



# Android Based Electronic Travel Aid System for Blind People

Paraskevas Diamantatos, Ergina Kavallieratou

► **To cite this version:**

Paraskevas Diamantatos, Ergina Kavallieratou. Android Based Electronic Travel Aid System for Blind People. Lazaros Iliadis; Ilias Maglogiannis; Harris Papadopoulos. 10th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2014, Rhodes, Greece. Springer, IFIP Advances in Information and Communication Technology, AICT-436, pp.585-592, 2014, Artificial Intelligence Applications and Innovations. <10.1007/978-3-662-44654-6\_58>. <hal-01391364>

**HAL Id: hal-01391364**

**<https://hal.inria.fr/hal-01391364>**

Submitted on 3 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Android Based Electronic Travel Aid System for Blind People

Paraskevas Diamantatos Ergina Kavallieratou  
Dept. Information and Communication Systems Engineering  
University of the Aegean  
Samos, Greece

Diamantatos@aegean.gr, kavallieratou@aegean.gr

**Abstract.** Blindness is the condition of lacking visual perception due to physiological or neurological factors. Blind people do not have the full perception of the surrounding environment, though navigating, in an unknown environment or/and with obstacles on route, can be a very difficult task. In this paper, an information mobile system is presented, that acts as an electronic travel aid, and can guide a blind person through a route, inform him about imminent obstacles in his path and help him orientate himself. The current prototype consists of a mobile phone, and the developed application.

**Keywords:** Blind, navigation, mobile phone

## 1 INTRODUCTION

It is absolutely crucial in our time to be able to help blind people move in urban environments effectively, with a simple, easy accessible, low cost system. According to existed research there are almost 40 million blind people [1] that naturally use a simple cane, or a guide dog to help them move around. Both ways have disadvantages: a cane cannot detect obstacles further away than its height, and the guide dogs while more expensive, are not enough for all the blind people.

The majority of the suggested systems [2-9] are often too expensive and unpractical in use, while most of the available navigation systems, cannot detect obstacles in a specific route, they simply give directions to the user for a point A to a point B [2]. The proposed system helps the user to move in the environment, detecting obstacles and finding the best path, in order to avoid them. It can give vocal instructions to the user, or use the mobiles vibration in order to give navigation instructions. It is also able to orientate the user, using military vocal orientations e.g. north 12 o'clock.

In the following section 1.1, a short description of previous work is given, while in section 2, the proposed system is described in detail. Finally, some comparison results and conclusion are mentioned in sections 3 and 4, respectively.

## 1.1 LITERATURE REVIEW

There are two major elements of the mobility problem that a Blind person has to deal with. First, the blind must sense obstacles in his environment and avoid them. The second problem, equally serious as the first, is the person navigation [3].

In the recent years a number of navigation system products became available to the public. Systems, like Google maps [2], that offer navigation from a point A to a point B, and can also support voice interaction with the user, braille interface, and screen reader interface especially for blind people [2]. While the navigation problem, seems to be a solved problem for outdoor navigation, mostly, due to the quality of the offered services and products to the public, the obstacle detection problem remains.

ETAs, Electronic Travel Aids, try to deal with this problem. ETAs development started at the mid 60's. Some of the most known applications include the LaserCane [4], which is a regular long cane with a built-in laser ranging system. The Mowat Sensor is an example of a pocket-sized device containing an ultrasonic air sonar system [5]. When it detects an obstacle, the device vibrates, thereby signaling to the user. Similar technology is used on NavBelt [6], a belt that utilizes an ultrasonic sonar system, and a guide cane, a cane that has a robot attached that also utilizes ultrasonic sonar system in order to guide the user.

While most of the ETAs try to address the obstacle detection problem, some ETAs try to provide navigation for indoor uncharted buildings. More specifically, Navatar [7] utilizes several sensors that can be found on modern smartphones like accelerometer and compass to determine the movement of the user in the space. It also takes a user input into account, e.g. if a user finds a door he can save it on the system. By this information, Navatar can determine the user's location inside a building [7]. Another system for indoor navigation is a case study from the school of engineering of University of California where a system was developed that utilizes mobile phone cameras and searches for specific markers on the walls [8].

A different approach for developing an ETA was used on tyflos system [9]. This system uses stereoscopic cameras in order to detect the depth of obstacles in the user's route and notify the user through a vibration belt. The research problem of designing a better ETA is a tough one. Despite 50 years of effort, no one has been able to design an electronic device that can replace the long cane.

## 2 THE PROPOSED SYSTEM

An application was developed that makes use of an android mobile phone with cameras along with sensors that can be found on every android smartphone, like accelerometer and compass. By analyzing the image taken by the camera and creating a depth map it is possible to detect obstacles in the user's path and redirect him by a different route.

In order to achieve that, two consecutively photos are considered and their common points are located. Then the optical flow between them is established, allowing extracting the homography and the fundamental matrices. This step is important in order to rectify the images and transform them to stereo images. This step allows the

creation of a depth map that contains any objects found, along with their relative position in space. Next the routing subsystems computes a route with no obstacles and notifies the user. In this section, all the mentioned modules of the proposed system (fig.1) will be further analyzed.

The proposed system is able to help blind people to be orientated in their environment by informing them orally on the location they are staring: north, south, east, west and their combinations. It consists of a total of 12 subsystems, as it is presented in figure 1. The system was developed in the Android mobile platform, and it also requires the Opencv for android library.

The Camera subsystem is responsible to capture two sequential frames using the user's mobile phone camera. The output is two bmp images with resolution 352\*288. Low resolution images are used in order to achieve lower computational time. Thus, the system can analyze the image in real time. In the prototype version, the two frames are shown side by side on the user's mobile phone screen, just for testing purposes.

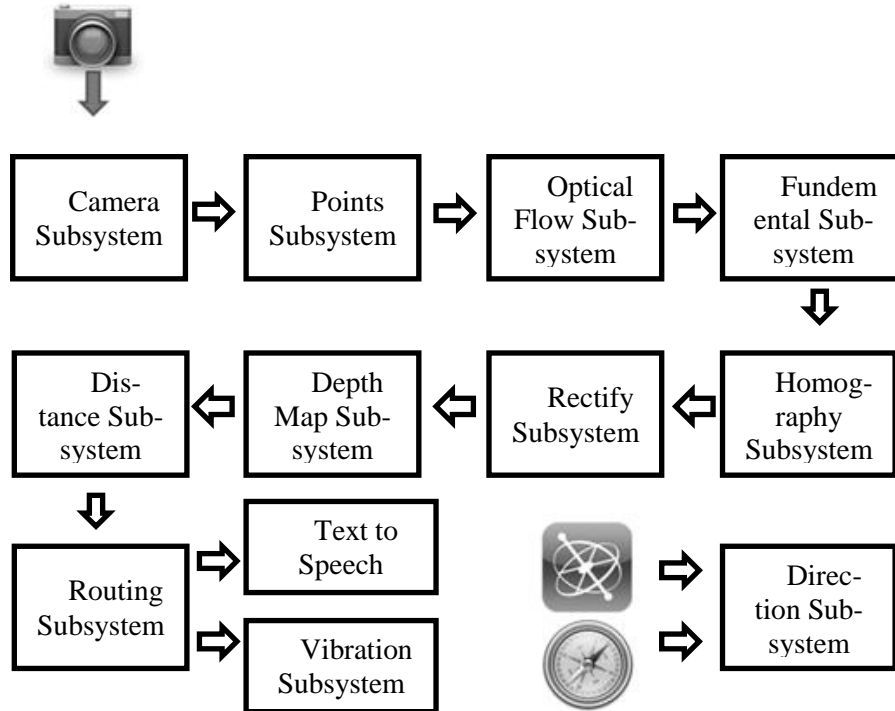
The Points subsystem is responsible to identify common points between the two images acquired by the Camera subsystem, using a detector. Various detectors were tried, offered by opencv library for this task, like SURF [10]: a robust local feature detector that is based on sums of 2D Haar wavelet responses, ORB [11]: a very fast binary descriptor based on BRIEF, and GFTT [12]: a detector that finds the most prominent corners in the image. The results proved that GFTT (good features to track) detector was ideal for the task, because of its very low computational time.

The Optical Flow subsystem is responsible to calculate the optical flow, i.e. the distance of the corresponding points (shift) between the first and the second frame, detected in the previous subsystem. It is absolutely crucial for the entire system that the points and optical flow subsystems return exact results, since all the next steps depend on those points and their optical flow.

The Fundamental subsystem, along with the Homography subsystem are responsible to calculate the fundamental [13] and the homography [14] matrices, respectively. The fundamental matrix is a  $3 \times 3$  matrix which relates the corresponding points in stereo images. While homography relates the transformation from a projective space to itself that maps straight lines to straight lines. Homography has many practical applications, such as image rectification, which our system performs on the Rectify subsystem. It is also used for image registration, or computation of camera motion (rotation and translation) between two images. Once camera rotation and translation have been extracted from an estimated homography matrix, this information may be used for navigation, or to insert models of 3D objects into an image or video, so that they are rendered with the correct perspective and appear to be part of the original scene.

The rectify subsystem is responsible to transform the two bmp images in stereoscopic images [15], images taken by a stereoscopic camera. Thus, their common points are located on the same axis. Opencv provides the tools for the image rectification process [16]. It is suggested that the rectification is made by using the camera characteristics along with the homography matrix. The problem here is that the camera characteristics are not available for every camera that someone can use in a mo-

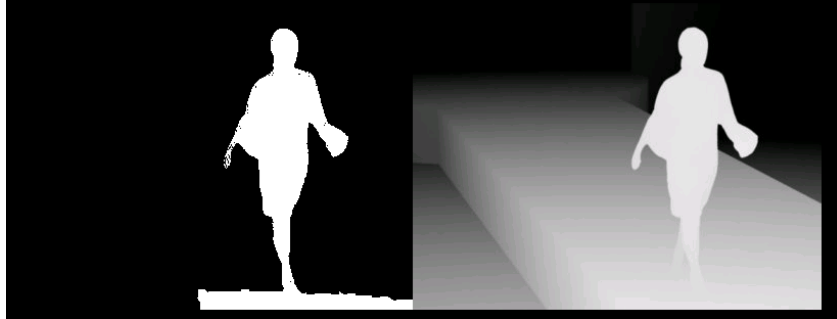
bile phone. However, the rectification is also possible by using only the homography matrix.



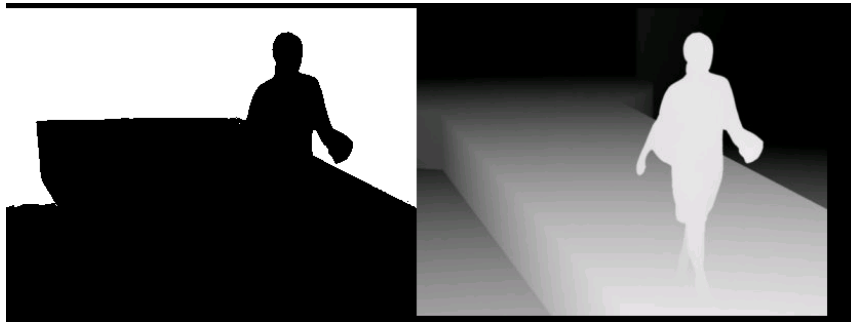
**Fig.1.** Overview of the proposed system.

The Depth Map subsystem is responsible to calculate the disparity depth map of the two stereo images of the rectify subsystem. A disparity depth map [16] is a 2D image where the color of each pixel represents the distance of that point from the camera. In more detail, light pixels are near and dark pixels are far. The disparity depth map will be used in order to identify obstacles near the user, for the calculation of the path that the user should follow. Block matching algorithms [16] offered by openCV were used, in order to produce the depth map.

The Distance subsystem is responsible to roughly calculate the distances of all the objects that can be found on the depth map. This was implemented by dividing the densities of the pixels from the depth map in 5 sub images. For objects very near to the camera (figure 2), are considered those with densities in the range of 220 – 255. Respectively a rough estimation (figure 3) is done for all the other objects (near, far, very far, and extremely far).



**Fig.2.** Objects located very near. Depth Map on right, Left only pixel densities between 220-255 are on



**Fig.3.** Objects located extremely far. Depth Map on right, Left only pixel densities between 0-80 are on

The Routing subsystem is responsible to calculate the area in the image (left, right or center) that includes the freest space. This is done by calculating the horizontal histograms on each of the images extracted by the Distance subsystem. After an iterative process, the area in the image with the freest space is localized. For example on figures 2 and 3, on the right side, the depth map of a scene is presented, while on the left only specific pixel densities are selected to be on. By considering the objects located very near (fig. 2, left image), and calculating the horizontal histogram of the left image, it is computed that on the left side of the scene, the user has free space to move, while on the right side of the scene, the user has no free space.

The Text-to-speech subsystem is responsible to notify the user with voice commands, while the Vibration subsystem, if selected, is responsible to notify the user with vibration codes. One vibration stands for left, two vibrations for right.

A Direction subsystem was also implemented that is responsible to notify the user of the direction he is headed, orally, using military codes. For example, if the user is headed north, the system will notify the user that he is headed at 12 O' clock.

### 3 COMPARISON RESULTS

It is hard to actually measure the results, mostly because there is no standard evaluation technique. However, it is possible to check, if we get the expected results from the subsystems.

For the Optical Flow subsystem, that is based on the Points subsystem, it was possible to measure the detected distances of the points, manually. The followed procedure should save the 2 sequential frames, along with the points detected, and their optical flow. The extracted information for 20 different sets of frames was reviewed. The success rate of this subsystem was 85%, due to some misplaced points, that is a normal behavior for this kind of subsystem.

The Rectify subsystem that is based on the Fundamental and Homography subsystems is the one responsible for the bad output of our algorithm. The accuracy of this system was also measured manually. 50 rectified images were kept, and later reviewed. Only the rectified images onto a common image plane in such a way that the corresponding points have the same row coordinates, where considered to be the ones expected. Specifically only the 4 % of the frames work as expected, making things worse.

The Depth Map subsystem works as expected. Known datasets were used to test the performance of this subsystem, along with their ground truth data. By tuning the characteristics of this subsystem an identical image to the ground truth was achieved. A total of 20 sets of stereo images where used, to output their depth maps. Overall the success rate of this subsystem found to be 95%.

The results revealed that the depth map creation is exceptional. On figure 2 we can see a depth map created from our system on a moncamera device, while on figure 3 a depth map created from images taken from a stereoscopic camera device is presented.



**Fig.4.** Depth map from moncamera device





**Fig.5.** Depth map from stereoscopic camera device

For the Routing subsystem that is based on the Distance subsystem several depth maps were used, and on each one of the depth maps, a decision was made over which direction the user should go. The success rate of our algorithm was 70%.

## **4 CONCLUSION**

Although our system works in some frames, it is not the case for all of them. Thus, it cannot be used in a real time application. Responsible for that is the camera characteristics that are not available on most of the mobile phones. Usually mobile phone manufacturers do not care to have available any of the camera characteristics per device. They just care for the camera to work.

This is making the development of such application, which utilizes only one camera, difficult, because it also requires a subsystem for camera calibration. Implementing one in our tests, different camera characteristics were extracted every time. Without camera characteristics, another function for camera calibration (stereoBM) was used that did not give so good results.

Although the final results are not perfectly good, and this system cannot be used by blind people yet, for safety reasons, we strongly believe that this kind of application is possible to be implemented in 3D mobile phones with 3D stereo cameras.

## **References**

1. World Health Organization 2012 ,Global Data on Visual Impairment 2010.
2. "Official Blog: Speech-friendly textual directions from Google Maps." 2006. 4 Mar. 2013 <<http://googleblog.blogspot.com/2006/12/speech-friendly-textual-directions.html>>.
3. National Research Council (U.S.). Committee on Vision. Working Group on Mobility Aids for the Visually Impaired and Blind, 1986 "Electronic travelaids: new directions for research".
4. J. M. Benjamine, N. A Ali and A. F. Schepis, a laser cane for the blind, proceeding of the San Diego Biomedical Symposium, 1973, Vol. 12, pp 53-57.

5. R. Nagarajan, S. Yaacob, and G. Sainarayanan, Role of Object Identification in Sonification System for Visually Impaired, In IEEE Tencon (IEEE Region 10 Conference on Convergent Technologies for the AsiaPacific), Bangalore, India, 2003, October 15-17.
6. Shoval, Shraga, Iwan Ulrich, and Johann Borenstein. "NavBelt and the Guide-Cane [obstacle-avoidance systems for the blind and visually impaired]." *Robotics & Automation Magazine*, IEEE 10.1 (2003): 9-20.
7. Apostolopoulos, Ilias. "Integrating Minimalistic Localization and Navigation for People with Visual Impairments." May. 2011.3. Warren, D.H. and Strelow, Edward R. *Electronic spatial sensing for the blind*. 1985.
8. Manduchi, Roberto, Sri Kurniawan, and Homayoun Bagherinia. "Blind guidance using mobile computer vision: a usability study." *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility* 25 Oct. 2010: 241-242.
9. Bourbakis, Nikolaos. "Sensing surrounding 3-D space for navigation of the blind." *Engineering in Medicine and Biology Magazine*, IEEE 27.1 (2008): 49-55.
10. Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU)*, Vol. 110, No. 3, pp. 346--359, 2008.
11. Rublee, Ethan et al. "ORB: an efficient alternative to SIFT or SURF." *Computer Vision (ICCV)*, 2011 IEEE International Conference on 6 Nov. 2011: 2564-2571.
12. J. Shi and C. Tomasi. Good Features to Track. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 593-600, June 1994.
13. Q. T. Luong and Olivier D. Faugeras (1996). "The Fundamental Matrix: Theory, Algorithms, and Stability Analysis". *International Journal of Computer Vision*.
14. Hartley, Richard, and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge, 2000.
15. N. Holliman, "3D Display Systems." to appear; *Handbook of Opto-electronics*, IOP Press, Spring 2004, ISBN 0-7503-0646-7.
16. Dröppelmann, Sebastian et al. "Stereo Vision using the OpenCV library." (2010).