Image-Based Location Awareness and Navigation: Who Cares?

Yung-Hsiang Lu and Edward J. Delp

School of Electrical and Computer Engineering, Purdue University

West Lafayette, Indiana 47907

yunglu@purdue.edu and ace@ecn.purdue.edu

Abstract

In this paper we describe some of the research issues and challenges in image-based location awareness and navigation. We will describe two systems being developed at Purdue University as testbeds for our ideas. The main system architecture combines image processing, mobility, wireless communication, and location awareness.

1 Introduction

Two decades ago, digital images were mostly confined in research labs and military applications. Over the last twenty years, imaging and graphics technology have become indispensable parts of consumer electronics. Apple's Macintosh was introduced in 1984 and brought graphical user interface into consumers' desktops. Video games with high-quality images are parts of many teenagers' lives. Image technology entered a new era in the late 1990s when digital cameras, personal digital assistants (PDAs), and mobile telephones became popular. These portable devices can show color images on LCD displays of several square inches. Meanwhile, wireless data networks (particularly IEEE 802.11 (WiFi) and wide-area networks such as GPRS) are widely deployed. Today, many university campuses, companies, or fast-food stores provide access points for wireless networks. In the early 1990s, US government established satellite-based positioning systems known as global positioning systems (GPS) for military purposes. Civilian applications in GPS-based navigation have grown quickly when higher accuracy became available in May 2000. The rapid advance of image technology, wireless communication, and GPS has provided new opportunities for research in image acquisition, processing, analysis, interpretation, and presentation.

(a)   (b)   (c)

Figure 1: (a) Navigation PDA. (b) GPS for vehicles. (c) Handheld GPS.

Among all possible applications that can benefit from the combination of portable image processing and wireless networks, we believe image-based navigation has a great potential to become a main-stream application in portable electronics. This prediction is based on the following trends: (a) World travel for leisure and business continues growing. One of the fastest growing regions is China. In 2001 there were more than 10 million tourists from China and the China National Tourism Administration projected this number will exceed 100 million by the year 2020. Portable navigation systems can be very helpful to tourists. The need for navigation systems for business travelers is obvious. (b) The public continues demanding better technology to reduce traffic accidents. One major factor in accidents is driver distraction, such as looking for street signs or maps. Fig-

ure 1 shows three commercial map-based navigation systems. Various studies have shown that maps are not intuitive for many people. Hence, these systems may actually become the sources of distraction and create dangers on the road. (c) Road conditions are frequently changing due to construction, accidents, or weather. Knowing a route is insufficient to safely reach one's destination. This is the reason a driver is needed to respond to road conditions. Technology can improve safety; for example, front infrared sensors can detect pedestrians at night and avoid collisions. More sophisticated image technology may provide assistance in detecting obstacles or lane-keeping.

This paper describes how image-based location awareness and navigation can address some of the above issues. We survey existing technology for navigation and describe the opportunities in using imaging technology in navigation. We describe a research project at Purdue University[1] to build testbeds for this technology [9].

2 Previous Work

Commercial navigation systems provide three common ways to help a user determine their current location: (a) longitude, latitude, and elevation, (b) waypoint marks, and (c) maps. For most people, (a) is meaningless because people do not navigate based on longitude or latitude. Instead, people navigate through roads and remember landmarks. Also, (b) requires the user to have been to the current location and marked a waypoint. As a result, maps become the dominant method for navigation.

Maps have many advantages. First, good maps reflect the geographic environments: the relative locations and distances among objects. Second, maps may encode multiple types of information associated with the geographical locations [13]. For example, cyan indicates rivers; circles mean cities; dark green is mountains. Third, maps are storage-efficient. The street maps of the entire United States can fit on a CDROM published by DeLorme. Fourth, paper maps are light-weight and easy to carry; they can be folded and put into a pocket. Fifth, electronic maps can be zoomed to show different levels of details. Finally,

maps do not require substantial rendering capabilities and can be presented easily on a portable device such as PDA [8].

On the other hand, several studies have demonstrated that maps are not the best way for route instruction and navigation. Maps are particularly problematic when they are displayed on palm-size LCD monitors, not on table-size paper. Ekman et al. [3] performed user studies and discovered that users preferred to have contextual information in small displays of mobile systems. They also found that active guidance using arrows could prevent users from making mistakes during navigation.

Satalich [12] surveyed many user studies in navigation and wayfinding and reached the following conclusions: (a) Reading maps alone was insufficient. Additional information could help users recognize location sequences, landmarks, and routes. (b) Even with customized maps, users could still make many errors in wayfinding. (c) "You Are Here" maps in buildings did not significantly reduce the time to find rooms.

Kray et al. [8] compared four different ways to present route instructions: textual, abstract direction (arrows), 2-D maps, and computer-rendered 3-D maps. They found the most commonly used strategy was matching buildings in the 3-D maps, especially for texture-shaded buildings. Their study also found that users preferred a flying view at 25 meter in altitude. An unexpected result was that users needed more time for orientation when they used 3-D maps, compared with 2-D maps.

Vainio et al. [14] reported that users preferred photo-realistic 3-D models so that landmarks could be recognized more easily even when a user was moving. Gronbak et al. [5] built spatial hypermedia to represent buildings for architecture design. Baus et al. [1] discussed designing navigation systems for users that combined several means of transportation to reach their destinations. Rogers et al. [11] used GPS information to provide safety advices such as lane keeping. Chan et al. [2] built a visualization system for 3-D geographical data. Jordan [7] predicted that future "intelligent maps" would include realistic images.

Most of the existing work generates navigation images using graphics techniques. Some 3-D images were created for virtual reality environments. This

approach encountered two challenges: First, it is impractical to generate polygon models for an entire city. Second, portable computers may not have the computation capability to render photo-realistic images fast enough. As a result, we believe navigation images should be taken directly from the streets, not computer generated.

"Image-based navigation" is an active research topic in robotics and computer vision, such as [4, 6, 10]. These robots used image sensors to detect obstacles or to recognize objects. Our problem is different from robot navigation. First, images are used to assist humans and they— not robots — are responsible for recognizing landmarks and avoiding collision. Second, these users can move significantly faster when they drive. This put a much tighter timing constraint in updating the images. In contrast, self-navigated robots move significantly slower. Third, we can assume maps (roads and intersections) are available to users while many research activities in robotics and vision intend to build maps from the image sensors.

3   Scenarios: *LAID* and *IEN*

This section describes the scenarios to illustrate image-based navigation. The first is the ability to use locally acquired images to determine the identity or function of an object that one observes as one roams in an area. For example, a user is walking through a university campus and comes to a building that is unknown to the user. The user takes a picture of the building with a handheld device. This image will be sent via the wireless network to a "location aware" image database. The database will index the image against its content and send back information to the user about the building, such as the name, use, and type of architecture. We shall refer to this scenario as using a Location-Aware Image Database or *LAID*.

The basic building blocks of the *LAID* system is a handheld client that has a camera, a GPS receiver, a compass, and network connection. The user connects via the network to a server that has location-aware image database and sends the image of the object. The server extracts content-dependent features from the image and uses this along with the location and orientation information to find the best match object in the database. This information along with the matched image (possibly a thumbnail image) is
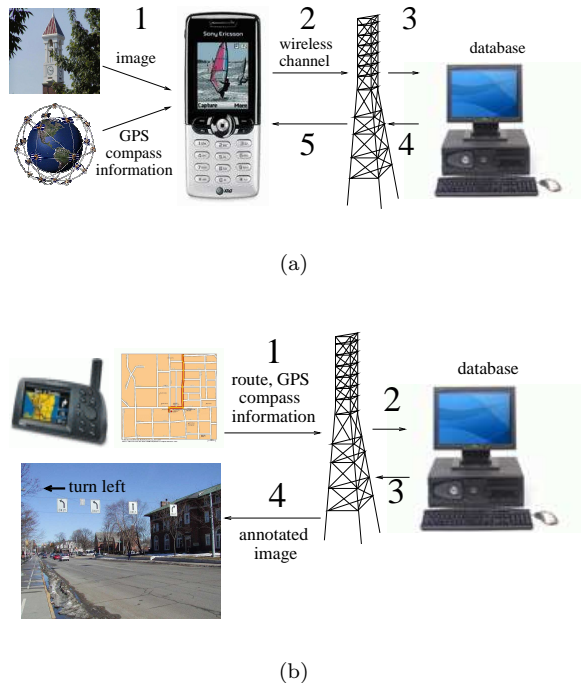

(a)

(b)

Figure 2: (a) Location-aware image database system (*LAID*) (b) Image-based navigation (*IEN*).

returned via the network to the user.

The second scenario is using images in a database to aid in vehicle navigation. The ultimate goal here is to provide the driver with an image of a recognizable object (for example, a building or a landmark) that can help the driver determine where the driver is located. The main principle here is that people are better at recognizing objects to determine their locations than reading maps particularly when they believe they are lost and may be disoriented. We refer to this as Image Enhanced Navigation or *IEN*. Images used in the *IEN* system are annotated with navigation instructions. The driver only needs to take a glance of the image to confirm the current direction. Voice-assisted GPS and the visual direction ensure that the driver follow the route correctly.

The *IEN* system consists of a GPS receiver, a map database, a route planner, a display, and wireless network connection that is located in the vehicle. The components of the *IEN* system that are not contained in currently available on-board navigation systems are the image database and the wireless connectivity.

Before a road trip, a user plans the route and downloads the associated images along the route. These images, in addition to maps, help the user navigate to the destination. While the user is moving, the system automatically updates the images periodically. Each image has information about its location, orientation, time of the day, and the weather. These features provide direct and intuitive visual information.

4 Research Issues

4.1 Database Construction

The first challenge of the system is to populate the database with location-annotated images. We are currently constructing a database with the images taken on the West Lafayette campus of Purdue University. When an image is taken, its location and orientation are recorded based on the GPS and compass readings. As the system scales up, an automatic method must be developed to construct the database efficiently. The city of West Lafayette has more than 100 miles of streets. In order to provide visual directions for navigation, four images are needed in each intersection. Suppose there are ten intersections for each mile of road, over 4000 images must be taken with their location information.

We are designing an automatic method to populate the database. One solution is using four webcams mounted on the roof of a vehicle. The webcams face four different directions and they are connected to a laptop computer in the vehicle. The webcams take images periodically. If a camera takes an image every five seconds, 24 images will be taken by one camera for each mile when the vehicle travels at 30 miles per hour. The laptop computer is connected to a GPS receiver and a digital compass. When the cameras take images, the laptop computer also record the current location and direction. We need only the direction of the vehicle and can compute the cameras' directions easily.

This approach can quickly populate the image database with another benefit: By comparing the images in the sequence, we can determine the occlusion relationships among objects. We can calculate which objects should appear first when viewed at a particular location and direction. This information will improve the matching accuracy of *LAID*.

4.2 Image Comparison and Background Adjustment

Another challenge in *LAID* is that the images sent by clients may be taken with angles or background lighting different from the images stored in the database. While human beings usually have no problem recognizing these images, it can be difficult for the database. Since each image contains the information of its location and direction, the search accuracy can be significantly improved. Moreover, the occlusion relationship obtained during image capturing can further assist enhancing the matching. This problem will have to be addressed in the way we extract features from the received image. Even though, in a general sense, we are doing "query-by-image-content", the location information will help us solve this problem.

The angle and lighting problem for *IEN* is the opposite. It is preferred to present the images with lighting and the angle matching what the user should be seeing. For example, if an image is taken at daytime and the user is driving at night, the images should be adjusted to match the night view. There are two solutions to this problem. First, we can use the sequence of images to separate background (sky) and objects because objects move faster in the sequence due to perspective projection on the camera. After separating the background from the objects, we can change the colors of the background to reflect the current time of the day. The second solution is to capture multiple images at different time of a day or in different seasons. If *IEN* is widely deployed, image capture may be achieved through the collaboration with car rental companies. Because *IEN* helps drivers navigate more comfortably, it has the potential to reduce accidents caused by nervousness due to feeling lost. Consequently, *IEN* may help rental companies reduce their cost due to accidents. Some rental companies consider using GPS to track their vehicles for better roadside assistance. These cars may also be equipped with digital cameras and capture images along the routes the customers travel. The privacy issue here are important but beyond the scope of this paper.

5 Conclusion

This paper describes how to use images to assist users in navigation. We have presented the user scenarios of using *LAID* and *IEN*. These systems may greatly improve the services for mobile users. These two systems require further studies and improvement in imaging technology. We also provide potential solutions to some of the problems.

References

[1] J. Baus, A. Kruger, and W. Wahlster. A Resource-adaptive Mobile Navigation System. In *International Conference on Intelligent User Interfaces*, pages 15–22, 2002.

[2] C. S. L. Chan, T. K. Y. Chan, and E. C. Prakash. A System for Geographical and Spatial Data Exploration on The Internet. In *Pan-Sydney Workshop on Visualisation*, pages 23–29, 2001.

[3] I. Ekman and P. Lankoski. What Should It Do? Key Isssues in Navigation Interface Design for Small Screen Devices. In *Conference on Human Factors and Computing Systems*, pages 622–623, 2002.

[4] R. Ghurchian, T. Takahashi, Z. D. Wang, and E. Nakano. On Robot Self-navigation in Outdoor Environments by Color Image Processing. In *International Conference on Control, Automation, Robotics and Vision*, pages 625–630, 2002.

[5] K. Gronbak, P. P. Vestergaard, and P. Orbak. Towards Geo-spatial Hypermedia: Concepts and Prototype Implementation. In *Conference on Hypertext and Hypermedia*, pages 117–126, 2002.

[6] J. Hong, X. Tan, B. Pinette, R. Weiss, and E. M. Riseman. Image-based Homing. *IEEE Control Systems Magazine*, 12(1):38–45, February 1992.

[7] L. E. Jordan. Dawn of A New Era: The Image Information Age. In *Geoscience and Remote Sensing Symposium*, pages 2228–2230, 1996.

[8] C. Kray, C. Elting, K. Laakso, and V. Coors. Presenting Route Instructions on Mobile Devices. In *International Conference on Intelligent User Interfaces*, pages 117–124, 2003.

[9] Y.-H. Lu and E. J. Delp. An Overview of Problems in Image-Based Location Awareness and Navigation. In *Visual Communications and Image Processing*, pages 102–109, 2004.

[10] E. Rivlin, I. Shimshoni, and E. Smolyar. Image-based Robot Navigation in Unknown Indoor Environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2736–2742, 2003.

[11] S. Rogers, P. Langley, and C. Wilson. Mining GPS Data To Augment Road Models. In *International Conference on Knowledge Discovery and Data Mining*, pages 104–113, 1999.

[12] G. A. Satalich. Navigation and Wayfinding in Virtual Reality: Finding Proper Tools and Cues To Enhance Navigation Awareness. Master's thesis, University of Washington, 1995.

[13] E. R. Tufte. *Envisioning Information*. Graphics Press, 1990.

[14] T. Vainio and O. Kotala. Developing 3D Information Systems for Mobile Users: Some Usability Issues. In *Nordic Conference on Human-computer Interaction*, pages 231–234, 2002.