of the user’s resultant mental map of the area.

To summarize, the contribution of this paper is four-fold:

- We introduce the *Pharos* navigation approach, which makes landmark-based navigation much more feasible by using global landmarks (relative to traditional landmark-based approaches that use local landmarks).
- We show that global landmarks are sufficiently broadly visible for pedestrian navigation tasks and that the locations at which global landmarks are visible can be computed at scale from geotagged imagery using computer vision and machine learning. This information can be computed beforehand so that no processing on the wearable or mobile device is needed.
- We develop a pipeline to generate navigation instructions for pedestrians on smartwatches.
- We demonstrate through a user study that the *Pharos* global landmark-based navigation approach outperforms the current-state-of-the-art turn-by-turn instructions in important navigation metrics.

Below, we first introduce related work. We then describe the process by which we determine the visibility of global landmarks and create visibility maps. We next discuss the *Pharos* approach of integrating global landmarks into navigation instructions and demonstrate the benefits of our approach based on the results of a user study. Finally, we conclude with a discussion of the *Pharos* approach, landmark-based pedestrian navigation in general, and open problems in this area.

RELATED WORK

This work draws from and builds on research in the domains of (1) pedestrian navigation and mobile guides, (2) landmark-based navigation in particular, and (3) landmark detection. Below, we discuss each of these areas in turn.

Mobile Guides and Pedestrian Navigation

Research on pedestrian navigation dates back almost two decades. From the beginning of this work, landmarks have played a role in guiding the user. For example, the GUIDE project [7] aimed at integrating landmarks in textual navigation instructions. Malaka and Zipf [34] used a 3D city model to create navigation instructions not only incorporating the visibility but also the look of landmarks (e.g. “turn right after the red building”) for pre-selected situations. Similarly, the LOL@ tourist guide for mobile devices [40] enhanced routing information with references to landmarks. The PhotoMap [45] uses images of taken with a GPS-enhanced mobile phone as background maps for on-the-fly navigation tasks.

While one of the earliest mobile guides, the DeepMap system [34], used a wrist-mounted display, wearable devices have become more relevant in research in the last five years. Indeed, in the context of pedestrian navigation, wearable devices do have an important benefit over their mobile counterparts as the user’s hands can remain mostly free. Wenig et al. [56] introduced StripeMaps, a cartographic approach for indoor navigation with smartwatches, which transforms 2D maps with route information into a 1D stripe. McGookin and Brewster [35] investigated unidirectional navigation for runners by designing

a navigation system called RunNav, which could also be used on a smartwatch. RunNav does not provide explicit routes, but rather a high-level overview to inform runners of areas that are good and bad places to run.

Even though research has explored how other modalities can be used to provide navigation instructions (e.g. auditory instructions [22] or haptic clues for mobile [42] and wearable devices [29, 39]), map-based navigation for mobile devices and turn-by-turn instructions for wearable devices have been established as de facto standards. For example, the current version of Google Maps for Android Wear smartwatches uses turns, street names, and distances to provide the user with instructions such as: “After 20 m turn left into Denver Road”. Importantly, current navigation systems and applications typically do not include landmarks in their instructions.

Landmark-based Navigation

There is already a large corpus of related work exploring the benefits of landmark-based navigation for pedestrians. While the term *landmark* was originally used to differentiate features with outstanding characteristics [33], the meaning of the term has changed over time and is now used more generically to describe well-known places [15]. According to Sorrows and Hirtle [48], “it is useful to understand landmarks in a way that supersedes knowledge in the environment”.

Landmark-based navigation instructions have been investigated in depth from the perspective of spatial cognition and cognitive psychology. Tom and Denis [52] showed that for guiding pedestrians, route information referring to streets is less effective (regarding the number of stops, instruction checks and time) than route information referring to local landmarks. In another experiment by Tom and Denis [53], the participants processed landmark-based instructions faster than street-based instructions and also remembered the route better with landmark information than with street information. Additionally, Ross et al. [43] showed that adding landmarks to basic pedestrian navigation instructions (turn information and street names) results in less errors and a higher user confidence. In contrast to our work, neither Tom and Denis [52, 53] nor Ross et al. [43] relied only on global landmarks in the navigation instructions.

From a more applied perspective, Wither et al. [59] showed that people can navigate solely using landmarks highlighted in panoramic imagery. Even ‘in the wild’ in natural environments, as shown by Snowdon and Kray [47], there are types of landmarks which are feasible to be used in mobile navigation systems. Recently, Bauer et al. [2] have found evidence, that indoors pedestrian navigation instructions on mobile devices should only depict a single prominent landmark (instead of four) for high navigation efficiency. In contrast, the Pharos approach is primarily targeted at navigation in urban environments. While our work is related to the work of Wither et al. [59], it is landmark-centric and instead of local landmarks the Pharos approach relies on global ones.

Landmarks do also play a role in designing pedestrian navigation aids for people with disabilities. For example, landmarks can be used to support low-vision and blind people in loca-

ting bus stops [18] or to help mobility impaired users with navigating [19]. In addition, landmarks can be used for other purposes in location-based services. For example, Kray and Kortuem [25] interactively determined the user’s positions based on the visibility of nearby landmarks. Lu et al. [32] generated traveling routes from geo-tagged photographs.

In general, landmarks have to be chosen carefully as their selection is highly important for the resulting quality of navigation instructions [36] (e.g. the personal salience of a landmark could be effected by language or gender [55]). More specifically, Sorrows and Hirtle [48] examined landmarks in real and electronic spaces and classified landmarks in terms of visual, cognitive and structural dimensions. Global landmarks have to be outstanding in all dimensions. Steck and Mallot [49] investigated the role of global and local landmarks in virtual environment navigation. They found out that in virtual environments both local and global landmarks are used for wayfinding, but different participants rely on different strategies; some of the participants used only local landmarks while others only used global ones. This results in a key question for our work, as with Pharos we aim at improving pedestrian navigation instructions with only a single global landmark rather than multiple local landmarks.

Landmark Detection

Detecting landmarks is a crucial step in the pipeline of all location-based services relying on landmarks. While researchers have often opted to perform the selection of landmarks by hand [34, 36] or using crowd-based [18] approaches, commercially successful applications require a solution that works on a global scale across a large set of users. As such, some automated approaches have been developed that make use of large databases of geographic features (e.g. OpenStreetMap [OSM]) [10, 41]. For example, Raubal and Winter [41] explored landmark-enriched wayfinding instructions by automatically extracting local landmarks from a spatial dataset based on a formal measure for landmark saliency. More recently, Dräger and Koller [10] generated landmark-based navigation instructions for car navigation systems using OSM data. They used proximity to certain geographic features as a criterion to select landmarks. For prominent landmarks, the proximity can be used to estimate if a geotagged image shows a particular landmark [38].

As global landmarks can be seen from longer distances and therefore the mere location information is not sufficient, visibility tests are always necessary. Visual landmark detection is related to the problems of location recognition, object recognition and image retrieval. For computer vision it is common to use simple image features, such as edges or corners in an image. For example, SIFT [31], SURF [3] and ORB [44] are popular image feature descriptors, which can be used together with machine learning approaches, e.g. support vector machines [5] to predict whether an image contains a particular landmark. Most recently convolutional neural networks [26] have become very popular; they work directly on the pixel data.

Approaches for detecting different landmarks are often based on these computer vision algorithms in combination with

the location of the landmark [6] or other contextual information [28]. Zheng et al. [60] built a large database from geotagged photos to detect the most popular landmarks around the world. As landmarks are visible from different positions, the visibility from different angles has to be computed [23]. Wither et al. [59] automatically created landmark-based navigation instructions by detecting salient landmarks in panoramic street imagery by using additional data, such as LiDAR information and text understanding. Such 3D information can often be used to enhance computer vision algorithms, e.g. [21]. Wakamiya et al. [55] combined this information also with social data (Twitter or Foursquare) to determine the best local landmarks in an area.

PHAROS

In this section, we describe the Pharos global landmark-based navigation approach. Pharos is named after the famous lighthouse in Alexandria that was one of the “seven wonders” of the ancient world and was used for centuries as a navigation landmark for ships. Below, we first report on a comparison of methods to determine the visibility of global landmarks. Secondly, we report on how we include global landmarks in the navigation instructions.

Determining Visibility for Global Landmarks

We determined the visibility of selected global landmarks to investigate whether they are (a) sufficient visible for pedestrian navigation tasks and (b) that this visibility can be detected at scale from geotagged imagery. The first step for this, is to compute their visibility for a given region of interest (ROI), e.g. the inner city or a district.

Global landmarks are diverse and range from mountains and large buildings to entire downtowns. The visibility of global landmarks is dependent on their *prominence* in a ROI (rather than their absolute height). Prominence in a topographical context describes “the height of a mountain...summit by the vertical distance between it and the lowest contour line encircling it but containing no higher summit within it”[1].

Although global landmarks are often visible from large parts of a given ROI, they are not unconditionally visible from areas around the landmark as other buildings/structures might block line of sight. To create detailed visibility maps for global landmarks, we evaluated different computer vision and machine learning techniques to determine their visibility for a certain ROI.

Techniques

We compared the most common image features used in object detection tasks in combination with machine learning on three different landmarks: The Eiffel Tower in Paris (324 m, France), the Petronas Towers in Kuala Lumpur (452 m, Malaysia) and the Burj Khalifa in Dubai (828 m, United Arab Emirates). All three are very prominent landmarks, but have very different characteristics (e.g. height, appearance, symmetry and structure). For each landmark, we manually collected as training images 150 (not necessarily geotagged) images containing the landmark using Google image search and Flickr. We made sure that the images for each landmark were taken in different

	Sliding Window	Selective Search
SURF+SVM	75.39	73.47
SIFT+SVM	69.26	60.70
ORB+SVM	54.47	52.73
CNN	84.13	77.18

Table 1: Comparison of methods to determine the visibility of global landmarks. The $f1$ -score is the harmonic mean of precision and recall. Our refined convolutional neural network outperforms the other approaches. In general, the sliding window approach works better than selective search.

lighting and weather conditions as well as from multiple perspectives and distances. In addition, as negative samples, we collected 450 images from Google Street View (GSV), which do not contain the landmarks but instead typical content in the ROI around the landmarks.

With this training data we compared SIFT [31], SURF [3] and ORB [44] as image features in combination with a Support Vector Machine [5] and an approach using a convolutional neural network [26] (i.e. deep learning). For deep learning, it is common to use already-trained neural networks and refine them by training again on task-specific data to improve accuracy. Therefore, we used the convolutional neural network which was trained on the ImageNet data [8] and trained it in a second step with our dataset. For this, we used the DeCaf library [9].

We used a simple sliding window approach, in which we slide a window with the size $256 \times 256 px$ in $64 px$ steps over the image to create sub-images. Then, we tested for each sub-image whether it contains the landmark or not. We did this for different resolutions of the image by using an image pyramid to detect the landmark, even if the landmark is far away and very small or when it is larger than $256 \times 256 px$ in the GSV image. We tested the sliding window against the selective search algorithm by Uijlings et al. [54] to find the landmarks in sub-images.

Classifier Evaluation

We used a separate data set with 100 manually collected GSV images for each landmark to evaluate the classifiers. We categorized an image as *containing a particular landmark* when at least one sub-image (based on the sliding window or selective search approach) contains the landmark. The results of the comparison can be seen in Table 1; the $f1$ -score is the harmonic mean of precision and recall. We calculated the scores across all three landmarks. The evaluation shows that we achieve the best results with the convolutional neural network combined with a simple sliding window approach. All results are similar for each landmark ($f1$ -score Eiffel Tower: 77.3, Petronas Towers: 79.5, Burj Khalifa: 94.5). The results for the Burj Khalifa are particularly good due to its prominence in Dubai’s skyline. Overall, our results suggest that the visibility of global landmarks can robustly be detected at scale from geotagged images.

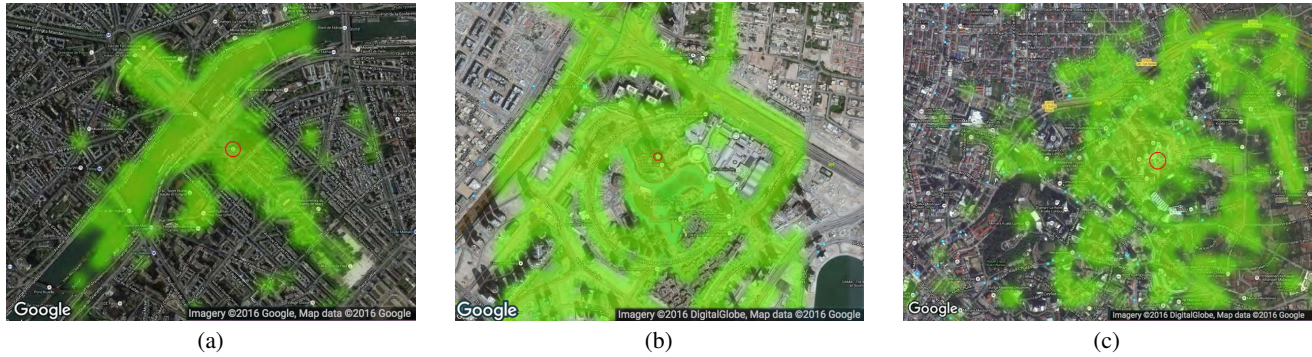


Figure 2: Visibility maps, automatically produced based on a CNN classifying Google Street View (GSV) images, for the three selected global landmarks: a) Eiffel Tower b) Burj Khalifa and c) Petronas Towers. Green spots indicate GSV images that contain the global landmark. Aerial images from Google Maps were used as base map (©Google 2016).

Visibility of Global Landmarks

As the results of the technical evaluation showed that the convolutional neural network works best, we used it to create visibility maps around the three landmarks using GSV images. In a radius of 2 km around the landmark, we downloaded a GSV image at regular 100m intervals. This resulted in around 1600 images for every ROI. For some points or areas GSV is not available. After excluding these areas, we computed the visibility of the landmark in all the remaining images. Figure 2 shows the resulting visibility maps for all three global landmarks. The green dots indicate that the landmark is visible in the GSV image. For the three selected global landmarks, the visibility maps show that these landmarks are likely sufficiently visible for pedestrian navigation tasks.

As the three selected global landmarks differ in shape, size and contour line of the surroundings, we conclude that determining the visibility of global landmarks based on public available imagery is feasible from the technical point of view. The visibility maps can be created on map and navigation servers beforehand so that no processing on the wearable or mobile device is needed. The visibility information than can be used to include global landmarks in navigation instructions.

Including Global Landmarks in Smartwatch Navigation Instructions

Turn-by-turn navigation instructions inform the user about upcoming turns, usually combined with a notification (e.g. vibration). Global landmarks can be included in these instructions to confirm that the user has taken the correct turn and is still on the correct route. For Pharos, we developed a pipeline to include the visibility information of the global landmarks in turn-by-turn based pedestrian navigation instructions. While our approach could be applied to both car and pedestrian navigation systems, we opted to focus on the integration of global landmarks into instructions for pedestrian navigation. More specifically, we focused on smartwatch-based pedestrian navigation as smartwatches have important benefits relative to other mobile devices (e.g. smartphones) for navigation, e.g. they remove the need to constantly take a device out of one’s purse or pocket [56] (Note: we expect that our findings will

generalize to a smartphone context, although further research is necessary to conform this hypothesis).

Current navigation systems for smartwatches, e.g. Google Maps for Android, use turn-by-turn-based instructions. They primarily rely on simple arrows showing the direction the user has to take at the next turning point (not at other decision points where the user does not have to turn), combined with the name of the street on which the user has to turn. Integrating global landmarks into such navigation instructions is not a trivial task. In general, situations in which navigation instructions can only rely on the global landmark are rare (e.g. “Head towards the landmark” when the route leads the user directly towards the landmark). Therefore, for Pharos, we aimed at enriching turn-by-turn navigation instruction with direct and indirect hints related to global landmarks. Direct hints precisely include the landmark in the navigation instruction, while indirect hints can be seen as additional information on the landmark’s position relative to the user at decision points.

Textual instructions, especially on very small screens, need to be both short and understandable. For Pharos, we identified four different types of textual navigation instructions including a global landmark: the route heads towards towards or away from the global landmark, the global landmark is on the left or on the right, the global landmark is in a direction other than a cardinal direction (e.g. soft left), and the global landmark is not visible at the turning point.

The most simple situations are situations when the user, after a turn, walks straight towards the landmarks or walks away from the landmark. For such instructions, it is straightforward to add a direction to turn-by-turn instructions:

“Head towards the landmark” or
“Head away from the landmark”

Similar are situations in which the landmark is on the left or on the right of the user after the turn, but cannot be directly included in the instruction. For such situations, we add indirect hints about the landmark’s location, e.g.:

“The landmark will be on your right” or
“The landmark will be on your left”

More difficult are situations in which the landmark is neither in front or behind the user nor on the left or on the right (e.g. 45 degrees or 135 degrees to the user’s walking direction). Here, the English language struggles to unambiguously describe such situations without referring to angles. To address these situations, we use instructions such as:

“The landmark will be in front of you to your right” or
“The landmark will be behind you on your left”

Additionally, there will be situations in which the landmark is not visible at the turning point. For these situations, we propose to include landmark information in the following form:

*“At the end of the street,
the landmark will be in front of you”*

From the technical point of view, all instructions can be easily generated using the visibility maps. For each turning point the angle between the direction of the following path segment (towards the next turning point) and the direction towards the landmark indicates the kind of instruction that should be used. For the last type of instruction, the visibility during the path segment also has to be considered. Whenever the landmark is not visible, simple turn-by-turn instruction can be used.

USER STUDY

We performed a user study to evaluate the benefits of the Pharos approach and to explore how the inclusion of global landmarks changes the navigation experience. The study was focused on standard navigation evaluation metrics (time to reach destination, number of errors made) and the users’ confidence that they were on the correct route. Confidence is not only an important aspect for the usability of a system in general [4], it is particularly important for navigation systems (and its instructions) as the user is usually navigating in an unfamiliar environment with potential safety risks (e.g., other traffic participants). In addition, we also measured how well users built up spatial knowledge of the route using cognitive maps. Spatial knowledge supports users in performing the same or a similar navigation task without technological assistance the next time they are in the area and also aids them in identifying possible shortcuts [15].

We compared the following two conditions in the user study, illustrated in Figure 3.

- 1) Turn-by-turn navigation instructions (TBT) as a baseline.
- 2) Turn-by-turn navigation instructions enriched with global landmarks through the Pharos approach (PHA).

In both conditions, the instructions are based heavily on Google Maps navigation instructions for Google Android Wear smartwatches. We used the same instructions as generated by Google Maps for Android Wear as well as the exact “look and feel” in the baseline (TBT) but enriched these instructions with the global landmarks in the PHA condition.

In the current version of Google Maps for Android Wear, the upcoming turn is shown with the remaining distance in

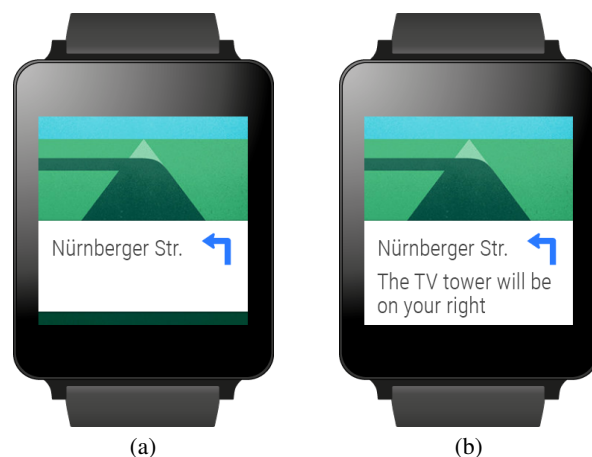


Figure 3: The two different conditions compared in the user study. Turn-by-turn navigation (TBT) on the left and the turn-by-turn navigation instructions including global landmarks following the Pharos approach on the right (PHA).

meters (in steps of 10 m). For the user study, we decided against this approach for two reasons. First, in pre-tests, the distance measures and the position of the notifications were insufficiently accurate as they occurred often with an offset of around 10 m. Second, pedestrian navigation instructions should neither require nor encourage the users to constantly check the system, distracting them from their primary task of walking.

To achieve optimal comparability, in the baseline condition (TBT) as well in the PHA condition we followed a wizard-of-oz study approach with instructions manually triggered by an experimenter. Participants were notified of new instructions via vibration of the smartwatch.

Participants & Apparatus

The study was conducted in a residential district of Bremen (Germany), a mid-sized city in northern Europe. As a global landmark, we used a telecommunication and television tower (referred as TV tower in the rest of the paper), which is about 235 meters high. It can be seen from significant distances (see Introduction & Motivation). The ROI of the study features wide and narrow streets with mid-sized row houses (up to three or four stories) such that the global landmark is not visible at all times.

For the user study, we selected two routes. Each of the routes had a very similar lengths (route 1: 1.12 km, route 2: 1.17 km) and featured an equal number of turns (seven turns plus the start and end point). The maximal length between two turns on both routes was 0.3 km. The global landmark was similarly visible in both routes. Figure 4 shows both routes on the visibility map for the TV tower.

The turn-by-turn-based navigation instructions were created as follows. First, we used the multi-destination feature of Google Maps to create the route. Secondly, we followed the route using Google Maps for Android Wear and at all turning

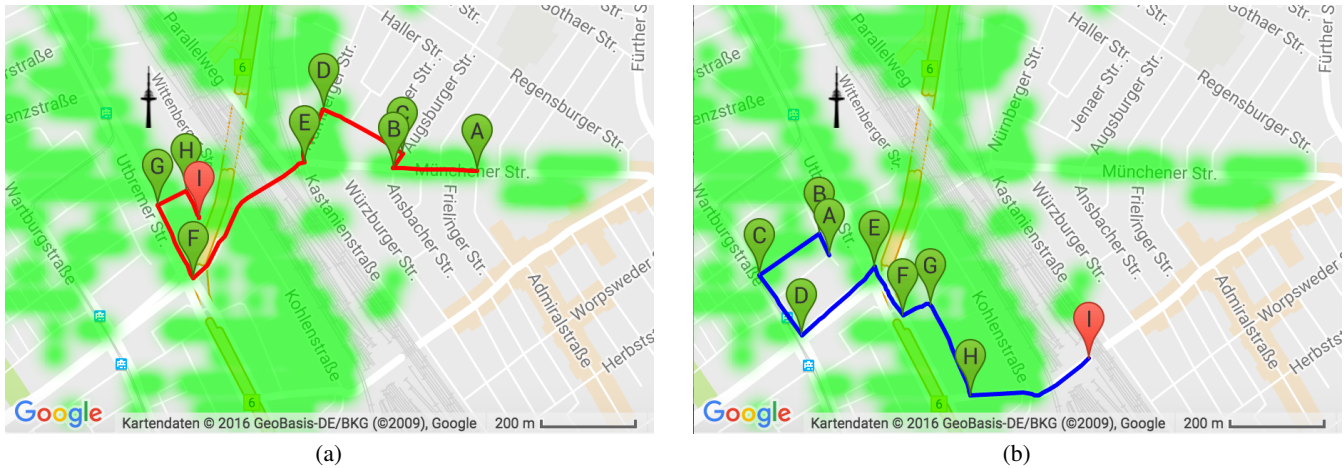


Figure 4: The two routes used in the user study. Green dots indicate the visibility of the TV tower based on GSV images. Route 1 (a, red) has a length of 1.12 km and route 2 (b, blue) is 1.17 km long. For both routes, A is the starting point and I is the end point. Google Maps was used as base map (©Google 2016).

points we took a screenshot of the smartwatch (including the background image showing the turn in addition to the arrow, see Figure 3). We used these screenshots to create the instructions and (for both conditions) removed the countdown distance measures (see above). For the Pharos condition, we enriched the instructions with information about the location of the TV tower (as described in the previous section). At one of the turning points and the following route segment of route 1, the landmark was not visible at all. At this point we used the baseline TBT instruction also in the PHA condition.

The instructions for the next turning point were shown immediately before the turn at the exact same locations (manually triggered by the experimenter accompanied with a vibration notification of the new instruction). As a result, on the next route segment (after the turn) the instruction would be outdated. To exclude irritations because of outdated instructions, for the path segments between the turning points we slightly adapted the navigation instructions for both conditions. These navigation instructions were based on the instructions for the previous turning point, but the turning arrow was replaced with a straight ahead arrow. Furthermore, we changed the tense of the textual instruction from future tense to present tense (e.g. “The TV tower is on your right” instead of “The TV tower will be on your right”). In contrast to the turning point instructions, these in-between instructions were not accompanied by vibration.

In accordance with our wizard-of-oz study design, we built a simple Android Wear smartwatch app to present the navigation instructions. An Android companion app we built for smartphones (showing all the turn-by-turn-based navigation instructions in a list view) allowed the experimenter to select and send the instructions to the smartwatch. The experimenter triggered the instructions always at the same predefined spots, approximately 5 m before the turning point. Furthermore, the app allowed the experimenter to count how often the participants looked at the smartwatch.

We recruited 12 participants (4 females, 8 males) with an average age of 27.1 ($SD=2.5$). Most of our participants had very limited familiarity with the study area, although some had traveled along minor segments of the routes before. All participants own a smartphone and two of the participants regularly wear a smartwatch, but only one of them had navigated via smartwatch before. Ten of the participants use a smartphone for navigation purposes on a regular basis.

The user study was conducted with a LG G Watch. All participants performed the test in both conditions (within-subject design). The orders of the two conditions (TBT and PHA) as well as the two routes (route 1 and route 2) were counterbalanced.

Task & Procedure

The participants were introduced to the experiment and told to follow the two different routes, one after the other. We explained the navigation task but did not mention the role of the TV tower. The participants did not have to select a route or target destination on a mobile device or on the smartwatch. For both conditions, we oriented the participants in the direction of the first movement and then started the navigation task with the first instruction.

As participants walked the route, the experimenter followed a few meters behind, selecting the navigation instructions on the companion app, collecting timing information, counting the number of times the user looked at her watch, and assessing the number of navigation errors. An error was assessed when a participant took a wrong turn without noticing their mistake after 10 meters (after this point, the experimenter would then guide the participant back on the route). While they were navigating, we asked the participants to rate their confidence (i.e. whether or not they believed they were on the correct route) on a seven-point scale. We performed this confidence assessment three times: after the start instruction, in the middle of the route, and at the end of the route.

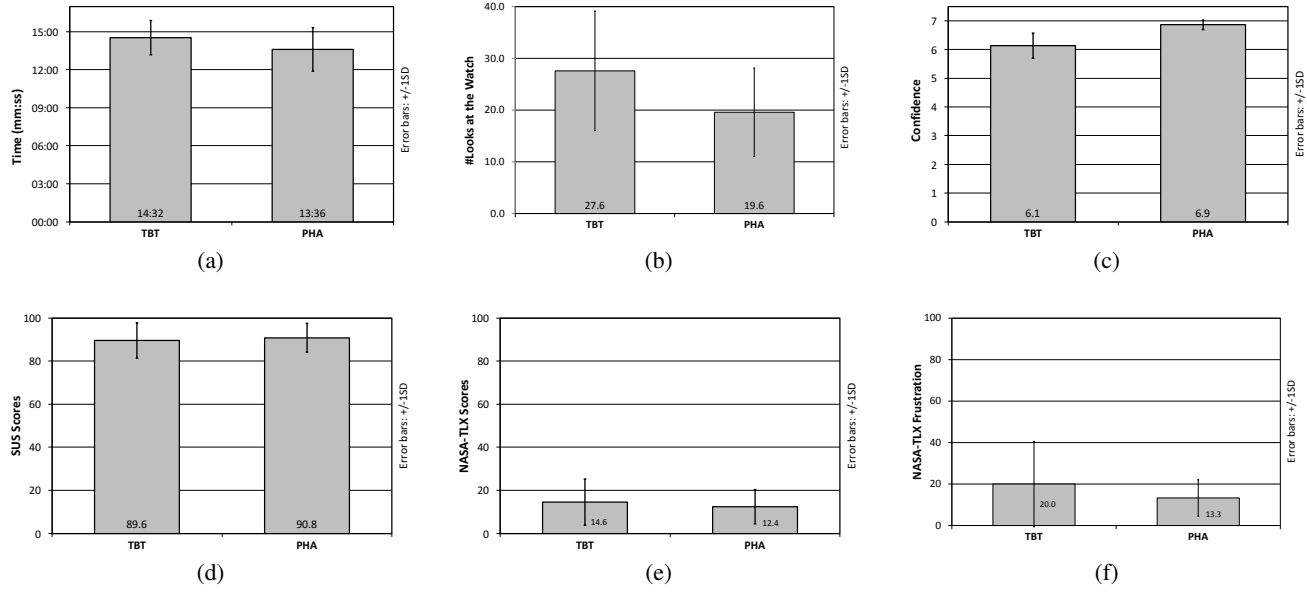


Figure 5: The results for the baseline (TBT) and Pharos (PHA) conditions for (a) the time the participants needed to complete the routes with both interfaces, (b) how often they looked at the smartwatch, and (c) how confident they felt during navigation. The second row shows the results from (d) the System Usability Scale, (e) the overall NASA-TLX Score, and (f) the results for the NASA-TLX sub-scale of frustration.

After each condition, participants were asked to draw a cognitive map of the route they were instructed to follow. In the PHA condition, they were also asked to include the location of the TV tower. The concept of cognitive maps was first introduced by Tolman [51] and later adapted and extended to the domain of spatial computing, where cognitive maps are a “mental representation of people’s perception of the real world” [14, 13]. They provide a representation of the spatial knowledge of a user [37]. The goal was to measure if the users gained more spatial knowledge in the PHA condition.

We also used the NASA-TLX [20] to measure the perceived workload and the System Usability Scale (SUS) [4] to measure the perceived usability. All questionnaires were filled out for both conditions after the participants had drawn the cognitive map. The total time taken by each participant for the whole study was about 60 minutes. Participants were encouraged to think aloud and to ask questions if necessary. Noteworthy incidents were recorded in writing. A semi-structured interview was conducted with each of the participants after finishing both routes.

Results & Analysis

All participants were able to complete all the tasks. Figure 5 summarizes the results of the user study.

On average, the participants took 14 minutes and 4 seconds per route (Route 1: 14:39 *m*, Route 2: 13:29 *m*). The fastest participant needed 11 minutes and 15 seconds while the slowest one needed 17 minutes and 50 seconds. Participants made a maximum of two errors per route. Most of the errors happened on route 1 ($M=0.92$) due to a missing sidewalk at one

turning point and only one participant made an error on route 2 ($M=0.08$). The participants were slightly faster in the PHA condition compared to TBT (TBT: $M=14:32$ *m*, $SD=01:21$ *m*; PHA: $M=13:36$ *m*, $SD=01:43$ *m*), while they committed almost the same number of errors in both conditions (TBT: $M=0.5$, $SD=0.65$; PHA: $M=0.5$, $SD=0.5$). Statistical analysis did not reveal significant differences in either the speed or error rate measures.

Regarding the perceived usability, the SUS scores were high for both the TBT condition ($M=89.6$, $SD=8.2$) and the PHA condition ($M=90.8$, $SD=6.8$), see Figure 5d. That means that the participants had no serious usability problems in both conditions influencing the results. Regarding the perceived taskload, the NASA-TLX values are low and almost the same for both the TBT condition ($M=14.6$, $SD=10.7$) and the PHA condition ($M=12.4$, $SD=8.0$), the sub-scale of frustration (“How insecure, discouraged, irritated, stressed, and annoyed were you?”) differs (TBT: $M=20.0$, $SD=20.6$; PHA: $M=13.3$, $SD=8.8$), see Figures 5e and 5f. This means that both conditions evoke a low workload with a slightly higher frustration for the TBT condition. However, statistical analysis did neither reveal a significant difference for the sub-scale of frustration, nor for the NASA-TLX overall values and the SUS scores.

The confidence ratings resulted in three values per participant per route. On average, the confidence was higher in the Pharos condition ($M=6.9$, $SD=0.1$) than in the TBT condition ($M=6.1$, $SD=0.4$), see Figure 5c. A paired t-Test revealed that the difference is statistically significant ($t(11) = -5.370$, $p < 0.001$) with a large effect size $r=0.81$ (Cohen’s

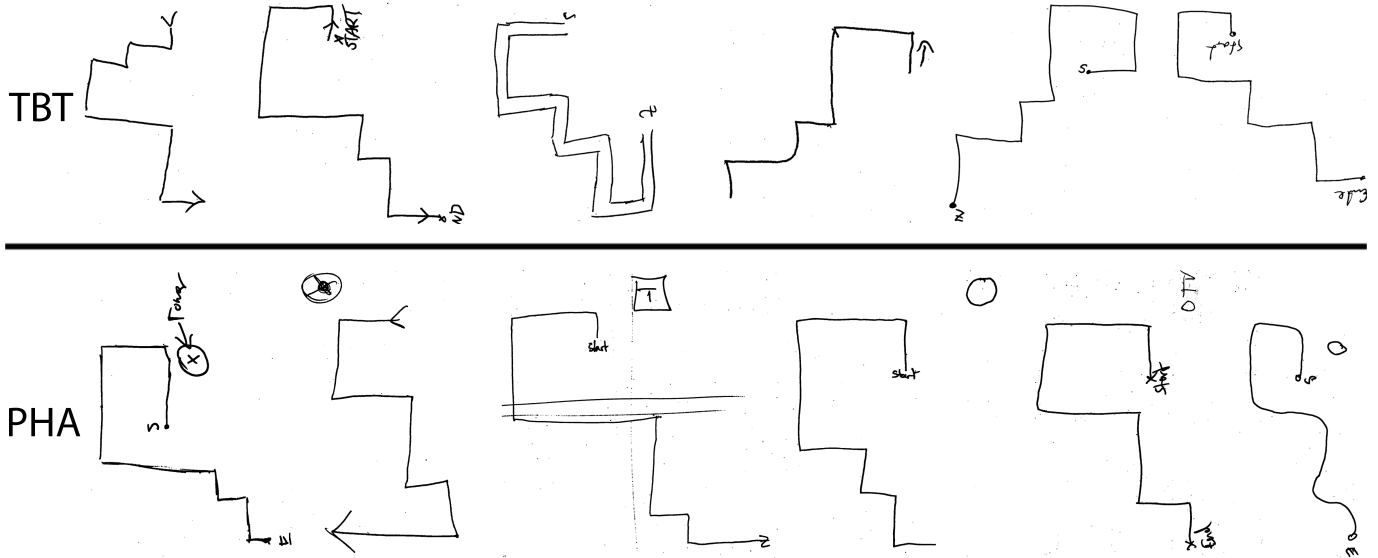


Figure 6: Mental map route sketches for route 2 from all participants in the baseline (TBT) and Pharos (PHA) conditions (rotated and scaled for comparability). First row TBT and second row PHA condition.

$d=7.243$). The reduced confidence in the TBT instructions is also apparent in the number of looks at the navigation instructions on the smartwatch. The number of looks is substantially lower for the PHA condition ($M=19.6$, $SD=8.6$) than for the TBT condition ($M=27.6$, $SD=11.6$), see Figure 5b. A paired t-test revealed a statistically significant difference between the conditions ($t(11)=2.339$, $p=0.039$), with a medium effect size $r=0.36$ (Cohen’s $d=0.74$). In other words, with the Pharos approach of integrating global landmarks into the navigation instructions, participants were not only more confident that they are on the correct route, they also looked at their smartwatch less often than was the case with baseline turn-by-turn instructions. We note that looking at one’s smartphone is often substantially more effortful than doing so using a smartwatch, so we expect that the benefits of Pharos may be greater in this respect for smartphone-based navigation.

To assess and analyze the representation of the spatial knowledge of a user while performing the task we used cognitive maps [37]. The cognitive maps drawn by the participants after each condition were digitized using QGIS¹. We performed a bi-dimensional regression following the procedure as described by Friedman [12]. In general, bi-dimensional regression [50] “requires an equal number of points between the configurations to be related” [14]. However, this was not true across participants’ sketches of the whole route. Therefore, we performed a bi-dimensional regression only on the turning points, excluding all of the segments between the turning points. A Wilcoxon signed-rank test on the retrieved correlation values of the TBT ($M=0.51$, $STD=0.17$) and PHA condition ($M=0.63$, $STD=0.11$) showed a significant difference between these two conditions ($Z = -3.04$, $p=0.043$). That means that participants could better remember the route with

the Pharos approach than with turn-by-turn instructions. Exemplarily, Figure 6 shows the sketches drawn by the participants for route 2 in the TBT and PHA condition.

Qualitative Feedback

In addition to observations, after walking both routes and testing both variants of instructions, we conducted a semi-structured interview with the participants. Overall they were very satisfied with smartwatch navigation and felt very safe: “I do not have to remember the way, better than with a smartphone [...]” as navigation instructions on a smartphone have to be constantly checked (P6). One participant stated that “smartwatch navigation is really cool” (P7).

When asking the participants which of the two prototypes they prefer, the answers were mixed, even though PHA outperformed TBT in terms of confidence and spatial knowledge. Seven of them preferred the PHA condition and five the TBT instructions. Participants who preferred the TBT condition thought that the additional information is “unnecessary” (P9, P10), because they have “more to read”. However, the participants thought that Pharos is extremely helpful especially at the starting point, whether they preferred PHA in general (P4, P11) or TBT (P10).

DISCUSSION & LIMITATIONS

In this paper, we present a novel way to include *global landmarks* in pedestrian navigation instructions, which makes landmark-based navigation much more feasible than approaches using local landmarks. From a technical point of view, we have shown that the visibility of global landmarks can be determined automatically from existing and public available geotagged image content. Our approach of using a convolutional neural network combined with a sliding window is robust and easy to train for new landmarks, as it requires only

¹<http://www.qgis.org>

around 100 images for each landmark. In addition, we presented a complete pipeline to not only determine the visibility of global landmarks, but also to include them in the navigation instructions. It allows to compute the visibility of landmarks beforehand on map and navigation servers so that no further computation on a mobile or wearable device is needed.

From the user’s perspective, Pharos offers a lightweight but effective navigation support. In a user study, the Pharos approach outperformed current-state-of-the-art turn-by-turn instructions in important navigation metrics. We showed that small textual changes to the navigation instructions including hints on the location of global landmarks led to significantly more confident users. Additionally, although the Pharos navigation instructions contain more information than traditional turn-by-turn instructions, the participants looked less often at their smartwatch while navigating. We also saw that these changes resulted in users building a better spatial knowledge of their environment. Regarding time and error, we did not find any differences. This is not surprising, as a perfectly working turn-by-turn-based navigation system — also in the baseline condition the navigation instructions were manually triggered with a very high accuracy current state-of-the-art pedestrian system do not provide (see User Study) — is probably impossible to outperform in these metrics. However, we are convinced that there is a need for navigation systems providing additional values, as Pharos does, but without resulting in slower navigation and more errors.

Additionally, all participants of the user study lived in the city the study took place, but were not familiar with most parts of the testing sites. As such, our study suggests that Pharos can be helpful when traveling to unknown places. This is particularly true in places where traditional navigation techniques break down (e.g. the street signs are written in a different language or with different characters).

The user study has also highlighted the benefits of smartwatches for pedestrian navigation in general as already pointed out by related work [56, 57]. The participants liked navigating via smartwatch and positively mentioned that, when using a smartwatch, they have their hands free for other interactions. Additionally, the study has also revealed current problems of pedestrian navigation systems for smartwatches that could be overcome by using Pharos. Due to relatively low position accuracy, they constantly provide the user with information about the next decision point, which leads to low confidence and might prevent the user to build up spatial knowledge of their environment.

We evaluated the Pharos approach with global landmarks using large buildings. However, mountains or downtown skylines are global landmarks and can also be included in navigation instructions. This might be more difficult for two reasons: First, skylines and mountains usually look different from different view angles, which is often not the case for buildings. This results in larger training sets that are needed to determine their visibility. Furthermore, in rural areas the availability of geotagged images (e.g. by services like GSV) to determine the visibility of the landmark could be limited as well. To overcome this, we could extend the pipeline to include approaches

that compute the visibility of landmarks in rural areas with the help of digital elevation models (DEMs) [27] or to use geotagged data from crowd-sourced approaches such as Open Street View².

CONCLUSION & FUTURE WORK

This paper demonstrates that the Pharos approach is both (1) feasible and (2) has benefits compared to traditional turn-by-turn (i.e. no landmark) instructions with a user study. This means that global landmarks can be included in the navigation instructions within cities for many routes, and that global landmark-enriched instructions can be necessarily cognitively when the landmark is visible to a user and when it is not.

Although we evaluated Pharos for pedestrian navigation in urban areas, we are convinced that it can also be useful in other contexts (e.g. outdoors while hiking) or in more rural areas using mountains as global landmarks. For the future, we also want to exploit the use of Pharos in rural environments. Besides that, we are interested to explore whether Pharos could be applied to other navigation domains, such as biking or driving.

To achieve this we will further extend our pipeline to cover a larger set of global landmarks and also include other global landmarks that make sense for other modalities. Furthermore, the visibility maps could also be used to calculate scenic routes (e.g. for tourists) that guide the users through areas where global landmarks are very often visible.

ACKNOWLEDGMENTS

This work is supported by the Volkswagen Foundation through a Lichtenberg professorship.

Note: This version of the paper contains a fix for a reference issue that appeared in the original version.

REFERENCES

1. 2017. Topographic prominence. (Jan. 2017). https://en.wikipedia.org/w/index.php?title=Topographic_prominence&oldid=759905037 Page Version ID: 759905037.
2. Christina Bauer, Manuel Müller, and Bernd Ludwig. 2016. Indoor Pedestrian Navigation Systems: Is More Than One Landmark Needed for Efficient Self-localization?. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*. ACM, New York, NY, USA, 75–79. DOI: <http://dx.doi.org/10.1145/3012709.3012728>
3. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up Robust Features. In *European conference on computer vision*. Springer, 404–417. http://dx.doi.org/10.1007/11744023_32
4. John Brooke. 1996. SUS - A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*, Patrick W. Jordan, Bruce Thomas, Ian L. McClelland, and Bernard

²<http://openstreetview.org/>

- Weerdmeester (Eds.). Taylor & Francis, London, England.
<http://usabilitynet.org/trump/documents/Suschapt.doc>
5. Olivier Chapelle, Patrick Haffner, and Vladimir N. Vapnik. 1999. Support Vector Machines for Histogram-based Image Classification. *IEEE Transactions on Neural Networks* 10, 5 (1999), 1055–1064. DOI:<http://dx.doi.org/10.1109/72.788646>
 6. David M Chen, Georges Baatz, Kevin Köser, Sam S Tsai, Ramakrishna Vedantham, Timo Pylvänäinen, Kimmo Roimela, Xin Chen, Jeff Bach, Marc Pollefeys, and others. 2011. City-scale Landmark Identification on Mobile Devices. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011)*. IEEE, 737–744.
<http://dx.doi.org/10.1109/cvpr.2011.5995610>
 7. Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, and Christos Efstratiou. 2000. Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, NY, USA, 17–24. DOI :
<http://dx.doi.org/10.1145/332040.332047>
 8. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A Large-scale Hierarchical Image Database. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*. IEEE, 248–255.
<http://dx.doi.org/10.1109/cvpr.2009.5206848>
 9. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. 2013. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. *CoRR* abs/1310.1531 (2013). <http://arxiv.org/abs/1310.1531>
 10. Markus Dräger and Alexander Koller. 2012. Generation of Landmark-based Navigation Instructions from Open-source Data. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL '12)*. Stroudsburg, PA, USA, 757–766.
<http://dl.acm.org/citation.cfm?id=2380816.2380908>
 11. Patrick Foo, William H Warren, Andrew Duchon, and Michael J Tarr. 2005. Do Humans Integrate Routes Into a Cognitive Map? Map-versus Landmark-based Navigation of Novel Shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, 2 (2005), 195. DOI :
<http://dx.doi.org/10.1037/0278-7393.31.2.195>
 12. Alinda Friedman and Bernd Kohler. 2003. Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and other Two-dimensional Data Sets. *Psychological methods* 8, 4 (2003), 468–49.
 13. Tommy Garling, Anders Book, and Erik Lindberg. 1984. Cognitive Mapping of Large-scale Environments the Interrelationship of Action Plans, Acquisition, and Orientation. *Environment and Behavior* 16, 1 (1984), 3–34.
 14. Ioannis Giannopoulos, Peter Kiefer, and Martin Raubal. 2013. The Influence of Gaze History Visualization on Map Interaction Sequences and Cognitive Maps. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on MapInteraction (MapInteract '13)*. ACM, New York, NY, USA, 1–6. DOI :
<http://dx.doi.org/10.1145/2534931.2534940>
 15. Reginald G. Golledge. 1992. Place Recognition and Wayfinding: Making Sense of Space. *Geoforum* 23, 2 (1992), 199 – 214. DOI :
[http://dx.doi.org/10.1016/0016-7185\(92\)90017-X](http://dx.doi.org/10.1016/0016-7185(92)90017-X)
 16. Reginald G Golledge. 1999. *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. JHU press.
 17. Joy Goodman, Stephen A Brewster, and Philip Gray. 2005. How Can We Best Use Landmarks to Support Older People in Navigation? *Behaviour & Information Technology* 24, 1 (2005), 3–20.
 18. Kotaro Hara, Shiri Azenkot, Megan Campbell, Cynthia L. Bennett, Vicki Le, Sean Pannella, Robert Moore, Kelly Minckler, Rochelle H. Ng, and Jon E. Froehlich. 2013. Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. ACM, New York, NY, USA, Article 16, 8 pages. DOI :
<http://dx.doi.org/10.1145/2513383.2513448>
 19. Kotaro Hara, Christine Chan, and Jon E. Froehlich. 2016. The Design of Assistive Location-based Technologies for People with Ambulatory Disabilities: A Formative Study. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1757–1768. DOI :
<http://dx.doi.org/10.1145/2858036.2858315>
 20. Sandra G. Hart and Lowell E. Stavenland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati (Eds.). Elsevier, 139–183. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000004342_1999205624.pdf
 21. Harlan Hile, Radek Grzeszczuk, Alan Liu, Ramakrishna Vedantham, Jana Košecka, and Gaetano Borriello. 2009. Landmark-based Pedestrian Navigation with Enhanced Spatial Reasoning. In *7th International Conference on Pervasive Computing (Pervasive 2009)*. Springer, 59–76.
 22. Simon Holland, David R. Morse, and Henrik Gedenryd. 2002. AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal Ubiquitous Comput.* 6, 4 (2002), 253–259. DOI :
<http://dx.doi.org/10.1007/s007790200025>
 23. Yitao Hu, Xiaochen Liu, Suman Nath, and Ramesh Govindan. 2016. ALPS: Accurate Landmark Positioning at City Scales. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1147–1158.

24. Alexander Klippel and Stephan Winter. 2005. Structural Saliency of Landmarks for Route Directions. In *Proceedings of the 2005 International Conference on Spatial Information Theory (COSIT'05)*. Springer-Verlag, Berlin, Heidelberg, 347–362. DOI: http://dx.doi.org/10.1007/11556114_22
25. Christian Kray and Gerd Kortuem. 2004. *Interactive Positioning Based on Object Visibility*. Springer Berlin Heidelberg, Berlin, Heidelberg, 276–287. DOI: http://dx.doi.org/10.1007/978-3-540-28637-0_24
26. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*. 1097–1105.
27. Jan Lee. 1994. Visibility Dominance and Topographic Features on Digital Elevation Models. *Photogrammetric Engineering and Remote Sensing* 60, 4 (1994), 451–456.
28. Yunpeng Li, D. J. Crandall, and D. P. Huttenlocher. 2009. Landmark Classification in Large-scale Image Collections. In *2009 IEEE 12th International Conference on Computer Vision*. 1957–1964. DOI: <http://dx.doi.org/10.1109/ICCV.2009.5459432>
29. Hyunchul Lim, YoonKyong Cho, Wonjong Rhee, and Bongwon Suh. 2015. Vi-Bros: Tactile Feedback for Indoor Navigation with a Smartphone and a Smartwatch. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 2115–2120. DOI: <http://dx.doi.org/10.1145/2702613.2732811>
30. Kristin L Lovelace, Mary Hegarty, and Daniel R Montello. 1999. Elements of Good Route Directions in Familiar and Unfamiliar Environments. In *International Conference on Spatial Information Theory*. Springer, 65–82.
31. David G Lowe. 1999. Object Recognition from Local Scale-invariant Features. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision 1999*, Vol. 2. 1150–1157.
32. Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. 2010. Photo2Trip: Generating Travel Routes from Geo-tagged Photos for Trip Planning. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 143–152. DOI: <http://dx.doi.org/10.1145/1873951.1873972>
33. Kevin Lynch. 1960. *The Image of the City*. MIT press.
34. Rainer Malaka and Alexander Zipf. 2000. Deep Map: Challenging IT Research in the Framework of a Tourist Information System. In *Information and Communication Technologies in Tourism 2000*. Springer, 15–27.
35. David K. McGookin and Stephen A. Brewster. 2013. Investigating and Supporting Undirected Navigation for Runners. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 1395–1400. DOI: <http://dx.doi.org/10.1145/2468356.2468605>
36. Alexandra Millonig and Katja Schechtner. 2007. Developing Landmark-based Pedestrian-navigation Systems. *IEEE Transactions on Intelligent Transportation Systems* 8, 1 (2007), 43–49.
37. Daniel R Montello and Andrew U Frank. 1996. Modeling Directional Knowledge and Reasoning in Environmental Space: Testing Qualitative Metrics. In *The construction of cognitive maps*. Springer, 321–344.
38. Symeon Papadopoulos, Christos Zigkolis, Yiannis Kompatsiaris, and Athena Vakali. 2011. Cluster-based Landmark and Event Detection for Tagged Photo Collections. *IEEE MultiMedia* 18, 1 (2011), 52–63.
39. Max Pfeiffer, Tim Dünste, Stefan Schneegeass, Florian Alt, and Michael Rohs. 2015. Cruise Control for Pedestrians: Controlling Walking Direction Using Electrical Muscle Stimulation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2505–2514. DOI: <http://dx.doi.org/10.1145/2702123.2702190>
40. Günther Pospischil, Martina Umlauft, and Elke Michlmayr. 2002. Designing LoL@, a Mobile Tourist Guide for UMTS. In *Proceedings of the 4th International Symposium on Mobile Human-Computer Interaction (Mobile HCI '02)*. Springer-Verlag, London, UK, UK, 140–154. <http://dl.acm.org/citation.cfm?id=645739.758125>
41. Martin Raubal and Stephan Winter. 2002. Enriching Wayfinding Instructions with Local Landmarks. In *International Conference on Geographic Information Science*. Springer, 243–259.
42. Simon Robinson, Matt Jones, Parisa Eslambolchilar, Roderick Murray-Smith, and Mads Lindborg. 2010. “Did It My Way”: Moving Away from the Tyranny of Turn-by-turn Pedestrian Navigation. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 341–344. DOI: <http://dx.doi.org/10.1145/1851600.1851660>
43. Tracy Ross, Andrew May, and Simon Thompson. 2004. *The Use of Landmarks in Pedestrian Navigation Instructions and the Effects of Context*. Springer Berlin Heidelberg, Berlin, Heidelberg, 300–304. DOI: http://dx.doi.org/10.1007/978-3-540-28637-0_26
44. Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. 2011. ORB: An Efficient Alternative to SIFT or SURF. In *International Conference on Computer Vision 2011*. IEEE, 2564–2571.
45. Johannes Schöning, Antonio Krüger, Keith Cheverst, Michael Rohs, Markus Löchtfeld, and Faisal Taher. 2009. PhotoMap: Using Spontaneously Taken Images of Public Maps for Pedestrian Navigation Tasks on Mobile Devices. In *Proceedings of the 11th International*

- Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*. ACM, New York, NY, USA, Article 14, 10 pages. DOI: <http://dx.doi.org/10.1145/1613858.1613876>
46. Alexander W. Siegel and Sheldon H. White. 1975. The Development of Spatial Representations of Large-scale Environments. *Advances in Child Development and Behavior* 10 (1975), 9–55.
 47. Caroline Snowdon and Christian Kray. 2009. Exploring the Use of Landmarks for Mobile Navigation Support in Natural Environments. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*. ACM, New York, NY, USA, 13:1–13:10. DOI: <http://dx.doi.org/10.1145/1613858.1613875>
 48. Molly E. Sorrows and Stephen C. Hirtle. 1999. The Nature of Landmarks for Real and Electronic Spaces. In *Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science (COSIT '99)*. Springer-Verlag, London, UK, 37–50. <http://portal.acm.org/citation.cfm?id=646127.758644>
 49. Sibylle D. Steck and Hanspeter A. Mallot. 2000. The Role of Global and Local Landmarks in Virtual Environment Navigation. *Presence: Teleoperators and Virtual Environments* 9, 1 (2000), 69–83.
 50. Waldo R Tobler. 1994. Bidimensional Regression. *Geographical Analysis* 26, 3 (1994), 187–212.
 51. Edward C Tolman. 1948. Cognitive Maps in Rats and Men. *Psychological Review* 55, 4 (1948), 189.
 52. Ariane Tom and Michel Denis. 2003. *Referring to Landmark or Street Information in Route Directions: What Difference Does It Make?* Springer Berlin Heidelberg, Berlin, Heidelberg, 362–374. DOI: http://dx.doi.org/10.1007/978-3-540-39923-0_24
 53. Ariane Tom and Michel Denis. 2004. Language and Spatial Cognition: Comparing the Roles of Landmarks and Street Names in Route Instructions. *Applied Cognitive Psychology* 18, 9 (2004), 1213–1230.
 54. Jasper RR Uijlings, Koen EA van de Sande, Theo Gevers, and Arnold WM Smeulders. 2013. Selective Search for Object Recognition. *International Journal of Computer Vision* 104, 2 (2013), 154–171.
 55. Shoko Wakamiya, Hiroshi Kawasaki, Yukiko Kawai, Adam Jatowt, Eiji Aramaki, and Toyokazu Akiyama. 2016. Lets Not Stare at Smartphones While Walking: Memorable Route Recommendation by Detecting Effective Landmarks. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 1136–1146. DOI: <http://dx.doi.org/10.1145/2971648.2971758>
 56. Dirk Wenig, Johannes Schöning, Brent Hecht, and Rainer Malaka. 2015. StripeMaps: Improving Map-based Pedestrian Navigation for Smartwatches. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 52–62. DOI: <http://dx.doi.org/10.1145/2785830.2785862>
 57. Dirk Wenig, Alexander Steenbergen, Johannes Schöning, Brent Hecht, and Rainer Malaka. 2016. ScrollingHome: Bringing Image-based Indoor Navigation to Smartwatches. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, New York, NY, USA, 400–406. DOI: <http://dx.doi.org/10.1145/2935334.2935373>
 58. Stephan Winter, Martin Tomko, Birgit Elias, and Monika Sester. 2008. Landmark Hierarchies in Context. *Environment and Planning B: Planning and Design* 35, 3 (2008), 381–398.
 59. Jason Wither, Carmen E. Au, Raymond Rischpater, and Radek Grzeszczuk. 2013. Moving Beyond the Map: Automated Landmark Based Pedestrian Guidance Using Street Level Panoramas. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 203–212. DOI: <http://dx.doi.org/10.1145/2493190.2493235>
 60. Yan-Tao Zheng, Ming Zhao, Yang Song, Hartwig Adam, Ulrich Buddemeier, Alessandro Bissacco, Fernando Brucher, Tat-Seng Chua, and Hartmut Neven. 2009. Tour the World: Building a Web-scale Landmark Recognition Engine. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*. IEEE, 1085–1092.