

# Multimedia for Mobile Environment: Image Enhanced Navigation

Shantanu Gautam, Gabi Sarkis, Edwin Tjandranegara, Evan Zelkowitz,  
Yung-Hsiang Lu\*, and Edward J. Delp

School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana

## Abstract

As mobile systems (such as laptops and mobile telephones) continue growing, navigation assistance and location-based services are becoming increasingly important. Existing technology allow mobile users to access Internet services (e.g. email and web surfing), simple multimedia services (e.g. music and video clips), and make telephone calls. However, the potential of advanced multimedia services has not been fully developed, especially multimedia for navigation or location based services.

At Purdue University, we are developing an image database, known as LAID, in which every image is annotated with its location, compass heading, acquisition time, and weather conditions. LAID can be used to study several types of navigation problems: A mobile user can take an image and transmit the image to the LAID sever. The server compares the image with the images stored in the database to determine where the user is located. We refer to this as the “forward” navigation problem. The second type of problem is to provide a “virtual tour on demand.” A user inputs a starting and an ending addresses and LAID retrieves the images along a route that connects the two addresses. This is a generalization of route planning. Our database currently contains over 20000 images and covers approximately 25% of the city of West Lafayette, Indiana.

**keywords:** multimedia, navigation, GPS, image processing

## 1 Introduction

As world-wide travel continues growing, navigation assistance for mobile users becomes increasingly important. Concurrently, broadband wireless networks are being widely deployed. Existing technology can allow mobile users to check email, trade stocks, watch sports, or make phone calls. We believe the potential of multimedia has not been fully explored for mobile users, especially multimedia for navigation. We use two user scenarios to demonstrate how multimedia may help future travelers. In the first example, a family takes a road trip. As the vehicle is moving, they can preview the route. One of the family members says to the driver, “In one minute, we will *see* a gas station on the left.” “In two minutes, we will *see* Burger King on the right.” “You need to turn left when you *see* the second traffic light.” “We will *see* a shopping mall and I want to *see* which items are on sale.” Existing navigation technology is mostly map-based: users read maps and cannot *see* the route. Some studies have shown that maps are not intuitive to many people and people usually navigate based on direct visual cues, i.e., images. In the second example, an airline passenger uses the personal display to navigate the destination city. The passenger can *see* the route from the airport gate to the rental car company and from the company to the hotel. The passenger discovers several landmarks along the route. On the same airplane, another passenger wants to *see* a city that is not included in this trip. The passenger can select a city and drive-through the city virtually and *see* real images taken in the city. The two scenarios share three important common requirements. First, they need high-density (spatial) high-quality images to provide a realistic drive through experience. Second, they are personalized based on

---

\*Please address all questions to Yung-Hsiang Lu, yunglu@purdue.edu.

the users' interests and destination; existing pre-recorded virtual tours are insufficient. Third, the navigation experience is interactive; the users may decide to change routes and the virtual tour has to respond correctly. We envision that a future navigation system will be multimedia based. In addition to maps and voice guidance, users can *see* the landmarks, *see* the routes using a bird's eye view, *see* the buildings and stores along the routes, visit the web sites of the stores or tourist attractions along the routes, or preview several routes and choose a route with scenery. A foundation of such multimedia-based navigation is a densely populated image database in which every image is annotated with the location and compass heading. Image sequences can be created on-demand to show the routes for different users.

This paper presents a location-aware image database (LAID) being built at Purdue University for the city of West Lafayette. In order to provide high-quality image sequences for navigation, we mount two 3.2 million pixels (MP,  $2048 \times 1536$ ) cameras on the roof of a sedan (at height of 1.4 meters) to acquire images. Digital cameras are used due to their higher resolution as most digital camcorders provide only 340K pixels. In the continuous-acquiring mode, each camera can obtain one image approximately every 1.5 seconds. Each image is associated with the GPS location of the camera, the time the image is taken, the date, and the weather condition. The city of West Lafayette has approximately 67,000 residents, including 38,000 Purdue students. Its location is ( $86^{\circ}52'W$ ,  $40^{\circ}21'N$ ) with 188 meters of elevation. The city covers  $14.2 \text{ KM}^2$  of area and has 100 KM of streets. If the vehicle moves at 18KM/H, each camera obtains one image every 7.5 M. For a typical driver in the city at 36 KM/H, this image density provides a frame rate of 1.3 frames per second. This frame rate is sufficient to provide smooth visual cues. LAID can be extended to many related projects. For example, we can use images for two parallel cameras (stereo view) or a sequence of images to calculate the location of an object in the images. It is also possible to provide localization without GPS: finding a user's current location from the objects seen by the user without using GPS signals. This application can be useful when GPS signals are unavailable. This paper focuses on the construction of the image database and reports the current status of our image-based navigation. This paper presents an implementation of the concepts described in our previous publications [19] [20].

## 2 Background

### 2.1 Navigation Assistance

Navigation assistance is one of the fastest growing areas in portable electronics. Vehicle-mounted as well as handheld GPS (global positioning system) terminals are expected to grow 31% per year until 2009 (ABI Research) and the market will reach \$50 billion dollars before 2010. IEEE Signal Processing Magazine has a special issue called "Location Is Everything: Positioning in Wireless Networks" published in July 2005. When a navigation system has wireless connection, it is possible to provide the most up-to-date traffic information and advise drivers to avoid congested areas. Many cities have announced plans to provide complete and free WiFi services, including Long Beach in California, Philadelphia, San Francisco, and Taipei.

All existing GPS systems, however, rely on an indirect way to guide users by using maps. Unfortunately, maps are abstract and do not provide direct visual cues. Figure 1 shows three commercial GPS units and all of them use maps. Vehicular GPS units usually also provide audio instructions, such as "Turn right in 100 meters." or "Turn left at the Fifth Street." These instructions have several disadvantages. First, it is difficult to estimate the distance of 100 meters especially while driving. Second, the street signs are often too small to see until the driver is very close to the intersection. We believe that multimedia can be applied to navigation systems and help users (including drivers, passengers, and pedestrians) recognize landmarks more easily and improve safety. A multimedia navigation system provides different types of information to assist users in finding their current locations and destinations. An image-based virtual tour helps the users "drive through" (or walk through) with actual landmarks and buildings along the routes. The navigation system can use wireless networks and be connected to existing navigation web sites, such as Google Maps to obtain aerial views.

Even though maps are the most popular form of navigation, maps are abstractions of actual objects.

Hence, reading maps requires mental efforts and may be a distraction for drivers. Several studies have been conducted to present navigational guides other than traditional street maps. Bessa et al. [4] point out the limitation of existing map-based navigation; they propose using 3-D images to assist users in finding their locations. Baus et al. [3] discuss designing navigation systems for users that combined several means of transportation, including driving and walking, to reach their destinations. Ekman et al. [7] discuss how to design the user interface for navigation systems. They perform user tests of a palm-size and a wrist-held systems. Hornbaek et al. [10] perform user study of reading maps with and without zoomable interfaces and compare their response in navigation. Even though users generally prefer an interface with overview, there is no convincing support that such an interface can help users navigate faster. Jost et al. [12] evaluate user responses to a multimodal system that provides zoomable maps and tour planning. They discover that users respond differently for indoor and outdoor applications. Lee et al. [16] study how people read maps and observe the demand of drivers' attention by existing navigation systems. Li et al. [17] expect the rapid growth of location-based application and build a tool for prototyping such applications. Rogers et al. [23] combine GPS and lane models and develop a set of land-level navigation assistance, such as keeping and choosing a lane. Satalich's master thesis [24] discusses the mental process of navigation and wayfinding; in the process, landmark recognition is a crucial step. The study presents the following conclusions: (a) Reading maps alone is insufficient. Additional information can help users recognize location sequences, landmarks, and routes. (b) Even with customized maps, users can still make many errors in wayfinding. (c) "You Are Here" maps in buildings do not have significant benefit for reducing the time to find rooms. (d) Virtual reality is less effective in helping users navigate than the experience in real environments. Warren et al. [31] use portable music devices for navigation by adjusting the volume in each side of the earphone. Unlike existing technology using audio navigation commands, their system embeds navigation hints without disrupting the music. Wijk et al. [30] develop a model for handling smooth zooming for navigating on a large 2-dimensional space, such as a US map. Zhao et al. [33] mount a laser range on a vehicle to acquire the structures of buildings. Kruger et al. [14] construct a navigation system that integrates desktop computers, PDAs, and in-vehicle navigation systems for both drivers and pedestrians.

## 2.2 Navigation Using Images and Videos

Kray et al. [13] compare four different ways to present route instructions: textual, arrows, 2-D maps, and computer-rendered 3-D maps. They find the most commonly used strategy is matching buildings in the 3-D maps, especially for texture-shaded buildings. Vainio et al. [29] report that users prefer photo-realistic 3-D models so that landmarks can be recognized more easily when a user is moving. Using images in traffic management has been studied by some researchers. Angel et al. [2] develop a system that can automatically detect and track vehicles with a sequence of aerial images acquired from a helicopter. Chang et al. [5] develop automatic license plate recognition for enforcing traffic laws, toll collection, and managing parking lots. Chen



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

et al. [6] use image sequences to detect overhead structures because existing radar-based collision warning produces false alarms by failing to distinguish an overhead structure from a stalled vehicle. Ghurchian et al. [8] use video input to guide a robot navigate along a road that is not well-paved; their method is based on differentiating the colors along the road and the surrounding obstacles. Gronbak et al. [9] develop a hypermedia system combining GPS and GIS and present augmented reality for building construction. Jirawimut et al. [11] use stereo cameras as an odometer for blind or visually impaired pedestrians. Liang et al. [18] improve the stability of the images captured by video cameras mounted on vehicles for intelligent driver-assistance systems. Miura et al. [22] use computer vision to assist drivers in responding to traffic conditions. Schafer et al. [25] use multiple formats of information for navigation. Their user interface includes texts, images of the landmarks, and maps that indicate the location of other users. A company called **a9.com** is creating an image database for Yellow Pages. The database includes the front views of stores in 10 US cities. The World Wide Media eXchange (WWMX) project [28] adds location tags to images and these locations are marked on maps. Users can browse the thumbnails attached to the maps. A fundamental difference between WWMX and our work is the density of images. Our image database contains densely populated images and allows us to study several issues that require such a high density of images. First, our database allows continuous virtual drive through in the city. Second, because of the high density, some buildings appear in multiple images and we can extract the structure information of these buildings using the techniques presented in [1] [27]. Third, we can calculate the location of an object in an image. Both WWMX and our project record the location of the camera, not the location of the object (or objects) in the image. Determining the location of the object from a single image is difficult because an object, especially a landmark, can be seen at a great distance. Also, an image may contain many objects at different locations and it is not possible to use a single location for all objects. By using techniques similar to extracting building structures, we can estimate the distance of an object from the camera and determine the object's location.

### 2.3 Three-Dimensional Modeling

Some recent projects construct 3-D object models for navigation. For example, Google Earth uses aerial images to construct building models; Google also provides satellite images for navigation. Google Maps and Google Earth has mapped most of USA cities and the major cities of the entire world by aerial satellite images. Locations of restaurants, lodging accommodations, roads and borders are also visible from a birds' eye view and these locations are marked on the map itself. Some major cities have 3-D reconstructed images of buildings, such as the Chicago downtown. One major disadvantage of using aerial or satellite images is that a user on the ground cannot match the images directly with the scene observed by the user. Our work uses images taken at the eye height (1.4 meter, on the roof of a car) so the images can be matched directly with the user's perception. Kwon et al. [15] use two rotating cameras to form stereo views to determine the distance of an object. Our approach is different in the following ways. First, the cameras are mounted on the same roof rack so the cameras' axes are parallel and do not intersect. Second, we do not select the object before capturing images in order to calculate the object's distance from the camera. Instead, we capture the images of streets and buildings first and later match the objects among the images taken at the same and nearby locations. Our system is more informative because it provides real images of the location seen at eye level. Landmark locations may be found from GPS coordinates of images. The system provides a closer to real-life navigation system.

## 3 Location-Aware Image Database (LAID)

The center of multimedia-based navigation is a location-aware image database (LAID). Figure 2 illustrates how to use LAID to create the different contents described in Section 1 for navigation. In the database, each image has the location and the compass heading of the camera. We use the database and a commercial map and navigation program to generate a virtual tour on demand using the images from the database. We

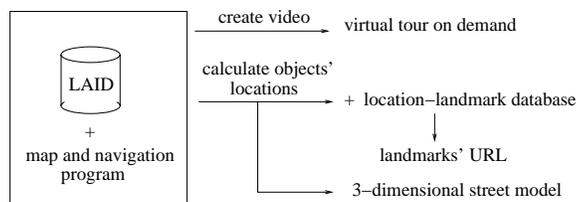


Figure 2: Overview of the navigation system using LAID

can calculate the location of the objects in the images. Using the locations, we can query the map program to determine the names of the buildings, stores, or landmarks and find the URLs of the web sites for these objects.

### 3.1 Image Capturing System



Figure 3: Two cameras are mounted on the roof rack of a sedan. (a) (b) The directions of the two cameras are parallel. (c) The directions are  $45^\circ$  apart. (d) The directions are  $90^\circ$ .

We mounted two digital cameras (Canon PowerShot A510) on the roof rack of a sedan, as shown in Figure 3. This camera was chosen because it can continuously take photographs at a rate of approximately one image every 1.5 seconds and store them to a 1 GB Secure Digital card. We did not use a camera with built-in GPS because the product reviews were mixed (especially the shorter battery life); also, we were unable to find a vendor with such a camera in stock. Webcams were not chosen due to their low resolution and image quality. Timestamps are used to calculate the location of the camera when an image is taken. Before we capture the images, the camera time is synchronized with an NIST Internet time server. A handheld GPS receiver is connected to a laptop computer that is also synchronized with the NIST server. When an image is taken, the EXIF timestamp of the image is recorded with the JPEG file. Meanwhile, the GPS receiver continuously sends the GPS location to the laptop. Each location is also associated with a timestamp from the laptop. We match each image's timestamp and the location's timestamp to calculate the location of the camera when the image was taken. If the vehicle moves at 18 KM/H, this procedure can introduce at most 15 meters of errors from three sources. The first is clock drifting between the camera and the laptop after synchronization. The second comes from the one-second accuracy of Internet time. The third is the GPS coordinates and the update rate. The update rate is one second of time so the coordinates recorded in the laptop are accurate to only one second. Overall, the three sources can introduce at most 3 seconds of difference, equivalent to 15 meters of error at 18 KM/H. The error from the GPS receiver is usually within 10 meters. Therefore, the total maximum error is within 25 meters. The images are stored in a database using MySQL with the locations as indexes.

### 3.2 Image Retrieval



Figure 4: A sequence of images in a virtual tour along Columbia Street in Downtown Lafayette, Indiana.

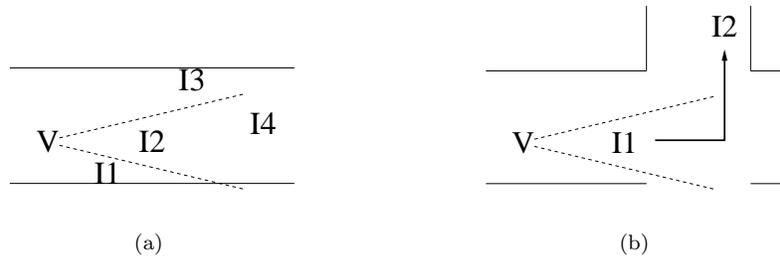


Figure 5: (a) If multiple images are taken along a route, a virtual tour selects images that fall within a viewing angle (dash line) of the vehicle. (b) A turn is needed in a virtual tour but this turn is not taken during image capture.

Each acquired image includes its location, time of capture, compass heading, and weather as the image's metadata. The metadata of an image is stored in MySQL to facilitate image search. When a user inputs a starting address and an ending address, LAID sends a query to Microsoft's MapPoint Web Service to obtain the GPS coordinates of the route, including the locations of turns. The result from MapPoint is used to look up the images in our database for creating a virtual tour on demand. Figure 4 shows a sequence of images along a route as an example. Important information includes (a) the street sign of State Road 26 in the second image and (b) a right-turn-only lane after passing the sign. Such information is not available in any existing navigation system. Three scenarios may occur while finding images along the route. The first scenario is the simplest: a unique image matches the GPS coordinate of a location along the route. This



Figure 6: Screenshots of the LAID user interface. (b) shows a turn at an intersection.

occurs when the route is traveled only once during image capture and the traveled direction is the same as the route. The second scenario occurs when multiple images are available along the same route. This can happen if the same street is traveled multiple times during image capture. The third scenario occurs when no image is available because the route requires a turn but this turn is not taken during image capture.

Figure 5(a) illustrates how to select an image when multiple images are available. In this example, “V” is the virtual vehicle’s current location (the current image), the next image is selected based on its relative location. Images I1 and I3 are outside the view angle of the virtual tour so I1 and I3 are not selected. These two images may be acquired because the street has multiple lanes. If I1 or I3 is chosen, the virtual tour shows rapid horizontal movement between adjacent images and may cause feelings of motion sickness. Within the viewing angle, both I2 and I4 may be selected if their timestamps are close to the timestamp of the current image for V. If I2’s timestamp is different from V’s timestamp by more than one minute, I2 and V are taken in two trips and the environment lighting (weather and shadows) can be different. Figure 5(b) shows how to handle a turn that is not taken during image capture. When this occurs, the image sequence from the two streets are concatenated to create the virtual tour and image I2 follows I1 directly. Figure 6 are screenshots of the LAID user interface. The interface shows the current location, the speed of the tour, an image, and a map on the right side. Each circle on the map indicates an image to be shown in the virtual tour. In order to alleviate the perception of an abrupt change during turning, arrow signs are added as shown in Figure 6(b).

### 3.3 Linking Landmarks and Images

The LAID system is used for helping users in navigation so it is important to highlight landmarks in the images. Since each image is annotated with the camera’s location, we need to calculate the location of a landmark in the image. This requires three steps: (a) identifying the landmark in the image (b) finding the landmark’s location (c) linking the image with the landmark’s information using the location information. When an image is captured, the GPS location is the camera’s location. Using a pair of stereo images or a sequence of images, it is possible to calculate the distance of an object in the images [21, 32]. We use stereo images to calculate landmarks’ locations.

Identifying a landmark in an image is non-trivial; finding the same landmark in two or more images is even more difficult. Thus, we use a semiautomatic approach. We manually select the same landmark

in multiple images and then use match algorithms to find the landmark in the selected region. Figure 7 shows an example. The landmark at the stereo images' center is selected and then our algorithm detects the object enclosed in the selected regions and computes the center of the object, similar to the match algorithm described in [26]. After finding the center of the object, our system computes the distance of the object from the cameras' current locations. With the cameras' locations, their compass heading, and the distance of the landmark, the landmark's GPS location can be obtained and the GPS location is used to find the address. A link of the landmark's web site is added to the enclosed regions in the images so that users can click on the landmark in the image to retrieve more information about the landmark. Reverse geo-coding is performed on the GPS coordinates using Microsoft Mappoint Web Service to locate a landmark name or a street address which corresponds to that specific location. The search results can then be used to query a search engine such as Google or Yahoo and retrieve a website which contains information about the landmark.

Using images to find landmark is a direct and intuitive method. Meanwhile, several reasons may cause errors in calculating the camera to object distance. First, the angle between the two cameras' directions must be known. A small error in the directions can lead to large errors for distant landmarks. Second, the two cameras have to acquire images at the same time (i.e. synchronized) since the vehicle is moving. Otherwise, the two images are taken at different locations. Third, the cameras have finite number of pixels ( $2048 \times 1536$ ) and this leads to quantization errors. Raising the resolution can improve accuracy so we use two 3.2MP cameras instead of two 340KP camcorders. Quantization errors become more serious for distant landmarks because the calculation become more sensitive to small changes in the images. Finally, the object matching algorithm may not find the same corresponding point in the two images.

### 3.4 Current Status of LAID

We have been building a prototype of LAID for the cities of West Lafayette and Lafayette, Indiana. We have captured nearly 20000 images to cover approximately 25% of West Lafayette and two major streets in Lafayette. LAID also contains the downtown of Indianapolis, Indiana in order to study how tall buildings block GPS signals and reduce the accuracy of location information. We continue acquiring images while testing and improving the functionality and accuracy of the system.

## 4 Conclusion

This paper presents the concept of navigation using images. We explain a prototype system, LAID, being built for West Lafayette, Indiana. LAID can be used to provide images at specific locations and help users



Figure 7: Use a pair of stereo images to calculate the location of an object. The center of the two images is the Soldiers' and Sailors' Monument in Indianapolis, Indiana

with direct visual cues. For future work, we plan to conduct user studies of our prototype to understand users' perception of such an image-enhanced navigation system. We are also investigating how to improve the accuracy in calculating landmarks' locations from images.

## 5 Acknowledgments

This project is supported in part by Intel and Purdue Research Foundation. Gautam, Sarkis, and Zelkowitz are supported by Purdue SURF (Summer Undergraduate Research Fellowship) in summer 2005.

## References

- [1] P. M. Q. Aguiar and J. M. F. Moura. Three Dimensional Modeling from Two-Dimensional Video. *IEEE Transactions on Image Processing*, 10(10):1541–1551, October 2001.
- [2] A. Angel, M. Hickman, P. Mirchandani, and D. Chandnani. Methods of Analyzing Traffic Imagery Collected From Aerial Platforms. *IEEE Transactions on Intelligent Transportation Systems*, 4(2):99–107, June 2003.
- [3] J. Baus, A. Kruger, and W. Wahlster. A Resource-adaptive Mobile Navigation System. In *International Conference on Intelligent User Interfaces*, pages 15–22, 2002.
- [4] M. Bessa, A. Coelho, and A. Chalmers. Alternate Feature Location for Rapid Navigation using a 3D Map on a Mobile Device. In *International Conference on Mobile and Ubiquitous Multimedia*, pages 5–9, 2004.
- [5] S.-L. Chang, L.-S. Chen, Y.-C. Chung, and S.-W. Chen. Automatic License Plate Recognition. *IEEE Transactions on Intelligent Transportation Systems*, 5(1):42–53, March 2004.
- [6] Y. Chen. Highway Overhead Structure Detection Using Video Image Sequences. *IEEE Transactions on Intelligent Transportation Systems*, 4(2):67–77, June 2003.
- [7] I. Ekman and P. Lankoski. What Should It Do? Key Issues in Navigation Interface Design for Small Screen Devices. In *Conference on Human Factors and Computing Systems*, pages 622–623, 2002.
- [8] 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.
- [9] 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.
- [10] K. Hornbaek, B. B. Bederson, and C. Plaisant. Navigation Patterns and Usability of Zoomable User Interfaces With and Without An Overview. *ACM Transactions on Computer-Human Interaction*, 9(4):362–389, December 2002.
- [11] R. Jirawimut, S. Prakoonwit, F. Cecelja, and W. Balachandran. Visual Odometer for Pedestrian Navigation. *IEEE Transactions on Instrumentation and Measurement*, 52(4):1166–1173, August 2003.
- [12] M. Jost, J. Haubler, M. Merdes, and R. Malaka. Multimodal Interaction for Pedestrians: An Evaluation Study. In *International Conference on Intelligent User Interfaces*, pages 59–66, 2005.
- [13] 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.

- [14] A. Kruger, A. Butz, C. Muller, C. Stahi, R. Wasinger, K.-E. Steinberg, and A. Dirschl. The Connected User Interface: Realizing a Personal Situated Navigation Service. In *International Conference on Intelligent User Interfaces*, pages 161–168, 2004.
- [15] H. Kwon, Y. Yoon, J. Park, and A. C. Kak. Person tracking with a mobile robot using two uncalibrated independently moving cameras. In *IEEE International Conference on Robotics and Automation*, 2005.
- [16] J. Lee, J. Forlizzi, and S. E. Hudson. Studying the Effectiveness of MOVE: A Contextually Optimized In-Vehicle Navigation System. In *Conference on Human Factors in Computing Systems*, pages 571–580, 2005.
- [17] Y. Li, J. I. Hong, and J. A. Landay. Topiary: A Tool for Prototyping Location-Enhanced Applications. In *ACM symposium on User Interface Software and Technology*, pages 217–226, 2004.
- [18] Y.-M. Liang, H.-R. Tyan, S.-L. Chang, H.-Y. M. Liao, and S.-W. Chen. Video Stabilization for a Camcorder Mounted on a Moving Vehicle. *IEEE Transactions on Vehicular Technology*, 53(6):1636–1648, November 2004.
- [19] 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.
- [20] Y.-H. Lu and E. J. Delp. Image-Based Location Awareness and Navigation: Who Cares? In *Southwest Symposium on Image Analysis and Interpretation*, pages 26–30, 2004.
- [21] M. Mikhalsky and J. Sitte. Active Motion Vision for Distance Estimation in Autonomous Mobile Robots. In *IEEE Conference on Robotics, Automation and Mechatronics*, volume 2, pages 1060–1065, 2004.
- [22] J. Miura, M. Itoh, and Y. Shirai. Toward Vision-Based Intelligent Navigator: Its Concept and Prototype. *IEEE Transactions on Intelligent Transportation Systems*, 3(2):136–146, June 2002.
- [23] 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.
- [24] 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.
- [25] W. A. Schafer, D. A. Bowman, and J. M. Carroll. Map - Based Navigation in a Graphical MOO. *Crossroads*, 9(1):8–15, September 2002.
- [26] M. Seul, L. O’Gorman, and M. J. Sammon. *Practical Algorithms for Image Analysis: Descriptions, Examples, and Code*. Cambridge University Press, 2000.
- [27] M. C. Shin, D. B. Goldgof, K. W. Bowyer, and S. Nikiforou. Comparison of Edge Detection Algorithms Using a Structure from Motion Task. *IEEE Transactions on Systems, Man, and Cybernetics- Part B: Cybernetics*, 31(4):589–601, August 2001.
- [28] K. Toyama, R. Logan, A. Roseway, and P. Anandan. Geographic Location Tags on Digital Images. In *ACM International Conference on Multimedia*, pages 156–166, 2003.
- [29] 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.
- [30] J. J. van Wijk and W. A.A.Nuij. A Model for Smooth Viewing and Navigation of Large 2D Information Spaces. *IEEE Transactions on Visualization and Computer Graphics*, 10(4):447–458, July 2004.
- [31] N. Warren, M. Jones, S. Jones, and D. Bainbridge. Navigation via Continuously Adapted Music. In *CHI Conference on Human Factors in Computing Systems*, pages 1849–1852, 2005.

- [32] M. Xie, C. M. Lee, Z. Q. Li, and S. D. Ma. Depth Assessment by Using Qualitative Stereo-Vision. In *IEEE International Conference on Intelligent Processing Systems*, volume 2, pages 1446–1449, 1997.
- [33] H. Zhao and R. Shibasaki. A Vehicle-Borne Urban 3-D Acquisition System Using Single-Row Laser Ranger Scanners. *IEEE Transactions on Systems, Man, and Cybernetics- Part B: Cybernetics*, 33(4):658–666, August 2003.