Evaluating Automatically Generated Location-Based Stories for Tourists

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

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

Nicole Starosielski
Department of Film Studies
University of California Santa Barbara
Santa Barbara, CA 93106
n_star@umail.ucsb.edu

CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 978-1-60558-012-8/08/04

Abstract

Tourism provides over six percent of the world's gross domestic product. As a result, there have been many efforts to use technology to improve the tourist's experience via mobile tour guide systems. One key bottleneck in such location-based systems is content development; existing systems either provide trivial information at a global scale or present quality narratives but at an extremely local scale. The primary reason for this dichotomy is that, although good narrative content is more educationally effective (and more entertaining) than a stream of simple, disconnected facts, it is time-intensive and expensive to develop. However, the WikEar system uses narrative theory-informed data mining methodologies in an effort to produce high-quality narrative content for any location on Earth. It allows tourists to interact with these narratives using their camera-enabled cell phones and an innovative interface designed around a magic lens and paper map metaphor. In this paper, we describe a first evaluation of these narratives and the WikEar interface, which reported promising, but not conclusive, results. We also present ideas for future work that will use this feedback to improve the narratives.

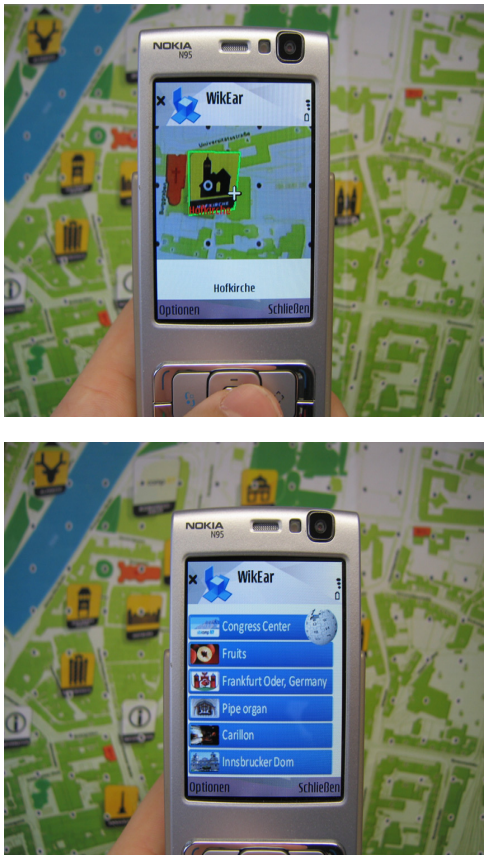

Figure 1: User interacting with Wikear (top). Selection of a destination the system play back an auto generated story (bottom).

Keywords
Tourism, mobile maps, auto-generated content, user feedback

ACM Classification Keywords
H.5.1 Multimedia Information Systems Information.

Introduction

While tourists choose to travel for many reasons, one significant market segment is that of tourists who wish to learn about the places they visit. It is for these tourists that mobile tour guide systems are severely lacking. In our research we do not seek to replace tour guides with mobile systems, but rather to improve the mobile systems to be appealing as a tourist guide. Anyone who has taken a high-quality tour from an excellent tour guide who uses narrative strategies to explain her/his surroundings knows the value of such an experience. Several researchers and well-known geographers have elaborated on and confirmed this sentiment [7, 8]. Automated approaches to *digital* tour guide content development have eschewed the narrative component in favor of the presentation of simple facts. This ignores the vital utility and enjoyment of narrative comprehension of a location. There are some digital tour guides that present information in a narrative format, but content development on these systems is never automated, and the spatial extent for which they are designed is usually quite small (i.e. Kyoto, Japan and a virtual art museum in [7]).

In our approach, called WikEar, we try to overcome the problems of hand-made content by generating stories with data mined from Wikipedia that is automatically organized according to principles derived from narrative theory and woven into an educational audio tours starting and ending at stationary city maps. The system generates custom, location-based "guided tours" that are never out-of-date and are ubiquitous – even at an international scale. WikEar uses a magic lens-based interaction scheme, allowing the continued use of paper maps, which have been shown to be particularly important in the tourist experience [4]. The user can select a start and end location by sweeping the mobile camera device over the map as illustrated in figure 1. The device is tracked relative to the paper map using tiny black dots printed on the maps [9]. By leveraging the wide availability of large public city maps and the magic lens metaphor, WikEar avoids the costs of GPS and the interaction problems of small screen map programs.

Such automated narrative generation and magic lens systems may prove increasingly popular as narrative generation and magic lens technology improves. Of course these stories can also be used in completely different systems like our Virtual Globe 2.0 [12]. As such, we developed and present a novel means of evaluating such systems. Because the automated narratives are not of a standard format, we could not simply test what the user thought about the quality of the story. Thus, our evaluation represents an innovative narrative theory-derived approach in which we assess whether or not testers of WikEar used narrative schema in comprehending the audio stories.

Related Work

Baus et al. [3] provide a good overview of the major mobile guides (most of them are research prototypes). Abowd et al. [1] present one of the first digital tour guides. Cheverest et al. also describe in [5] one of the

first electronic tour guides – GUIDE - and their experiences developing and evaluating it. They point out that "population of the information model with appropriate content was a time consuming task but was absolutely necessary to enable the validation of the system's design and an evaluation of the systems usability, "(p. 30) [5]. REXplorer - more a mobile pervasive game than a normal tour guide – was launched in Regensburg, Germany in the summer of 2007 [2]. Its goal was to combine a tour guide with a pervasive game. "Spell-casting" is REXplorer's main interaction paradigm. In order to support this mode of interaction, a gesture recognition system for camera-based motion was developed for Nokia N70 mobile camera phones. Again Ballagas et al. underlined the importance of good content.

Auto-Generation of Narrative Content

The audio tours for the WikEar system [11] are generated using the Minotour system [6]. First, the closest georeferenced Wikipedia article (*A*) is selected as a starting location for the story. Next, the user chooses a final destination that also corresponds to a georeferenced Wikipedia article (*B*). The Minotour software then chooses the best path through the Wikipedia Article Graph (WAG) between *A* and *B*. The optimality of the path through Wikipedia is defined as that which corresponds best to a narrative theory-based function. Specifically, the algorithm optimizes for three characteristics: (1) a narrative arc from specific to general and back to specific; (2) progression from article *A* through intermediary articles to article *B*; and (3) unity between *A*, *B*, and intermediary articles. After the optimal path is selected, a snippet from article *A*, a series of snippets from the intermediary Wikipedia

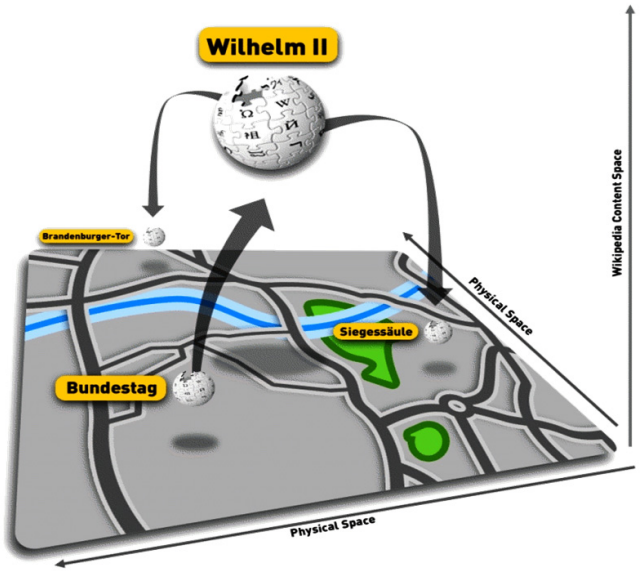

Figure 2: Simplified diagram depicting a location based narrative from Wikear.

articles, and a snippet from article *B* is delivered to the user as they walk from point *A* to *B*. In the WikEar system, the narrative is experienced in audio form (via human narration in our demo unit due to the poor long-term listening pleasure of current text-to-speech software). The user thus receives a tour in a narrative structure from the starting location to the ending location. A sample narrative is still accessible at: ifgi.uni-muenster.de/~j_scho09/wikear.wav, illustrating the walk from the *Kongresshaus* in Innsbruck, Austria to the cathedral. The story's length is adjusted to the walking time (story length := walking time / 2).

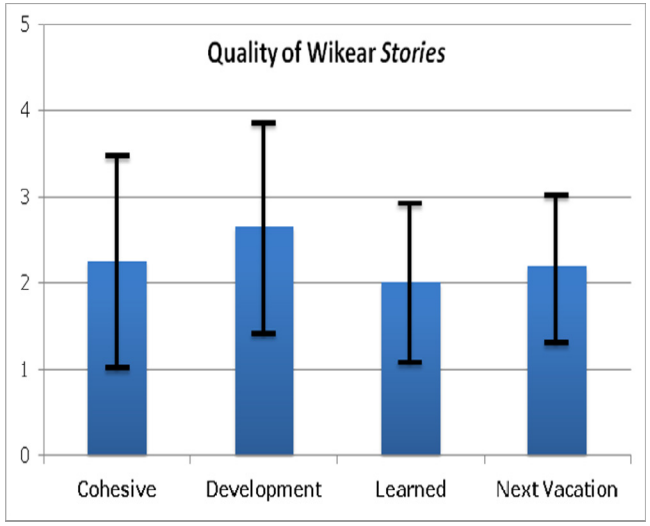

Figure 3: Quality of Wikear Stories.

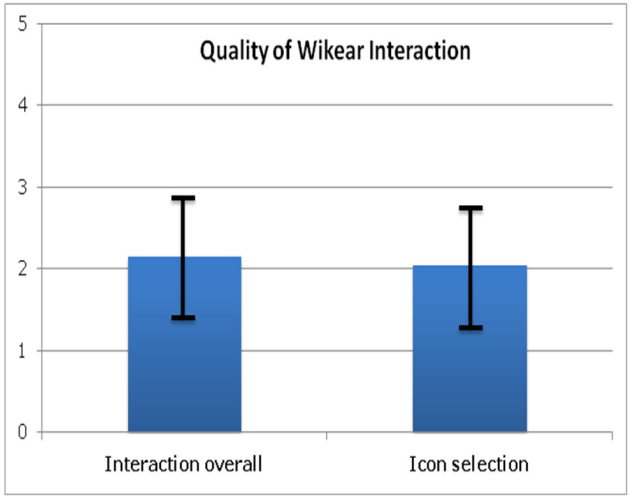

Figure 4: Quality of Wikear Interaction.

Evaluation

In conjunction with the Ninth International Conference on Ubiquitous Computing taking place in Innsbruck in September 2007, we conducted a semi-formal evaluation of the quality of our generated stories and the new interaction metaphor. We demonstrated our WikEar prototype at the conference and asked interested attendees to give us feedback on the stories and the interaction. We distributed maps (A4 size) and testers were given the option of using one of three loaner Symbian S60 (third edition) phones. We set up a web-based evaluation system where users could listen to two stories again and answer the eight following questions with a six-point ranking system (1=good to 6=bad):

- How cohesive was this story? Did all the pieces seem to make sense together? (Ranking: from "*It was unified*" to "*It was completely disjointed*". A "1" on this question indicated that the stories were "unified" and a "6" indicated that the stories were "completely disjoint").

- Was there a sense of development from beginning to middle to end? Did you feel that the story was progressing to a specific point? (Ranking: from "*The story developed as I read. I often thought about what might happen next.*" to "*I felt lost. The story had no clear direction*". Similarly, a "1" on question two indicated that the user found that the "story developed as [they] read" and that they often thought about "what might happen next". A "6" on the same question signaled that the "felt lost" and that the story had no clear direction.)
- Do you think you learned more about Innsbruck than you would have if just using a paper map? (Ranking: from "*Yes*" to "*No*").

- If this service were available for your next vacation, how much would you use it? (Ranking: from "Frequently" to "Never").

- How well did the interaction between the paper map and the camera phone work? (Ranking: from "Very good" to "Very difficult").

- How difficult was icon selection by pointing with the camera phone? (Ranking: from "Very Easy" to "Very difficult").

- How did you like the service? What else would you expect from such a service? (Free text field).

Results and Discussion

We got feedback from 21 conference attendees, three female and 17 male (one participant did not specify gender). All were conference attendees at UbiComp and had a mean age of 34.1 years (SD=6.8). One participant had problems with the evaluation website so we excluded that person from the data that we analyzed. UbiComp attendees are a peculiar population to do an evaluation of UbiComp technology, but we believe this to have had only little effect on our results. The main problem of evaluating auto-generated stories is that these stories take an experimental form. Since most traditional forms of narrative are produced in relationship to cultural, stylistic, and genre conventions, these stories may be evaluated with reference to a norm. Thus when a user encounters a Wikear story, it will likely seem deficient when compared to a traditional literary narrative (if the user even considers it a traditional "narrative.") This is why we had to test the user directly to see if they used narrative schema in comprehending the material. We isolated two distinct indicators of narrative comprehension and tested for these: unity and development. The overall opinion on the quality of the WikEar stories and the level of narrative comprehension is shown in figure 3. The overall opinion on the interaction is described in figure 4. Because the standard deviation on both narrative questions (questions one and two) was so high, very few conclusive results can be drawn. The mean of question one was 2.25, but because the standard deviation was 1.37, the 95 percent margin of error covered nearly the entire range of possible values. The same occurred with question two, which had a mean of 2.65 and a standard deviation of 1.39. We would like to see a more definite result for both unity and development. However, the fact that a large majority of people report positive results of "1s" (3) and "2s" (8) is very encouraging. We conclude that although there are some indicators of success from our survey methodology, we need to make some improvements to better bind the snippets together and to underline the development in the stories. We are currently working on an enhanced version of the narrative generation algorithm (see the future work section) and are also testing interface-based methodologies, such as showing the paragraphs on the screen (see figure 1 bottom). It is important to note that we got the best rating - a mean of 2.02 - for the question about education, an indication our narratives accomplished our pedagogical goal . Additionally, most people would like to take such a system with them on their next vacation, another encouraging statistic. As is shown in figure 4, the user had no problems interacting with the paper map. We also conducted a formal user study to compare the performance of this magic lens approach against two other methods (joystick navigation and a dynamic peephole method without visual context - the map - in

the background) for map navigation with mobile devices. Detailed results can be found in [10].

Conclusion and further Work

The first item of future work is developing a more formal user study around our narrative evaluation methodology discussed above. We are also working on a new version of the Minotour system that we believe will better capture both unity and progression in generated texts. We will be comparing this system with the old system, which was evaluated in this study. Another evaluation idea we will be exploring is the importation of our generated narratives into travel wiki websites. It is our intent to judge the value of the narratives based on the wiki community's reaction to them.

Acknowledgments

We would like to thank Michael Rohs, who helped develop the WikEar prototype and Michael Schellebach for giving us constructive feedback. We are grateful to the Deutsche Telekom Laboratories and the NSF (#DGE-0221713) for partially funding this research.

References

[1] Abowd, G.D., Atkeson, C.G., Hong, J., Long, S., Kooper, R. and Pinkerton, M. Cyberguide: A mobile context-aware tour guide. In *Wireless Networks. 3,10* Springer (1997). 421-433.

[2] Ballagas, R.A, Kratz, S.G, Borchers, J, Yu, E., Walz, S.P., Fuhr, C.O, Hovestadt, L. and Tann, M. REXplorer: a Mobile, Pervasive Spell-casting Game for Tourists. In *Ext. Abstracts CHI 2007*, ACM Press (2007).

[3] Baus, J., Cheverst, K. and Kray, C. A Survey of Map-based Mobile Guides. In *Mobile Guides: Map-based Mobile services-Theories, Methods, and Implementations*. (2004), 197-216.

[4] Brown, B. and Chalmers, M. Tourism and Mobile Technology. In *Proc. of ECSCW 2003*, Kluwer Academic Press (2003).

[5] Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C. Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences. In *Proc. of HCI 2000*, ACM Press (2000), 17-24.

[6] Hecht, B., Starosielski, N., and Dara-Abrams, D. Generating Educational Tourism Narratives from Wikipedia. In *Proc. of* AAI *Fall Symposium on Intelligent Narrative Technologies*, (2007), 37-44.

[7] Isbister, K. and Doyle, P. Web Guide Agents: Narrative Context with Character. *Narrative Intelligence.* John Benjamins Publishing Company (2003), 229-243.

[8] Lanegran, D. Discussion on question, "What makes a good field trip?" With B. Hecht. St. Paul, (2005).

[9] Rohs, M., Schöning, J., Krüger, A. and Hecht, B. Towards - Real-Time Markerless Tracking of Magic Lenses on Paper Maps. In *Adjunct Proc. of Pervasive 2007*. Austrian Computer Society (2007), 69-72.

[10] Rohs, M., Schöning, J., Raubal, M., Essl, G. and Krüger, A. Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context. In *Proc. of ICMI 2007*, ACM Press (2007), 146-153.

[11] Schöning, J., Hecht, B., Rohs, B. and Starosielski, N. WikEar: Automatically Generated Location-Based Audio Stories between Public City Maps. In *Adjunct Proc. of Ubicomp 2007*. (2007), 128-131.

[12] Schöning, J., Hecht, B., Raubal, M., Krüger, A., Marsh, M. and Rohs, M. Improving Interaction with Virtual Globes through Spatial Thinking: Helping users Ask „Why?". In *Proc. of IUI 2007,* ACM Press (2007).