A General Framework for a Collaborative Mobile Indoor Navigation Assistance System

Huaming Rao[1,2] and Wai-Tat Fu[2]

[1]Department of Computer Sci & Tech, Nanjing University of Sci & Tech

[2]Department of Computer Science, University of Illinois at Urbana-Champaign

[1,2]{huamingr,wfu}@illinois.edu

ABSTRACT
Reliance on a fully autonomous system often requires well specified maps and sensors in the environment, making the system not easily extensible to new or changing environments. Given the increasingly ubiquitous network connections, one way to provide quick and robust spatial assistance is to connect local users to an expert user at a remote location, such that direct guidance can be provided. This workshop paper proposes a general framework for such a remote spatial assistance system. The goal of the system is to provide a cost-effective method to effectively transfer what the user is seeing to a remote expert who is familiar with the area (e.g., providing museum tours, guiding a lost pedestrian, providing guided emergency response to an area struck by hurricane), such that interactive assistance can be provided to the local user using augmented reality techniques. We propose that this framework contains three major components: 1) Spatial plans generation; 2) Context-aware positioning and error correction; 3) Collaborative guidance with augmented reality interfaces.

Author Keywords
Remote spatial assistance, augmented reality, mobile indoor navigation, collaborative system

ACM Classification Keywords
K.4.3 Organizational Impacts: Computer-supported collaborative work

INTRODUCTION
Given the relative lack of unique environmental cues, people often lose their sense of direction when navigating indoors. Indeed, despite the fact that Global Positioning Systems (GPS) has been widely used in outdoor navigation systems, they are not suitable for the indoor use because buildings may block satellite signals, making them unreliable, if not unusable. Rapid advances in mobile devices as well as ubiquitous wireless connections have, on the other hand, provided many other options that can be used for indoor navigation systems. For example, location information can be identified by sensor signals, such as those from the gyroscope, accelerometer, bluetooth, and wifi available in most mobile devices, while communication to other users can be augmented by augmented reality techniques. By integrating these technologies with methods from machine learning and computer vision research, mobile systems have great potential to augment users by providing useful and relevant information by recognizing the context under which users are located.

While research on developing autonomous agents has made significant progress, significant challenges remain. On the other hand, many tasks that are difficult to be handled by computing algorithms can be easily handled by humans. An optimal mix of computing and human agents can therefore provide a cost-effective approach for many practical problems. The main challenge for designing such a system is how one can seamlessly connect humans with computing agents, such that their tasks can be accomplished efficiently. This workshop paper proposes a general framework for a remote spatial assistance system which connects local users to an expert user at a remote location, such that direct guidance can be provided. The main components are shown in Figure 1.

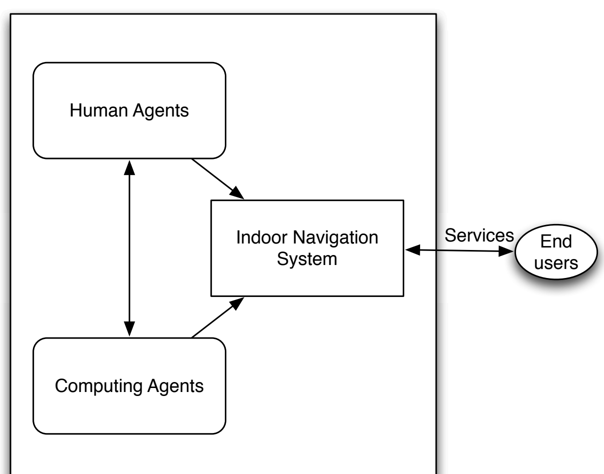

Figure 1. Main ideas of the system

LAMDa'13 in conjunction with *IUI'13*, March 19-22, 2013, Santa Monica, CA, USA.

EXAMPLE SCENARIO

We illustrate the framework with a scenario: An area was just hit heavily by a hurricane, and a local resident volunteers to go inside the building to search for persons who need help. The local rescuer has a ready-to-use smartphone or tablet computer that has connection to a remote expert. The remote expert has partial information about the whereabouts of a potential list of target persons, a floor map of the building (which might be outdated), and a live view of the environment through the local rescuers smartphone camera. The remote expert collaborates with the local rescuer to identify crucial anchor points in the building, work out an initial plan based on the relative locations of these anchor points, and incrementally modify and change the plan as they explore the environment. The remote expert leverages on the system that combines the live camera view with the floor map to generate additional spatial views that capture the relative locations of the anchor points, and create annotations on the camera views that facilitate communication with the local rescuer. The remote expert also has tools to quickly figure out new directions to verbally communicate to the rescuer, as well as to modify and generate new plans based on new information, queries, and inferences from the local rescuer. As a result, the system provides valuable spatial assistance to the local rescuer to search for the injured persons.

This scenario is just one example of the types of application that the proposed framework attempts to address. Further examples include remote collaboration between an interior designer and an architect or helping a lost pedestrian to navigate in an unfamiliar environment. The proposed framework does not aim at any of them specifically. Rather, the goal is to provide a proof-of-concept system to generate a set of principles for designing more effective remote spatial assistance systems in general.

RELATED WORKS

Huang et al.[8] review the recent mobile indoor navigation systems and provide an evaluation framework which combines the key aspects of building indoor navigation systems, among which indoor positioning and context-aware adaption are the ones that may mostly affect the performance of indoor navigation. Afyouni et al.[1] surveys indoor spatial models developed for research fields, also focusing on the perspective of context-aware while delivering services within indoor scenarios. So these two factors are the most considered ones when the system is designed.

Despite that, Gu et al.[6] access different indoor navigation systems using different criteria. Hightower et al.[7] describes the properties of location systems for ubiquitous computing and developed a taxonomy to help developers better evaluate their options when choosing a location-sensing system. Both of them look into indoor navigation in other ways, extending our field of view to design a better system by considering user preferences, complexity, commercial availability and limitations.

Many researchers[2][3][4] have studied remote collaboration and have identified unique challenges and solutions in different task scenarios. In particular, many have used augmented reality to allow remote users to communicate with the local users. Kim et al.[10] propose a vision-based indoor positioning system using augmented reality technique, which automatically recognizes a location within indoor environments from image sequences and overlays the user's view with the resulting location information.Huey et al.[9] also build a similar system, whose main difference from Kim's is that it handles the processing locally (on mobile phones or PDA) instead of transmit the images to the remote PCs. All of these works indicate the potential for incorporating AR interfaces into indoor positioning system to facilitate user's spatial cognition.

OVERVIEW OF THE PROPOSED SYSTEM

The goal of the system is to effectively transfer what the user is seeing (shared camera view, see Figure 2) to a remote expert who is familiar with the area (e.g., providing museum tours, guiding a lost pedestrian, providing guided emergency response to an area struck by hurricane), such that interactive assistance can be provided to the local user.

We adopt the general sensor-based indoor positioning techniques to approximately estimate a user's position. The sensors include wifi, gyroscope and accelerometer. Based on the local user's position, the remote experts work collaboratively with the local user to perform general spatial planning (e.g, sketch out a floor plan of the building), identify critical points in the indoor environment, and provide directional guidance with AR interfaces. The system is designed as a mobile-browser app with a server transferring information between local users and remote experts and restoring building floor plans. Figure 2 shows the general framework of the system, which is designed with three major components: 1) Spatial plans generation; 2) Context-aware positioning and error correction; 3) Collaborative guidance with augmented reality interfaces.

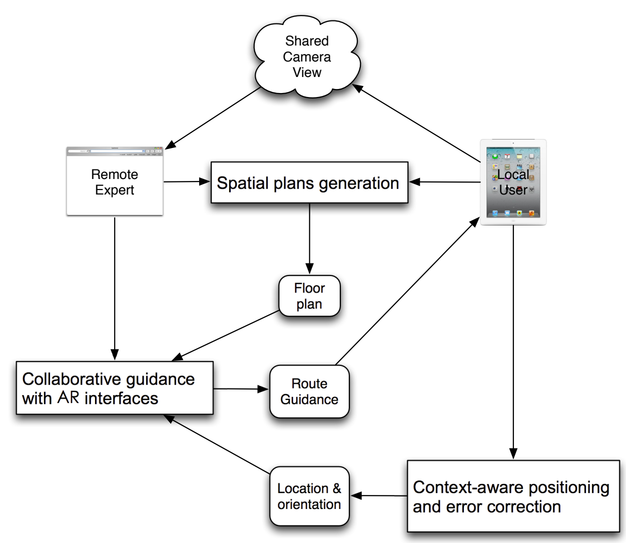

Figure 2. Framework of the proposed system

Spatial plans generation

One of the obstacles for implementing mobile indoor navigation systems is the difficulty to generate a detailed spatial plan for every building in advance. While large buildings usually provide floor plans at entrances or other salience spots, they may not be always available. So if an expert user can collaboratively work with the local user to sketch out a general spatial plan, it will help the user to have a good overview of the indoor environment. The workflow of this component is shown in Figure 3.

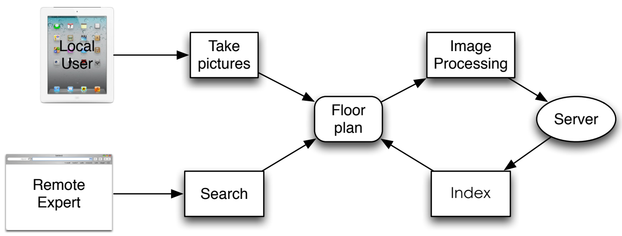

Figure 3. The workflow of spatial plans generation

The local user collaborates with the remote expert to generate a spatial plan (e.g., a floor plan). The local user can use the camera on the phone to take pictures of the environment in the building. The remote expert can interpret the image and, sometimes search the web for relevant information to help sketch out the spatial plan. The goal is to combine both snapshots of the environment with schematic representations of the indoor environment, such that a general spatial plan can be generated with critical landmarks that are meaningful to the local user for navigation and to gain a general sense of the spatial layout of the environment. The resulting spatial plans will be uploaded to the server and restored in database along with the buildings location as index keys after the image is adjusted and processed, such that other users can retrieve it. These indices will be useful for multi-user scenarios, as well as for future users.

Context-aware positioning and error correction

Sensor-based positioning

One common algorithm for sensor-based positioning is triangulation. It calculates the users locations by utilizing geometric properties. Depending on which property it measures, it has two variants: lateration and angulation. Lateration measures its distances from other reference points directly with received signal strengths or through the time of signal arrival or time difference of arrival. Angulation computes its angles relative to other reference points.

Besides radio signal sensors, the gyroscope and accelerometer are now available in most smart phones today can be also used to locate user's position. Pombinho et al.[11] propose a technique to record the foot path of the user's movement and calculate the location of the user inside a building. The algorithm detects when the user takes a footstep using the accelerometer and determine the direction of the footstep using the gyroscope, from which as well as the building floor plan the approximate result can be acquired.

Vision-based positioning

Nowadays almost every smart phone is equipped with a camera, which makes it possible to apply computer vision techniques to mobile indoor navigation systems. The object recognition community has developed some features that are invariant to changes in illumination, view point, scale, image translation and rotation. These features can be utilized to recognize the landmarks appearing in indoor environments. Sala et al.[12] propose a method to automatically select the minimum set of features by which the robot can navigate in its environments.

Context-aware positioning and error correction

Sensor-based positioning tends to have lower accuracy due to fluctuating signals, while vision-based positioning tends to have slow response times due to hardware limitation and time-consuming computations. Stand-along method may not work well to meet the high requirement of reliability. So the proposed system will use a context-aware way to select the best method to determine users location. Considering the hardware limitation of the device, vision-based may not be practical for consumer devices, but may become more applicable in the future.

Because WLAN is widely deployed in most large buildings. Some network infrastructure may even scan and tell users locations, which means instead of the client doing the work of scanning and calculating, the buildings network will do it. The app just need to ask the network where it is. But sometimes the infrastructure is not supported, then the app should do it itself. So the proposed system will prefer to use wifi signal and triangulation algorithm to estimate the approximate location of the user. When there is no wifi signal, the system will drop back to use gyroscope and accelerometer to record user's track to calculate the location [11].

But in practice, the environments are more complex indoor than outdoor due to the walls, the moving people and other electric equipments, which may interfere the signals of electromagnetic waves and the illumination conditions. To overcome these difficulties, the remote user need to be incorporated into the system to manually correct the inaccuracy when the signal is fluctuating. The processed signals are transferred to the remote side. While fluctuating signals are detected (changing sharply) the system will warn the remote expert to do the task of correcting location by seeing through the shared camera view. Note this task can be performed within agents by applying some scene analysis technologies, but it may benefit much from the expertise the remote side has especially when the agents fails to handle in some extremely severe situation. Figure 4 shows the workflow of this process.

Collaborative guidance with AR interfaces

The remote expert can see what the user is seeing through the shared camera view, such that he can know much about what's around the user. According to that, remote expert is able to give instructions as superimposed directional signs and other information into the shared camera view on the mobile device using AR techniques.To be more specified, the remote expert can sit in front of a computer and interact with the local user by drawing paths or click at some important points which then are transmitted to the local user in realtime

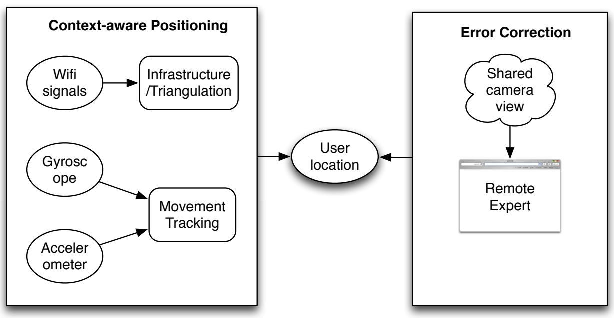

Figure 4. Workflow of context-aware positioning and error correction

and augmented on the camera view. The local user can see this information through his/her phone and get to know more about the environment that surrounds him/her.

Besides augmented visual information, verbal text and voice are also provided to assist the user in identifying his location easily particularly when the user is out of hands. The system is designed as a mobile-browser app, which means that remote side is rather flexible. This makes it much easier deployed to various consume applications.

CONCLUSION

Although we presented only our preliminary work on this system, our goal is to demonstrate the feasibility of the system. We hope that we can share this idea with the workshop participants, and receive valuable feedback to help us refine and extend the current system.

Future work will focus on:

- Perform systematic experiments to verify the degree to which it improves the local user's spatial cognition and helps him/her to navigate in indoor environments.

- Apply vision-based method to provide necessary information to the remote expert so as to assist him/her to provide better service to the end user. Given the challenges in identifying landmarks, identifying fiduciary markers are easier to achieve. By embedding information into 2d barcode, it makes the system more capable to know about its environments. However, these markers need to become more widely available to make them useful.

- Perform some empirical validations for the three proposed components: 1) to identify features that could most effectively describe the floor maps [12] within the communication between the local user and the remote expert, such that they can efficiently collaborate to generate the spatial plan; 2) to consider more context-aware factors that can meet user's requirements such as privacy, performance, cost, robustness, etc.[6], and identify cues in the shared camera view to highlight the conflicts between real locations and detected locations for more effective error correction; 3) to validate how augmented reality markers can provide better match guidance and instructions to the local users, and to test what types of markers can best help the local user know more about the surrounding environment and resolve ambiguities and confusion.

As Grudin[5] states, "Successful AI applications strengthen the tie to HCI by providing research foci as well as by creating a demand for new and better interfaces". Hopefully this work could be a good attempt towards this direction.

ACKNOWLEDGEMENT

The research is partially supported by a grant from the office of naval research N00014-12-1-0486 to the second author. And the first author gratefully acknowledges financial support from China Scholarship Council.

REFERENCES

1. Afyouni, I., Ray, C., and Claramunt, C. Spatial models for context-aware indoor navigation systems: A survey. *Journal of Spatial Information Science*, 4 (2012), 85–123.
2. Billinghurst, M., and Kato, H. Collaborative augmented reality. *Communications of the Acm 45*, 7 (July 2002), 64–70.
3. Dong, W., and Fu, W.-T. One piece at a time: why video-based communication is better for negotiation and conflict resolution. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 167–176.
4. Gelb, D., Subramanian, A., and Tan, K. H. Augmented reality for immersive remote collaboration. In *IEEE Workshop on Person-Oriented Vision (POV)* (2011).
5. Grudin, J. AI and HCI: Two fields divided by a common focus. *AI Magazine 30*, 4 (2009), 48.
6. Gu, Y., Lo, A., and Niemegeers, I. A survey of indoor positioning systems for wireless personal networks. *Communications Surveys & Tutorials, IEEE 11*, 1 (2009), 13–32.
7. Hightower, J., and Borriello, G. Location systems for ubiquitous computing. *Computer* (2001).
8. Huang, H., and Gartner, G. A survey of mobile indoor navigation systems. *Cartography in Central and Eastern Europe* (2010), 305–319.
9. Huey, L. C., Sebastian, P., and Drieberg, M. O. S. I. . I. C. o. Augmented reality based indoor positioning navigation tool. In *Open Systems (ICOS), 2011 IEEE Conference on* (2011), 256–260.
10. Kim, J., and Heesung Jun Consumer Electronics, I. T. o. Vision-based location positioning using augmented reality for indoor navigation. *Consumer Electronics, IEEE Transactions on 54*, 3 (Aug. 2008), 954–962.
11. Pombinho, P., Afonso, A. P., and Carmo, M. B. Indoor positioning using a mobile phone with an integrated accelerometer and digital compass. In *INForum* (2010).
12. Sala, P., Sim, R., Shokoufandeh, A., and Dickinson, S. Landmark Selection for Vision-Based Navigation. *IEEE Transactions on ROBOTICS 22*, 2 (Apr. 2006).