# *deSCribe*: A Personalized Tour Guide and Navigational Assistant*

Dheeraj Kota[2], Neha Laumas[1], Urmila Shinde[1], Saurabh Sonalkar[1], Karthik Dantu[1], Sameera Poduri[1], and Gaurav S. Sukhatme[1,2]

[1] Computer Science Department
[2] Ming Hsieh Department of Electrical Engineering
University of Southern California, Los Angeles, CA 90089, USA
{dkota,laumas,ushinde,sonalkar,dantu,sameera,gaurav}@usc.edu

**Abstract.** Mobile phones have become ubiquitous in daily life. They are also becoming more powerful with more computation and a variety of sensors embedded in them. We have built *deSCribe*, a context-aware phone-based navigational aid. The application provides turn-by-turn directions from the user's current location to a requested destination. It provides information about the surroundings by processing the images taken by the camera on the phone. It enables a novel user interface with the ability of using the phone as a remote to point at buildings to get further information about a particular floor within the building. Lastly, the application uses user feedback to control how much information is presented to the user.

## 1 Introduction

Mobile phones are ubiquitous. Commodity mobile phones with powerful microprocessors (order of 500 MHz), high resolution cameras (3.2 megapixel camera), and a variety of embedded sensors (accelerometers, compass, GPS) are becoming the norm. Coupled with good network connectivity, they can be used as a great source of information and user context.

One can envision using such mobile phones as mobile tour guides in trade shows, university campuses, downtowns, national parks etc where the user is assisted with information based on her location and situational context to better assist her in the task she is doing. We have implemented *deSCribe*, a mobile tour guide application that encapsulates several features required of such an application. Currently, the application has been tailored to the University of Southern California (USC) but given the data for other locations, the application can easily be tailored for them. We will now describe the design of our application and its various components.

This paper is organized as follows. Sec. 2 motivates the need for such an application. Sec. 3 describes some related work. Sec. 4 elaborates on the design of our application. Sec. 5 shows some results. Sec. 6 lists some of the limitations of our application due to

* This work was supported in part by NSF grant CCR-0120778 (CENS: Center for Embedded Networked Sensing), and by a gift from the Okawa Foundation. It was initiated as a project for the graduate course CS 546: Intelligent Embedded Systems taught at USC in Spring 2009.

S. Ystad et al. (Eds.): MobiCASE 2009, LNICST 35, pp. 393–403, 2010.

the tradeoffs we made and the contextual information at our disposal. Sec. 7 concludes our description followed by Sec. 8 which lists future directions.

## 2 Motivation

Every year University of Southern California gets hundreds of visitors who are not familiar with the campus and have trouble navigating and getting relevant information regarding the buildings on campus. The problem is further exacerbated by the fact that the campus spans a 226-acre area and has over a hundred buildings. This is not a problem unique to this or just other university campuses. We can imagine large downtowns, shopping complexes, fairs and exhibitions etc. posing a similar problem for a first time visitor. Such an application can be further extended using the user generated mapping in online sevices like wikimapia to further extract information about each building in any locality.

We hope to address this problem and have developed our application *deScribe* that would provide relevant information regarding the buildings in the vicinity of the users. The application also provides turn by turn directions to any destination on campus. Information of buildings is also provided in audio which makes it very user friendly, easy for hands free operation and provides better assistance for the visually impaired.


The input to our application was a detailed map of the USC campus at three resolutions and a database with building information and GPS coordinates. We use this to display the current location of the user on the USC campus and the point of the next hop in the route to the destination so that the user has an idea regarding her position on the map. Navigation is further aided by image processing algorithms to recognize side views of cars in the image which would help the visually handicapped to determine whether a car is going in front of them. This would aid in situations in case they need to decide whether they want to cross the road or not.

## 3 Related Work

There has been a lot of work in designing novel interfaces for mobile tourism and other interactive navigational services using a variety of embedded systems in the past few years. We cite a representative set here and draw from their experiences.

Schmidt-Belz et al. in [8] report their findings of a user validation study for the CRUMPET system which is a location aware mobile based tourism system focused on providing personalized tourism services. One of the findings of the study was that people need more textual/audio tour description as they find it difficult to interpret maps. Also, the study suggested that sights to visit are usually looked up both before traveling and also while on tour. The study also indicates that location based services would be very important for mobile based tourism applications.All these findings clearly indicate the need to provide users with relevant contextual information while on the go. Kray et al. in [5] review several navigation assisting services/devices available commercially or as research prototypes. The authors compare the various systems based on the services offered, positioning mechanism used, user interaction method, system architecture etc. Their observation is that more situational awareness is crucial for future mobile guides. The other key observation is that mobile guides need to adapt to real world situations like lack of network connectivity, visibility to GPS satellites etc. They also raise questions regarding the architecture to be used in the design of such guides and whether a client-server architecture is the most suitable. Gregory et al. in [4] talk about Cyberguide, a context aware tour guide. This is one of the first mobile applications to provide location aware services both indoor and outdoor. It also uses a past history of the user's locations to better assist the user.

[6] performed a multi-user mobile application game study. Their primary observation was that social interaction is a strong motive to participate in a multi-user service. They also concluded that it is equally influenced by three factors - context, communication and identification. The other novel outcome of their study was that user communication increased dramatically when the game was displayed on a large public display. This suggests that users are affected strongly by moving the viewership of their participation from their personal space to public space in some form.

[9] describes a location aware mobile tourist guide application where, it provides the users, in one of the mode they describe as a explorer mode, a way to provide details about buildings around the present location of the user. This data is provided on a map along with the current location of the user. We extend this to make the application more usable by utilizing the compass in the mobile phone to provide a mode intutive way of providing information to the user, thereby giving the flexibility of using the mobile phone as a virtual tourist guide who explains more about the building when one points to a building an asks more details about it.

There are numerous studies on using the phones in novel ways. [7] develops an application that allows the use of the phone as a pen to jot down small pieces of information quickly by writing in air. The Zagat NRU [1] and acrossair NY subway [2] are examples of virtual reality applications that use the GPS and the compass sensor to provide a direction for restaurants closeby and the nearest subway stations respectively. Wikitude [3] is the closest to our work. It is an augemented reality application which uses HTC G-1 camera view to display annotated landscape, mountain names, landmark descriptions, and interesting stories.

# 4 Design

The *deSCribe* application has three key functions. These are distinguished with three tabs on the application as shown in Figs. 1, 2.

- **Contextual Information:** Allows a user to point to any building on the USC campus and hear the name and other relevant details regarding the building being spoken aloud. The user has the option of setting the granularity of the details that they would like to hear. On changing the angle of elevation of the phone the user hears the floor number being pointed to. On pointing the phone vertically towards the sky they hear the current weather information.
- **Turn-by-Turn Directions:** Allows a user to find the shortest walkable route to any building on campus from her current location. The user is told the distance and direction she needs to walk to reach the desired location. It also displays a map of the USC campus and shows the current location of the user (a red pin is displayed) and the next hop in the path ( a blue pin is displayed) to the desired destination.
- **Phone as a remote:** The user can use the phone as a remote and point to buildings/surroundings of interest. She can then query for further information about the point of interest by the push of a button.

## 4.1 Platform

We used the Android Platform running on the HTC G1 phone for our development. The major factors that influenced the decision of choosing Android platform is the availabilty of various modes of localization available, with varying levels of power consumption and utilizing two different techniques which will come handy when the GPS direct signal is unavailable. Kray-C et al. in [5], provided a list of issues related to mobile application design after a survey of available tour guide applications. One of the criteria they mentioned is usage of application in a non networked mobile phone. We chose Android platform beacuse of the sqlite database they provide on the device itself. Considering the usability of the user, we provide the details of the building in text for which we required a text to speech library which is already available as an open source library on the android plaform. Since, the content regarding the buildings are saved on the local phone storage, it might become stale after some point, this problem can be solved once utilizing the already exisitng android market's infrastructures to provide applications and updates to the users. In addition to that, due to the open source licensing of android platform, more devices are likely to come, thereby increasing the target audience.

### 4.2 Turn by Turn Directions

As mentioned earlier, a database was provided with the GPS coordinates of all buildings on campus. The GPS sensor is switched on when the user requests route directions. This provides us updates regarding the user's current location after every two seconds. We also turn on the compass sensor on the phone to provide us with the user's orientation. We use the GPS to get her current location and then run the shortest path algorithm to find the appropriate route to the destination. On each location update from the GPS sensor, we determine the distance the user needs to walk by calculating her current distance to the GPS coordinate of the next hop in the route. A sample screenshot of the execution of the application is shown in Fig. 1.


**Fig. 1.** Turn-by-Turn Directions Initial Screen

We also determine the direction she should be walking in with the help of the onboard compass. We use this distance and direction to inform the user of where she should be heading in the form of a message displayed on the screen.

### 4.3 Contextual Information

We provide the user with contextual information by providing her the option to point to a particular building. The application determines the building and speaks aloud its name. The user can choose the granularity of details she hears about the building. The default option is *coarse* which only speaks aloud the name of the building. The user can switch to the *fine* option which would allow her to hear a detailed description of the purpose of the building along with any other interesting features such as names of cafes, restaurants etc., housed in the building. The text to speech library helps provide these details to the user in clear audio.


**Fig. 2.** Phone as a Remote Control Screen

As soon as the user clicks the option to determine the building ahead of her we query the GPS sensor on the phone to provide us with the user's current location and the compass to find his/her orientation. We have a database which has a GPS coordinate identifying each building on campus. It is stored in the android's sqllite database.

We use the database to filter out all the buildings within a specified range (currently the range consists of all buildings whose latitude and longitude is within 0.001 from our current latitude and longitude) from us. On obtaining this subset of buildings, we determine the distance from our current position to the GPS coordinates of each of these buildings. This distance is calculated in terms of Universal Transverse Mercator (UTM) coordinate system. The building description is obtained using the current location and the orientation of the compass. From the figure below consider the case where we are in between four buildings. The building to which the mobile device is currently pointing to is obtained by getting the angle the ends points of building makes to the current location of the device. We can conclude that we are pointing to that particular building if the compass is oriented within the range of this particular angle. For this purpose we need to assign angles dynamically from the current location to the buildings around. This would have been trivial in case we had the GPS co-ordinates of the four corners of the building. However, currently we have been able to obtain only one GPS co-ordinate of each building. Further, we think it would be the same case in many other scenarios, where we do not have complete information about the environment. We assume that the single GPS co-ordinate we have for each building, is at the center of the building. We now calculate the slope of the line joining the current location of the phone obtained via GPS to the building's latitude and longitude information available to us. Consider the building b1 in the figure which at a distance d1 which makes an angle of 135 degrees with the +x axis from the current view. Now we can tell for sure if the compass if oriented towards 135 degrees that we are pointing to b1. The building can be huge and we need identify the building even when the compass is pointing to any part of the building. In order to achieve this, we assign some range $(135 - dx, 135 + dx)$ to the building, which can be used as a measure to decide if we are pointing to that particular building. We obtain the value of $dx$ based on the current distance $d1$ which is the distance between the current location of the phone to the GPS co-ordinate of the building. The value of $dx$ varies with distance and the value with which it varies was obtained empirically by getting an estimate of about ten buildings at USC by calculating the angle made by each building from various distances. A larger value of $dx$ is assigned to a smaller distance and vice a versa and we made a lookup table assigning the value of $dx$ based on distances.

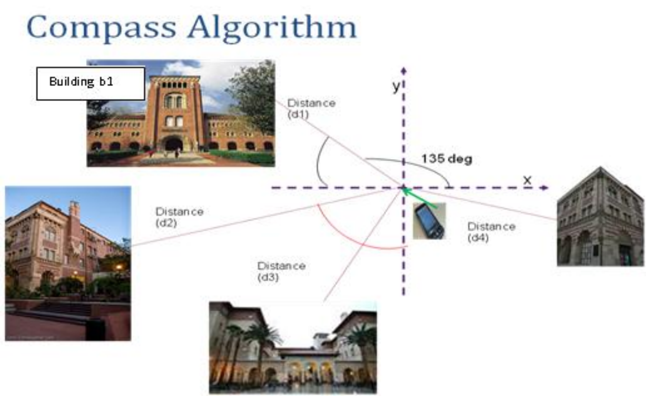


**Fig. 3.** Compass Algorithm

### 4.4 Phone as a Remote Control

Floor determination is done using the Y axis of the compass which indicates the pitch. When the phone is exactly horizontal, the Y axis indicates 0 degrees. When you start turning the phone upwards then angle becomes negative and is $-90$ degrees when completely vertical. When you take the phone downwards the angle increases from 0 to $+90$

degrees when the phone is completely pointing to the ground. To calculate which floor you are pointing to, one has to get the distance between the current location and the building. Then from the y axis angle and the distance you can get the height the phone is pointing to by the formula.

$$Height = \tan(angle) * Distance$$

The height of the floor is assumed constant at 3.7 meters. Based on this data we can calculate the floor number the phone is pointing to. One of the problems that we faced when calculating the height is that the GPS coordinates of the buildings we have is of the center of the building. Fig. 4 below illustrates the error introduced.

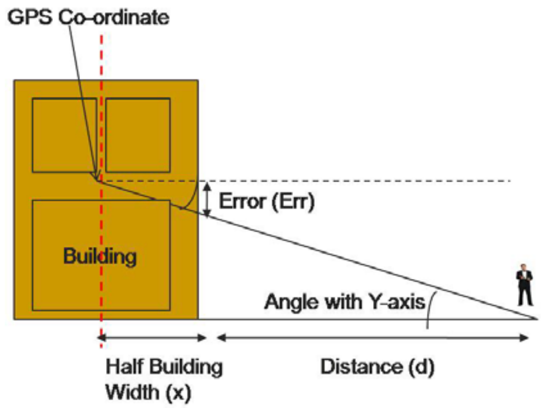


**Fig. 4.** Computing the Floor Number

As shown in the Fig. 4, the measured height is

$$H2 = \tan(angle) * (d + x)$$

when it should be $\tan(angle) * d$. So the error introduced is: $err = tan(angle) * x$. Fig. 6 shows the plot of error versus angle in which the building width is assumed constant at 30 meters. Graph shows the increase in error as the angle increases. So the floor information will work best only for smaller angles which means buildings having less floors or lower floors of tall buildings. Also if we keep the angle constant, the error becomes proportional to the width of the building. So the floor works best only for the small buildings or we need to have the GPS coordinates of all the four corners of the building. Another factor that the floor determination depends on is the users view and the phone view. Sometimes the user thinks she is pointing to a particular floor based on her line-of-sight angle but the phone angle is slightly different. This also introduces significant error depending on the distance from the building. If the angle of pointing increases beyond 80° or beyond the number of floors in a building, then we get the weather information. We are currently using the Google weather API. The API returns an XML file which is parsed and the information such as current weather condition, the temperature, humidity and wind conditions are spoken out by the TTS engine.

### 4.5 Current Location Display

We determine our current location with the help of the GPS as shown in Fig. 5. As mentioned earlier, we have been provided with the USC map at three resolutions. We align the tiles such that it displays the user's current location. We also show the point on the map which in the next hop to the destination. This can provide the user an understanding of her relative position and a sense of orientation on the USC map. Determining the presence of a car We give the users an option to determine the presence of cars ahead of her. This work has been motivated to aid the visually handicapped. The user is required to take an image of the scene ahead of him. We then upload this image to the aludra server running in USC. We use the GentleBoost image recognition algorithm on the server to identify the number of cars detected in the image. We then send this information back to the phone and provide the user the number of cars that were detected in the image. This information is helpful since the user is aware that a car is going side ways in front of him, indicating that it is probably dangerous for him/her to cross the road.


**Fig. 5.** Map Display with Current Location (Red) and Next Hop (Blue)

## 5 Results

We collected data from about 30 buildings on campus and our accuracy rate was 83.2%. We found that the buildings identified incorrectly were very big in size. We attribute our

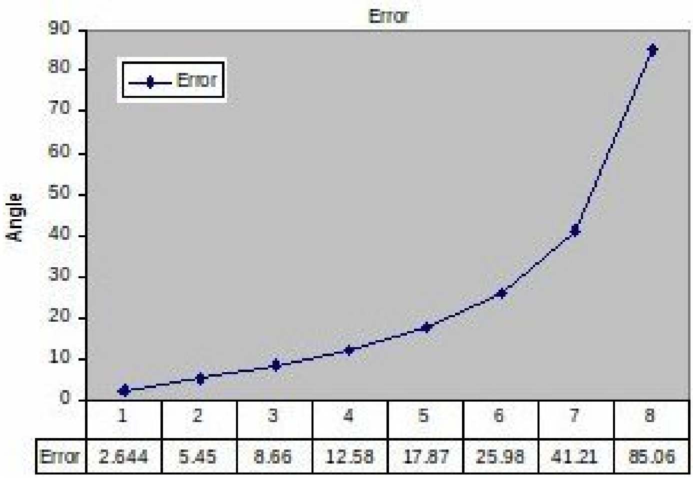


**Fig. 6.** Angular Error in Floor Computation

inaccuracy to the fact that our algorithm assigns angles based on the distance and for large buildings we would need to move very close to the building in order to get correct identification. In case we had been provided with GPS coordinates for all four corners of buildings we would not have to use an empirical angle based on distance and would have got more accurate readings.

## 6 Limitations

Our database currently identifies each building on campus by a single GPS coordinate. Also, these GPS coordinates are often not located in the center of buildings and building sizes are variable on campus. This sometimes leads to some inaccuracy in the estimation of the angle the phone is making with the buildings in the vicinity and leads to incorrect answers by the phone. The accuracy of our estimation of the building that is being pointed to by the phone would increase considerably in case we had GPS coordinates of all the four corners of each building since they would help us calculate the correct angle the phone is making with each building. There is an error introduced when estimating the height of a building since we do not have the GPS coordinate of corners of buildings. The GPS coordinate available being at the center of building introduces an error in the estimation of our distance from the building for the purpose of calculating the height of the building. We have assumed that the routing is point-to-point and that the roads are straight lines. This sometimes leads to somewhat ambiguous directions being given to the user. Both these limitations could be overcome by obtaining finer details in the map like the GPS coordinates of building corners and the size and coordinates of the roads.

Also, in our work we use GPS measurements both for finding turn by turn directions and finding the building that is being pointed to by the phone. GPS uses a lot of power and we find a significant reduction in the battery lifetime on using our application continuously. On an average the talk time on the phone is 5 hours and it reduces to nearly 3 hours.

## 7 Conclusion

We believe that this work is a pre cursor to develop such applications for schools, recreational parks etc. within which buildings and land marks are often not reachable by commercially available GPS devices like Garmin and Tom Tom. We believe that such an application is highly beneficial and in the long run will replace paper maps which can help find a route to a particular destination but do not provide any contextual information. Further, using the phone as a remote control is a novel user interface for such mobile devices and can see many other uses beyong mobile tourism.

## 8 Future Work

Currently the information about each and every building is saved on the mobile phones local storage. The information about a building can change. This new information has to be updated in the phones database. This has to be done with minimum amount of delay so that the users donot get incorrect information. The application can be extended to recognize more objects such as bicycles; pedestrians etc. that are present ahead of the user and integrate them to our application. This would make the application appealing for use for the visually impaired it would provide them information regarding what lies ahead of them.

## References

1. http://www.zagat.com/Blog/Detail.aspx?SCID=42&BLGID=20939
2. http://www.acrossair.com/apps_newyorknearestsubway.htm
3. http://www.wikitude.org
4. Abowd, G.D., Atkeson, C.G., Hong, J., Long, S., Kooper, R., Pinkerton, M.: Cyberguide: A mobile context-aware tour guide. Baltzer/ACM Wireless Networks 3 (1997)
5. Baus, J., Kray, C., Cheverst, K.: A Survey of Map-based Mobile Guides. Springer, Heidelberg (2005)
6. Leikas, J., Stromberg, H., Ikonen, V., Suomela, R., Heinila, J.: Multi-user mobile applications and a public display: novel ways for social interaction. In: Fourth Annual IEEE International Conference on Pervasive Computing and Communications, PerCom 2006, pp. 5–70 (March 2006)
7. Sandip Agrawal, I.C., Gaonkar, S., Choudhury, R.R.: Phonepoint pen: Using mobile phones to write in air. In: ACM Workshop on Networking, Systems, Applications on Mobile Handhelds, Mobiheld 2009 (2009)
8. Schmidt-Belz, S.B., Schmidt-belz, B., Nick, A., Poslad, S., Zipf, A.: Personalized and location-based mobile tourism. In: Proceedings of Mobile-HCI (2002)
9. ten Hagen, K., Modsching, M., Kramer, R.: A location aware mobile tourist guide selecting and interpreting sights and services by context matching. In: 2nd Annual International Conference on Mobile and Ubiquitous Systems (MobiQuitous 2005), pp. 293–304. IEEE Computer Society, Los Alamitos (2005)