WikEye – Using Magic Lenses to Explore Georeferenced Wikipedia Content

Brent Hecht
Department of Geography
University of California Santa Barbara
Santa Barbara, CA
93106, USA
bhecht@geog.ucsb.edu

Michael Rohs
Deutsche Telekom Laboratories
TU Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
michael.rohs@telekom.de

Johannes Schöning
Institute for Geoinformatics
University of Münster
Robert-Koch-Str. 26-28
48149 Münster, Germany
j.schoening@uni-muenster.de

Antonio Krüger
Institute for Geoinformatics
University of Münster
Robert-Koch-Str. 26-28
48149 Münster, Germany
antonio.krueger@uni-muenster.de

ABSTRACT

Traditional paper-based maps are still superior in several ways to their digital counterparts used on mobile devices. Namely, paper-based maps provide high-resolution, large-scale information with zero power consumption. Digital maps offer personalized and dynamic information, but suffer from small outer scales and low resolutions. In this paper, we present WikEye, an interdisciplinary project that demonstrates proof-of-concepts of three novel research ideas related to this topic: multi-dimensional interaction with a magic lens device, spatio-temporal Wikipedia data mining, and magic lens markerless tracking. Through its integration of the advantages of static and dynamic maps and its employment of Wikipedia as its primary knowledge repository, we envision WikEye being used to address mobile technology issues in the tourism and education fields. In particular, we focus on the need for effective tourism-related mobile technologies and the demand from educators for new methods that students can use to interact with information.

1. INTRODUCTION

Tourism provides over six percent of the world's gross domestic product [20]. However, despite the fact that mobility is at the heart of tourist activity [4], mobile technologies have yet to infiltrate the tourism industry [7]. Similarly, educators are always seeking out new technologies for student-information interaction, such as in [8]. Applications like RFID maps [13], Timmi (Timmi is Mobile Map Interaction) [18, 19] and Minotour [6] are members of a new generation of systems aimed at making mobile technology more appealing to tourists, and Minotour is designed specifically for tourism-based education. Minotour is a location-aware tour guide application for mobile devices that frees the content of Wikipedia from existing physical and organizational restrictions by generating appealing narratives relevant to a user's current activity space. Timmi [18, 19] combines the advantages of large scale and high-resolution maps with the interactivity and up-to-date nature of dynamically queried geoobjects. Through the integration of the premise of Timmi and its new markerless tracking interface with the geography education-oriented Wikipedia technology and research of Minotour, we can provide users of mobile devices with a novel approach to space and place understanding through dynamic display empowered static maps.

Like Semapedia and similar projects, WikEye takes Wikipedia off the computer and moves it into the real world. However, WikEye does so in an entirely novel fashion: by placing Wikipedia data into interactive spatial, temporal, and semantic augmented reality contexts that are referenced to static maps. In doing so, we meet the call for greater interaction between paper maps and guidebooks (in this case, Wikipedia) in [7], providing a prototype mobile technology appealing not only to the tourist, but to the educator as well.

Because of its Wikipedia article density and its history, a city map of Berlin provides an excellent test case for WikEye. Maps of Berlin alone cannot satisfactorily educate the user in the rich history of Berlin, and using a map and another information source at the same time can be a cumbersome experience [7]. Ideally, visitors to Berlin or students of the city would be able to easily view information contained in a book or travel literature through the context of a large-scale, high-resolution paper map. WikEye provides this exact functionality in an easy-to-use handheld application powered by a novel map-device interaction scheme and informed by the vast quantity of free information in Wikipedia.

WikEye has three main components, each of which comprises an innovative area of research. Section 2 covers the issues involved with mining Wikipedia for the dynamic information that drives WikEye. In Section 3, we discuss multi-dimensional interaction methods based on the interaction primitives of Rohs [16]. Our efforts towards markerless tracking, a vital technology that allows WikEye to be used on standard-looking static maps (thus reducing the infrastructure costs of the entire WikEye system), is briefly described in Section 4.

2. WIKIPEDIA

Wikipedia has been extensively studied as a phenomenon, but its utility as a vast repository of free world knowledge is little understood [3]. In this project, we mine Wikipedia for three key types information: spatial data, temporal data, and semantic connections between places (see Figure 1). Without such a massive and freely available data corpus, the utility of WikEye would be categorically diminished, and WikEye would require significant and impractical specialized content development.

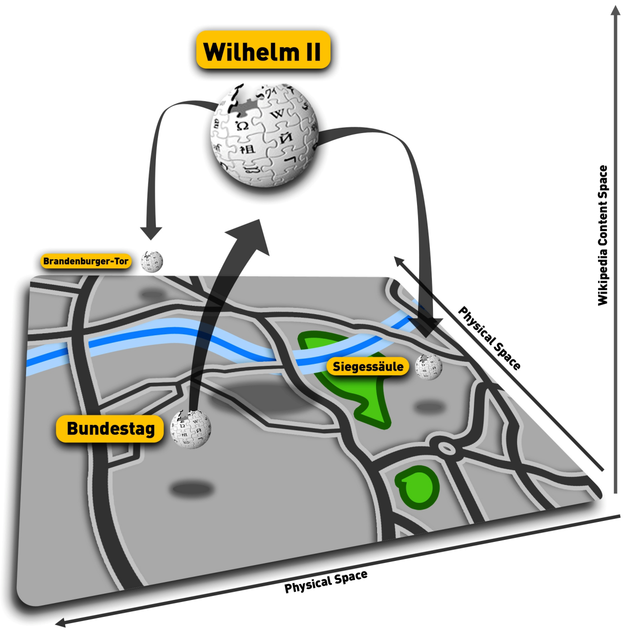

Figure 1: Spatial feature-edge-feature relationships in Wikipedia.

Wikipedia spatial data extraction is comprised of two tasks: identifying and evaluating the point-only and limited explicitly spatial Wikipedia data, and, more significantly, georeferencing Wikipedia articles, which provides the implicit two-dimensional geospatial extent of articles' subjects. In [6], it was found that while 20-30 percent of Wikipedia articles have an implicit spatial extent, only about 4-5 percent are tagged by Wikipedia users with explicit spatial data. For both the implicit and explicit tasks, the data mining approach taken is identical to that in [6]: explicit data is obtained through elaborate parsing techniques informed by a knowledge of Wikipedia custom; implicit spatial information is made explicit through a basic and customized georeferencing process. Some supplementary spatial data extraction techniques using external data sources are also being explored.

The goal of our Wikipedia temporal information extraction procedure is to discover the start and end years of spatial features via the temporal expressions in the features' corresponding Wikipedia articles. Critically, key historical eras in each article must also be identified. The mining of natural text for temporal information is a well-understood problem [9], but it has proven difficult to solve with high levels of accuracy. Fortunately, our needs are limited compared to experiments such as those in [10] and Wikipedia provides some inherent advantages for temporal information extraction. The majority of temporal information extraction research involves identifying the syntax and evaluating the semantics of all embedded temporal expressions – "April 29", "8:20:00 p.m. on April 10, 1983", "yesterday", "last Monday", "in the Fall", etc. – in corpi of natural texts. However, the functionality of WikEye demands only that we take a statistically unbiased sampling of temporal expressions, and that we use years as our maximum sampling precision. Since Wikipedia's encyclopedic style is replete with fixed-format temporal expressions ("2006", as opposed to "last year"), we have found that we can gather temporal information with both sufficient quantity and accuracy by using relatively simple temporal information extraction procedures, such as those described in [11]. In addition, temporal accuracy is improved by the pre-disambiguated temporal expressions in Wikipedia. These disambiguations take the form of wiki links to the large number of articles on specific years, such as the "1983" article, for example.

Once the temporal information is extracted, the application must then identify the start and end years of each spatial feature, as well as features' key historical periods and/or eras. This is done via the analysis of article *temporal reference profiles* with pattern recognition techniques and other statistical methods. Temporal reference profiles are essentially histograms that describe the number of references to each year on a timeline. Figure 2 shows the profile of the article on Berlin. Common-sense methods will also be employed in the temporal analysis process. For example, years written into headings of sections will be weighed differently than years buried in the text body.

The fourth dimension of interaction for our application is spatially-referenced semantic connections. Mining Wikipedia for semantic connections between places is a simpler problem than temporal and spatial information extraction. In fact, easily identified (although not easily interpreted [21]), semantic connections are a key benefit of the Wiki concept. Due to the restraints inherent to our novel interaction scheme (see below), it was determined that only spatial feature-edge-node-edge-spatial feature relationships (e.g. Bundestag - Wilhelm II - Brandenburg Gate) will be considered (see Figure 1). In other words, only relationships between two spatial features through a *single* intermediary Wikipedia article will be presented to the user. Spatial feature-edge-spatial feature relationships (Berlin - Bundestag) are not included, as they have been found to be overloaded with the spatial relationships that are already evident to the user via the map display. Delivery of the "story line" between the features and through the intermediary article will be similar to that described in [6].

3. INTERACTION WITH WIKIPEDIA DATA

By making explicit the spatial, temporal, and relational structures implicit in Wikipedia, we free the encyclopedia's

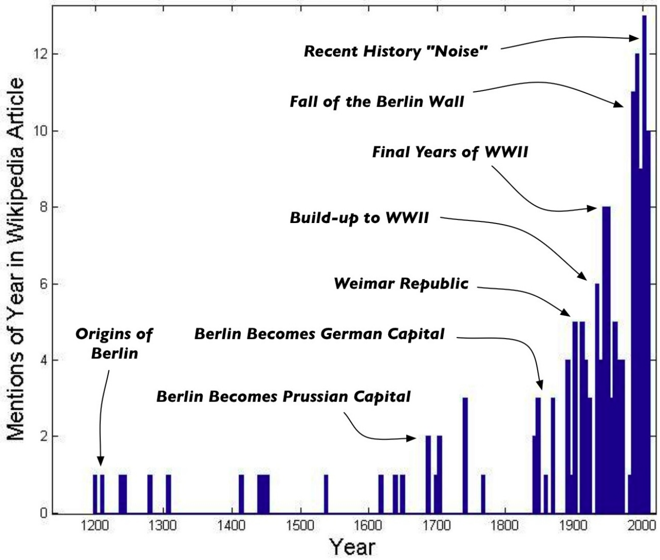

Figure 2: Berlin article temporal reference profile.

data not only from the constraints of the desktop, but also from its rigid interaction medium. This section describes how we provide novel methods for interaction with these structures.

As is noted in the introduction, the WikEye user's view of the physical map is combined with Wikipedia-derived information on the mobile display. The user acts on two layers of information: the "transparent" device screen with georeferenced objects and the physical map surface. The camera display unit acts as a movable window into an augmented view of the physical map. This configuration is known as the *see-through interface* [1, 5], the *magic lens* [2], or the *magnifying glass approach* [14]. Originally, see-through tools or toolglasses [1] are GUI widgets that are layered between application objects on the screen and the cursor. They can be positioned over application objects in order to modify their appearance or set the context for command execution. Here, the background layer is the printed map and the overlay is realized by the handheld device. Usability issues that arise in transparent layered user interfaces and see-through tools include "switching costs" for shifting attention between the two layers [5] and the visual interference of background and foreground objects.

It is in this "magic lens" context that WikEye must provide an interface to the aforementioned Wikipedia structures. Specifically, we enable the browsing of Wikipedia via four entirely new interaction tasks:

1. Space-time exploration
2. Space-time selection
3. Space-Wikipedia relationship exploration
4. Space-Wikipedia relationship selection

We construct these tasks via the interaction primitives developed by Rohs [16]: pointing, rotation, tilting, distance, stay, keystroke, and sweeping. This research extends previous work that derived from Rohs' primitives the set of mobile device spatial interaction tasks (exploration, spatial selection, annotation, navigation, guidance). The tasks from the previous work, based on [12] and the experience of Rukzio et al. [17], are also supported by WikEye and must be compatible with the new Wikipedia-oriented tasks.

A primary new interaction task in WikEye is that of space-time exploration. By *sweeping* the mobile camera device over the map, the user can explore Wikipedia articles on the map. By *rotating* the device – evoking a clock metaphor – the user can simultaneously wander the chronological eras relevant to the articles within the current spatial extent. A more specialized task is the space-time selection of an object or area with *point* and *shoot* (*keystroke*) in order to view Wikipedia text data from this spatio-temporal region. Again, by *rotating* the mobile camera device the user chooses a time interval. By clicking (*keystroke*) with the cross-hair on a georeferenced Wikipedia object the user gets a Wikipedia paragraph, or "snippet", from the object for the selected time period. For example, while exploring a map of Berlin with WikEye, a user may want to know more about the Reichstag, a spatial feature / Wikipedia article tuple extracted in Wikipedia data mining process. If the user wishes to know more about the Reichstag during the beginning of the Second World War, the user selects this period by rotating the device to the correct time period and clicking on the Reichstag's icon. The user would then receive a Wikipedia snippet about the Reichstag during that period.

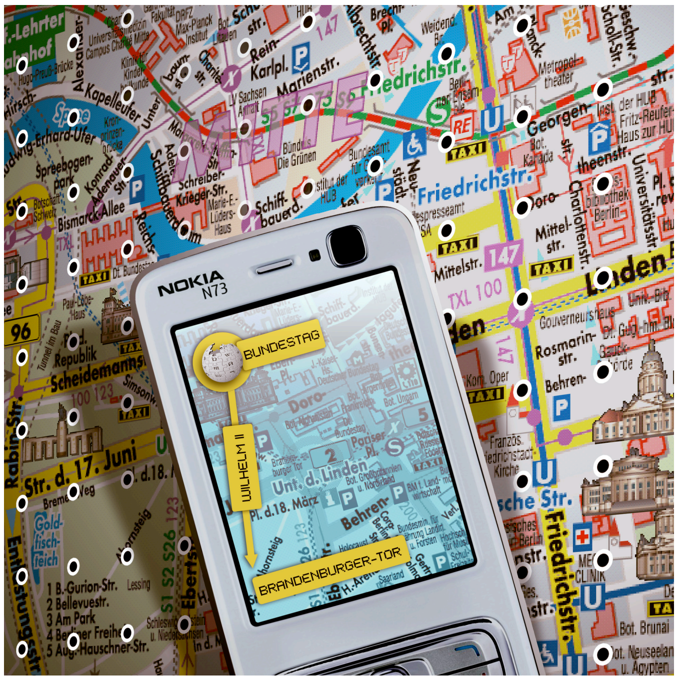

Figure 3: Application in use (constructed).

A second basic interaction paradigm is that of space-Wikipedia relationship exploration. The user interacts with space-Wikipedia relationships by sweeping the mobile device over the map to view arrows drawn on the screen between related georeferenced Wikipedia articles (see Figure 3). These arrows have labels that attempt to succinctly describe the relationships they represent, relationships that are limited to that of feature-edge-node-edge-feature connections. Space-Wikipedia relationship selection is a derivative of space-Wikipedia relationship exploration. By selecting a georeferenced Wikipedia article with a click, the user can explore feature-edge-node-edge-feature relationships beginning at the selected feature. By selecting a second object, a snippet-based "story line" between the two articles is displayed on the mobile device, as is done more elaborately in [6].

4. REAL-TIME MAP TRACKING IN WIKEYE

A key prerequisite to all of the interaction tasks described above is that the user's view of the physical map is accurately combined with dynamic information on the mobile display. In order to align the overlay graphics with objects in the camera view, the video stream of the camera is continuously analyzed. The tracking algorithm computes the position of the camera view on the map. For each camera frame it computes a projective mapping (planar homography) from the map coordinate system to the image coordinate system. We extend this mapping to a mapping from the image plane to real world map coordinates with a simple affine transformation.

This real-time tracking task is simplified for execution on a mobile phone with limited computing resources by introducing regularly spaced black dots onto the map (see Figures 3 and 4). The grid of dots subdivides the map into squared patches that are used to correlate the camera view with the map. Using the correlation between a precomputed map and the camera image, the system computes the actual position of the mobile device over the map (see Figure 4). The method is described in more detail in [15]. In a future version, we aim to replace the dots by horizontal and vertical lines, a so called *map grid*. This map grid is less obtrusive and commonly found on city maps, for example to simplify the lookup of street names.

5. IMPLEMENTATION

The Wikipedia data is processed offline using an extensive, custom-built Java Wikipedia parser (English Wikipedia only) and is transferred to devices in the same manner as in [6], with both online and offline device operation supported. Significant important implementation details are involved with the Wikipedia data processing. For instance, we prevent all time-focused articles (the article on "1983", for example) from being included as the node in the feature-edge-node-edge-feature relationships, as these articles generally lead to very uninteresting and uninformative relationships. Also, many special cases must be handled in both our spatial and temporal information extraction processes. More information on the implementation, as well as details about the results, will be available in an upcoming paper.

The magic lens prototype using our markerless tracking method is implemented on a Nokia N80 mobile phone. The view finder mode has a frame size of 176×144 pixels at 15 fps. We successfully tested the algorithm with 3 city maps, with

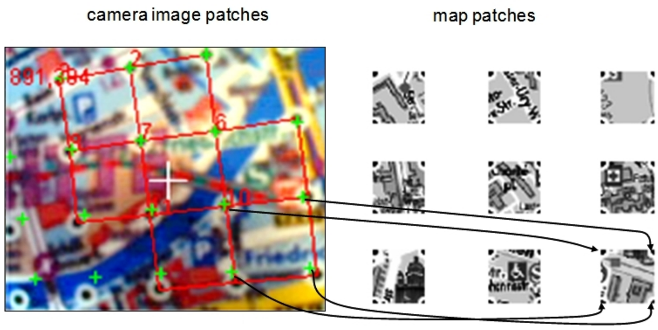

Figure 4: Markerless tracking.

a maximum of 386 patches. The main test map, a city map of Berlin, has $16\times12 = 192$ correlation patches. The template file has a size of 30 kB, which means that it can quickly be downloaded via the mobile phone network. Printed on a DIN A3 sheet (print area 36×27 cm), the size of each patch is 21×21 mm and the diameter of each black grid dot is 3 mm (surrounded by 1 mm whitespace). From each patch 12×12 gray-value samples are taken and compared to the stored template patches. In order to speed up the correlation process only a 5×5-patch window centered at the previously detected position is considered. With this optimization the algorithm processes 5 to 7 frames per second.

6. CONCLUSION AND FURTHER WORK

In this paper we have discussed the fundamental concepts and components of WikEye, a system that uses Wikipedia's spatial and temporal implicit data structures to allow tourists, students, and others to interact with Wikipedia content presented on mobile devices as a transparent overlay on physical maps. This paper is a first attempt to unify the aforementioned content selection process, the Wikipedia entity relationship technology of Minotour, the embedded interaction patterns developed in the context of the Timmi system, and the map tracking research described in section 4. We have highlighted here several portions of this project that we think are crucial for a rich tourism and/or educational experience: (a) organization of content along spatial, temporal, and semantic dimensions, (b) intuitive interaction patterns with the mobile device and the physical map and (c) introduction of unobtrusive tracking markers to reduce distraction of users to a minimum. In all three areas open questions still exist: the optimal temporal information extraction methodologies need to be identified, the interaction classes need to be fully implemented and evaluated in user studies, and the tracking of the mobile device needs to be improved by further reducing the visual appearance of the markers, e.g. by using only the existing structure of a map grid. Finally, while most the individual parts of the system exist in prototype form, we are still working on developing an integrated prototype. We believe that our proposed combination of a structuring process with adequate interaction mechanisms will allow people to explore the richness of Wikipedia in an entirely new and portable fashion, without being impeded with the confined size of mobile screens. Of course, this has still to be confirmed by user studies with the final prototype.

7. REFERENCES

[1] E. A. Bier, M. C. Stone, K. Fishkin, W. Buxton, and T. Baudel. A taxonomy of see-through tools. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 358–364. ACM Press, 1994.

[2] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80. ACM Press, 1993.

[3] E. Gabrilovich and S. Markovitch. Overcoming the brittleness bottleneck using wikipedia: Enhancing text categorization with encyclopedic knowledge. *Proceedings of the 21st National Conference on Artificial Intelligence (AIII-06)*, pages 1301–1306, 2006.

[4] C. M. Hall. Reconsidering the geography of tourism and contemporary mobility. *Geographical Research*, 43(2):125–139, 2005.

[5] B. L. Harrison, G. Kurtenbach, and K. J. Vicente. An experimental evaluation of transparent user interface tools and information content. In *UIST '95: Proceedings of the 8th annual ACM symposium on User interface and software technology*, pages 81–90. ACM Press, 1995.

[6] B. Hecht, D. Dara-Abrams, N. Starosielski, K. Goldsberry, J. Dillemuth, and J. Roberts. Minotour: A location-aware mobile tour application that weaves a spatial tale from wikipedia. *Proceedings of the 2007 Meeting of The AAG*, 2007.

[7] K. Kuutti and E. Karsten, editors. *Tourism and mobile technology*, Helsinki, Finland, 2003. Kluwer Academic Press.

[8] K. Leichtenstern and E. André. Social mobile interaction using tangible user interfaces and mobile phones. *Proceedings of the Conference of the Gesellschaft für Informatik, Workshop Mobile and Embedded Interactive Systems (MEIS)*, 2006.

[9] I. Mani. Recent developments in temporal information extraction. *Proceedings of Recent Advances in Natural Language Processing*, 2004.

[10] I. Mani and G. Wilson. Robust temporal processing of news. *38th Annual Meeting on Association for Computational Linguistics*, 38:69–76, 2000.

[11] D. McKay and S. J. Cunningham. Mining dates from historical documents. *University of Waikato Department of Computer Science Technical Report*, 2000.

[12] V. Paelke and C. Brenner. Development of a mixed reality device for interactive on-site geo-visualization. *Proceedings of 18th Simulation and Visualization Conference*, 2007.

[13] D. F. Reilly, M. E. Rodgers, R. Argue, M. Nunes, and K. Inkpen. Marked-up maps: combining paper maps and electronic information resources. *Personal and Ubiquitous Computing*, pages 215–226, 2006.

[14] J. Rekimoto. The magnifying glass approach to augmented reality systems. In *International Conference on Artificial Reality and Tele-Existence '95 / Conference on Virtual Reality Software and Technology (ICAT/VRST '95)*, pages 123–132, 1995.

[15] M. Rohs, J. Schöning, A. Krüger, and B. Hecht. Towards real-time markerless tracking of magic lenses on paper maps. In *Pervasive 2007 Adjunct Proceedings*, Toronto, Ontario, Canada, May 2007. Austrian Computer Society (OCG).

[16] M. Rohs and P. Zweifel. A conceptual framework for camera phone-based interaction techniques. In H. W. Gellersen, R. Want, and A. Schmidt, editors, *Pervasive Computing: Third International Conference, PERVASIVE 2005*, pages 171–189, Munich, Germany, May 2005. LNCS 3468, Springer.

[17] E. Rukzio, K. Leichtenstern, V. Callaghan, A. Schmidt, P. Holleis, and J. Chin. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. *Eighth International Conference on Ubiquitous Computing (Ubicomp)*, 2006.

[18] J. Schöning, J. T. Heuer, H. J. Müller, and A. Krüger. The marauders lens. *Proceedings of the 4th International Conference on GIScience, Extended Abstracts*, 2006.

[19] J. Schöning, A. Krüger, and H. J. Müller. Interaction of mobile camera devices with physical maps. *Adjunct Proceeding of the Fourth International Conference on Pervasive Computing*, 2006.

[20] United Nations World Tourism Organization. Tourism and the world economy, 2003. `www.unwto.org/facts/eng/economy.htm` [online; accessed January 24, 2007].

[21] M. Völkel, M. Krötzsch, D. Vrandecic, H. Haller, and R. Studer. Semantic wikipedia. In *Proceedings of the 15th International Conference on World Wide Web (WWW)*, volume 15, pages 585–594, Edinburgh, Scotland, May 2006. ACM Press.