ISSN 1975-8359(Print) / ISSN 2287-4364(Online)
The Transactions of the Korean Institute of Electrical Engineers Vol. 65, No. 12, pp. 2138~2144, 2016
http://dx.doi.org/10.5370/KIEE.2016.65.12.2138

컴퓨터 비전 기반 시각 장애 지원 모바일 응용

A Computer Vision-based Assistive Mobile Application for the Visually Impaired

Arnel A. Secondes* · Nikki Anne Dominique D. Otero* · Frank I. Elijorde* · 변 영 철†
(Arnel A. Secondes · Nikki Anne Dominique D. Otero · Frank I. Elijorde · Yung-Cheol Byun)

Abstract - People with visual disabilities suffer environmentally, socially, and technologically. Navigating through places and recognizing objects are already a big challenge for them who require assistance. This study aimed to develop an android-based assistive application for the visually impaired. Specifically, the study aimed to create a system that could aid visually impaired individuals performs significant tasks through object recognition and identifying locations through GPS and Google Maps. In this study, the researchers used an android phone allowing a visually impaired individual to go from one place to another with the aid of the application. Google Maps is integrated to utilize GPS in identifying locations and giving distance directions and the system has a cloud server used for storing pinned locations. Furthermore, Haar-like features were used in object recognition.

Key Words : Visually impaired, Smartphones, Assistive systems, Navigation systems, Object recognition

1. Introduction

Vision is an essential ability of a human being. The absence of such can negatively affect a person. It limits his capacity to be a functional member of society. Just recently, an estimate by the World Health Organization (WHO) states that 285 million people worldwide are visually impaired[1]; while 39 million of these are blind, an estimated 246 million have low vision. Without additional interventions, these numbers are predicted to significantly increase by the year 2020[2]. A huge percentage of visual impairments can be avoided and even cured; unfortunately, 90% of the visually impaired live in developing countries which do not have sufficient treatment and support options available to help them deal with such disability. In addition, 65% of the visually impaired are aged 50 years or older and this number is projected to increase[3].

Several accessibility and mobility problems are encountered by blind and visually impaired people in their everyday life. These tasks or situations usually involve physical constraints, which are sometimes impossible for them to overcome[4]. Such difficulties can be attributed to the lack of autonomy or information that could assist them in avoiding obstacles. For instance, moving through an unknown environment becomes a real challenge when we can't rely on our own eyes[5]. Since dynamic obstacles usually produce noise while moving, blind people develop their sense of hearing to localize them[6]. However they are reduced to their sense of touch when the matter is to determine where or what an inanimate object exactly is. Then the issues dealing with communication and access to information are pointed out. Here a significant help is offered by software applications for computers and touch-screen devices equipped with speech synthesizers that enable browsing the internet and access to text documents.

Many of the commercially-available assistive products are beyond the financial reach of the common population while most state-of-the-art technologies are hard, especially for the elderly, to grasp. Such a situation raises the fact that there is an urgent need to develop solutions which are cost-effective, intuitive and make use of commonly available technologies such as smartphones or tablets. Mobile phones are nowadays ubiquitous, even among blind and visually impaired users. As trend has it, it is observed that phones with physical keypads are becoming less popular. This introduced a new usability barrier for the visually-impaired who had to go through difficulties in handling small touch-screen devices. Fortunately, new user interfaces that utilize touch gestures have significantly improved user-friendliness of the touch-screen devices for the blind. If deployed in a capable device, computer vision technology has the potential to assist blind individuals to independently access, understand, and explore such environments.

† Corresponding Author : Dept. of Computer Engineering, Jeju National University, Korea.
E-mail: ycb@jejunu.ac.kr
* Institute of Information and Communications Technology, West Visyas State University, La Paz, Iloilo City, Philippines
Received : October 17, 2016; Accepted : November 21, 2016

Copyright ⓒ The Korean Institute of Electrical Engineers
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/)which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

The general objective of this study is to develop an android-based assistive mobile application that aids visually impaired people through object recognition and navigation. Specifically, the study aims to develop a mobile application that could recognize objects with the use of Haar-like features, identify user's current location and give distance information using Google Maps, respond to human speech and produce voice feedbacks.

2. Related Works

Robust and efficient object recognition systems can be of great help to people with severe vision impairment to allow them to access unfamiliar environments without frequent assistance. However, recognizing generic objects in the surrounding environment is a major challenge for the visually impaired for which few assistive technological solutions have been devised to date. Previous solutions make use of alternate sensing technologies, such as RFID[7], sonar[8] and infrared[9] for this task, thus, avoiding some of the inherent drawbacks associated with strategies based on computer vision. However, the aforementioned technologies also suffer from limitations of their own- they all require special sensing equipment. For instance, infrared and RFID require specific tags; on the other hand, sonar and infrared may not be very effective in indoor environments because the obstacles present in such surroundings may cause the reflected echoes to become distorted resulting in unreliable information being conveyed to the user.

Another application intended for the blind users is the EyeRing project[10]. It is a finger-worn device that communicates with an Android mobile phone. The EyeRing comprises a VGA mini-camera, a microcontroller, a Bluetooth connection module and control buttons. The task of the mobile device is running speech processing algorithms and all computer vision algorithms. The currently implemented functionality of the device is: detection of banknotes, recognition of colors and distance calculation which is supposed to work as a “virtual walking cane”. This solution, however, is costly and requires an additional device to be worn by a blind user.

DORA[11], an object recognition system, using artificial neural networks and capable of recognizing a set of pre-trained object that aims to provide visually impaired people a camera-based detection system that would help them recognize surrounding objects and their descriptive attributes (brightness, color, etc.), and announce them on demand to the patient via computer-based voice synthesizer. The similarity of the study with the proposed system is to assist visually impaired people recognized the surrounding objects and their descriptive attributes and announce them verbally via voice synthesizer.

In [12], location-based services provide the mobile clients personalized services according to their current location. The authors propose the implementation of location-based services through Google Web Services and Walk Score Transit APIs on Android Phones to give multiple services to the user based on their Location. Some feature is similar with the proposed system like using positioning methods to discover the position of the mobile to identify the location of the user. The proposed application supports Google API and generates data from the Google Maps. What is unique with our system is that the application processes the coordinates more frequently for added accuracy and it can tell the user's current location and the travel time for his desired destination.

The work in [13] utilized location based services are used for obtaining the location of the user and utilizing this to provide a service. LBS services can be categorized as triggered LBS services (push services) and user-requested LBS services (pull services) In a triggered (push) LBS service, the location of user's mobile device is retrieved when a condition set in advance is fulfilled. In a user-requested (pull) LBS service, the user decides whether and when to retrieve the location of his/her mobile device and use it in the service. The proposed system is akin with this study in a way that they both aim to obtain the location of the user and utilize the information to provide service. What's unique with our system is that the user can pin a location and have a codename for it, its data is stored in the system's server.

Currently there are several similar image analysis applications for mobile devices available on the market. LookTel[14], is a commercial application dedicated for iPhones, that is intended to recognize an object within the camera's field of view whose image was previously stored in a local database. The application is intended to help visually impaired people to recognize household objects.

Meanwhile, utter! Voice Commands is a free intelligent personal assistant for Android created and developed by Brandall. With the help of this app, a lot of tasks can be carry out by simply speaking to your tablet or smartphone. It can handle the basics, like navigating to a destination, speaking the time and weather, post to Facebook and Twitter, and set alerts or calendar appointments. utter! is accessible through speaking which is similar to what the proposed application aim to achieved, an application that would be easy for the visually impaired people to use because of its voice command.

3. System Architecture

Shown in Fig. 1 is the architectural design of the proposed system. The system is deployed in an Android phone which enables it to recognize objects and give location and direction information. Google API is integrated to determine locations while Open CV is incorporated for identifying objects. When the camera is activated and directed to the object the application identifies the object through speech feedback. In location and direction information, the system fetches data from Google API through post request, JSON decode will then be used to get the location and direction information and the output data is fetched and presented to the user through voice feedback.

In Fig 2, the activity sequence of the system is shown. The system has two main functions, the safest route and object recognition. Upon start-up if the user chose the function safest route the user will set the location and

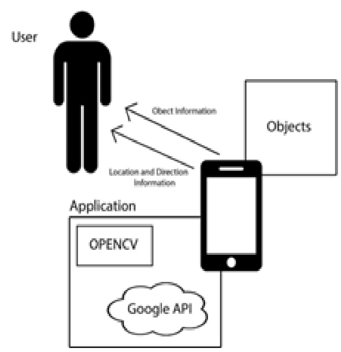

Fig. 1 System Architecture

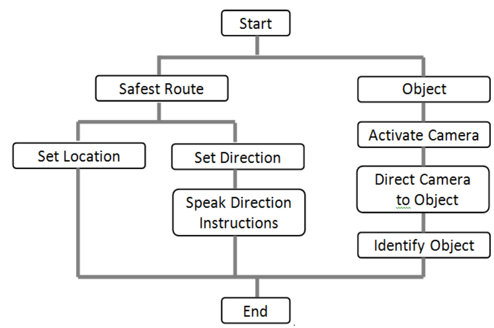

Fig. 2 Activity Sequence of the System

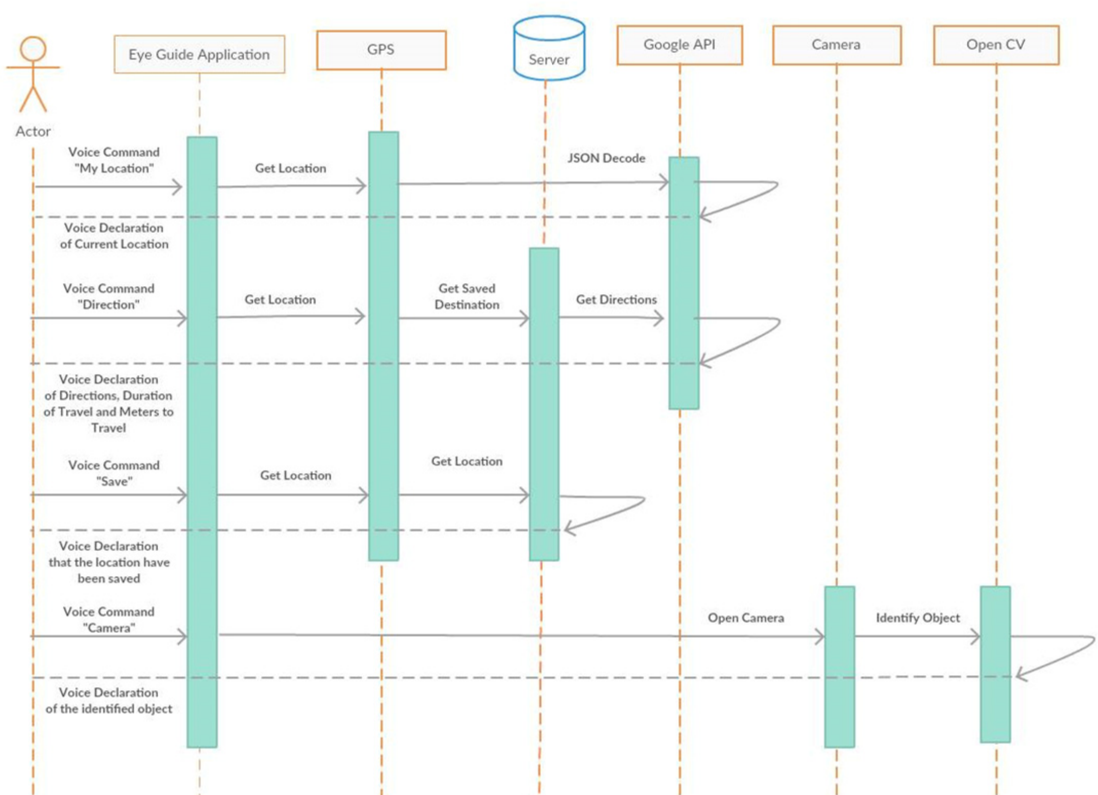

Fig. 3 Sequence Diagram of the System

destination and the output response will be a speech feedback. Otherwise, if the user chose object recognition the camera will be activated and the user will direct the camera to the object for it to be identified.

The sequence diagram in Fig. 3 shows the interaction of the objects within the system. The safest route function has other sub functions: set location, set direction, and save location. In set location, the user will say the command “My location” and the application will get its current location through GPS and the data will be sent to Google API for JSON decoding. In set direction, the same process with set location in getting the current location, after the getting the current location it will get the desired destination in the server and it will then be sent to Google API for processing of the direction, duration of travel and distance to travel. In saving location, the application will get the current location and save it to the database. Object recognition is the second function of the application, where the user will say the command “camera” to open the camera and identify the object using Open CV.

4. System Implementation and Evaluation

The researchers’ proposed system has two main functions: object recognition and safest route and has a feature of voice commands/feedbacks. Object Recognition is implemented using Open CV, a library of programming functions, and its Haar-like Features component is specifically used. As for Safest Route, the system will inform the user through speech feedback of his current location and direction instructions for his desired destination. The user can pin his usual locations and have a codename for it. Once pinned in the Google Map, the coordinates of the pinned location will be saved in the server. In this function, the system fetches data from Google API through post request, JSON decode will then be used to get the location and direction information and the output of the data fetched

Shown in Fig. 4 is the deployment plan for the system. The system is implemented in an android phone. Open CV is integrated for the object recognition and Google Maps for identifying locations. The user can access the system anytime as long as they are connected to the internet. The application is connected to a cloud server that has a database for stored locations. Google Authentication is used to access Google API for the location and direction.

Shown in Fig. 5a-5d, are sample screenshots of the system showing the home interface, the direction interface, and the location interface.

The Home User Interface in Fig. 5b, provides the time and date, as well as the instructions. There is a help button on the upper right corner and a compass parallel to it. The circle with a microphone logo is the button pressed to say

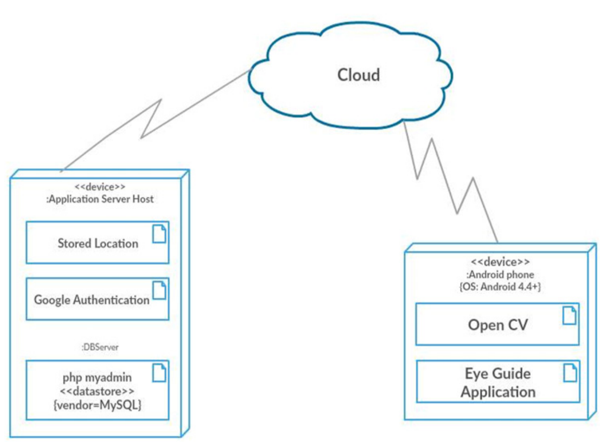

Fig. 4 Deployment Diagram of the System

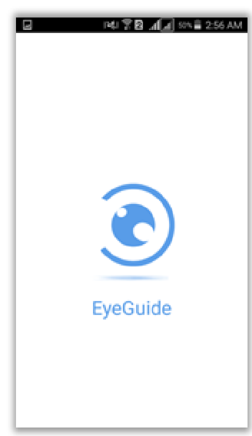

Fig. 5a Start-up UI

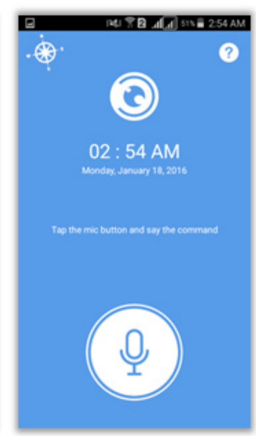

Fig. 5b Home UI

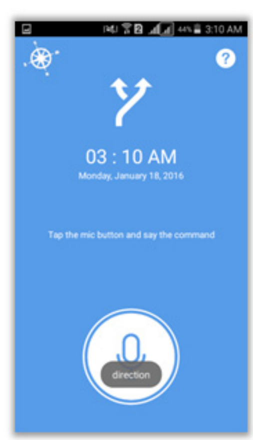

Fig. 5c Direction UI

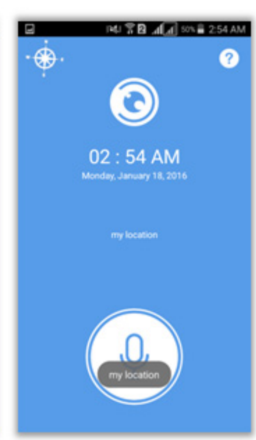

Fig. 5d My Location UI

the application's command.

Shown in Fig. 5c, is the Direction User Interface. The command is done if the user wants to know the directions to his desired location. Meanwhile, the My Location command can be useful is the user wants to know his current location; its interface is shown in Fig. 5d. The application will only identify the directions of his desired location if it was pinned. Table 1 shows the voice commands used in navigating and manipulating the application. For example, the command for pinning locations is "save", for object recognition, the command is "camera" to activate the camera, if the user wants the application to repeat its response he could use the command "pardon".

In Fig. 6, during object recognition, once the camera is activated and directed to the object it will identify the object.

Table 1 List of Voice Commands.

Commands	Description
Save	To save the location of the user.
My Location	To know the current location of the user.
Direction	To know the directions on how to go to a certain place.
Camera	To activate camera mode.
Time	To know the time.
Date	To know the date.
Pardon	To repeat the instruction.
Exit	To go back to home.

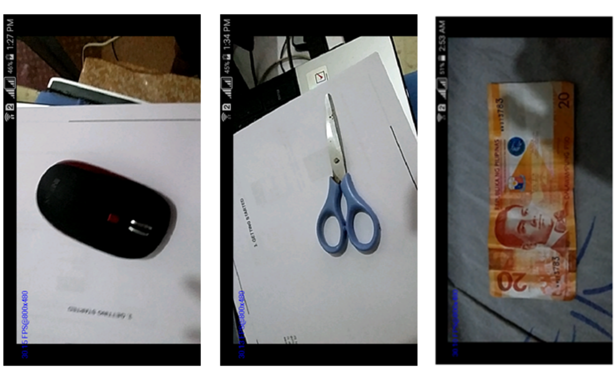

Fig. 6 Object Recognition User Interface

The system was evaluated by the visually impaired according to ISO 9126 Software Quality Standards. The criteria of the evaluation involved the following: functionality, reliability, usability, efficiency, maintainability, and portability.

As shown in Table 2, the visually impaired respondents have given an overall rating of "satisfactory". The visually impaired evaluated the application "very satisfactory" in terms of functionality, which indicates that the application is suited to serve its purpose well. The ability of the system to be easily operated by a given user in a given environment was given the rating "satisfactory" for its usability. In terms of efficiency it was evaluated "very satisfactory". The system is efficient in response to a given task. As for the reliability and maintainability, both were evaluated "satisfactory" which means that the system is able to recover from failure and can operate again. In terms of portability the visually impaired rated it "fair" because of their incapacity regarding the technicality of the system.

Table 2 ISO 9126 Software Quality Standards Evaluation Results

Characteristics	Mean	Description
Functionality	3.46	Very Satisfactory
Usability	3.11	Satisfactory
Efficiency	3.33	Very Satisfactory
Maintainability	3.22	Satisfactory
Reliability	3.22	Satisfactory
Portability	2.33	Fair
Grand Mean:	3.11	Satisfactory

5. Conclusion

The study aimed to develop a system that could aid visually impaired individuals perform significant tasks through object recognition and identifying locations through Google Maps. Computer vision-based techniques that utilize visual tags to recognize objects have only recently begun to be employed in the domain of assistive technologies for the visually impaired; however, the results reported so far are very promising and clearly demonstrate the potential of such system.

Tests were made to make sure that the system is fully functional and has fulfilled its objectives and architectural design. The visually impaired users as respondents evaluated the application and it turns out that system functionality has the highest mean rating. This indicates that the safest route and object recognition features perform the expected functions. However, the visually impaired respondents recommend that the system should not be internet dependent for its accessibility, and they also suggested that a 'Talk Back' feature should be integrated for them to navigate it easily due to the concern of the application's portability.

감사의 글

The authors would like to thank the rest of the development team (Avis A. Aligarbes, John Alphonse A. Jimena, Lorenz A. Lorca, Doryn Marie L. Tallador) for their significant contributions to the study.

References

[1] Pascolini, D., Mariotti S. P., “Global estimates of visual impairment: 2010,” British Journal Ophthalmology, 2011.
[2] “Elimination of Avoidable Blindness Report by the Secretariat,” World Health Organisation, Fifty-sixth World Health Assembly, 2003.
[3] Visual impairment and blindness: Fact sheet number 282, http://www.who.int/mediacentre/factsheets/fs282/en/, WHO media center, 2012.
[4] Jae-Han P., Seung-Ho, P., Moon-Hong, B., “An intelligent navigation method for services robots in the Smart Environment,” International Conference on Control, Automation and Systems, pp. 494-498, 2007.
[5] Espinosa, M. A., Ungar, S., Ochaíta, E., “Comparing Methods for Introducing Blind and Visually Impaired People to Unfamiliar Urban Environments,” Journal of Environmental Psychology, pp. 277-287, 1998.
[6] Schmidt, F., “Fundamentals of Sensory Physiology,” Springer, 1979.
[7] Murad M., Rehman A., Shah A. A., Ullah S., Fahad M., and Yahya K. M., “RFAIDE-An RFID Based Navigation and Object Recognition Assistant for Visually Impaired People,” 7th International Conference on Emerging Technologies, Islamabad, Pakistan, 2011.
[8] “A Buzzing Sonar Watch That Helps Blind People Navigate,” http://www.wired.com/2014/11/buzzing-sonar-watch-helps-blind-people-navigate/
[9] Crandall W., Brabyn J., Bentzen B. L., and Myers L., “Remote infrared signage evaluation for transit stations and intersections,” J. of Rehabil Res Dev., vol. 4, pp.341-355, 1999.
[10] Nanayakkara S. C., Shilkrot R. and Maes P., “EyeRing: An Eye on a Finger,” Intl. Conf. Human Factors in Computing, 2012.
[11] Fink, W., “DORA: Digital Object-Recognition Audio-Assistant for the Visually Impaired,” http://autonomy.caltech.edu/biomedicine/dora.html.
[12] Singhal, M., Shukla, A., “Implementation of Location based Services in Android using GPS and Web Services,” IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, 2012.
[13] Radhika Rani Ch., Praveen Kumar A., Adarsh D., Krishna Mohan K., Kiran K.V., “Location Based Services In Android,” International Journal of Advances In Engineering & Technology, 2012.
[14] LookTel Recognizer, http://www.looktel.com/ recognizer

저 자 소 개

Arnel N. Secondes

Arnel N. Secondes is the Chair of the Division of Information Systems at the College of Information and Communications Technology, West Visayas State University, Iloilo City, Philippines. He finished his BS in Computer Science at Computer College of the Visayas, and his MS in Computer Science at Iloilo Science and Technology University. His research interest is in the area of Optical Mark Recognition technology, Business Process, Software Engineering, Business/Systems Integration, Enterprise Systems (ERP), Information Systems and Decision Support Systems (DSS).

Nikki Anne Dominique D. Otero

Nikki Anne Dominique D. Otero received her B.S. degree in Information Technology from West Visayas State University in 2016. Her interests lie in the field of Multimedia Technologies, Software Development, and Healthcare Systems.

Frank I. Elijorde

Dr. Frank I. Elijorde received his B.S. degree in Information Technology and M.S. degree in Computer Science from Western Visayas College of Science and Technology (now Iloilo Science and Technology University), Philippines, in 2003 and 2007 respectively. He received his Ph.D. in Information and Telecommunications Engineering from Kunsan National University, Korea, in 2015. Currently, he is an Associate Professor at the College of ICT in West Visayas State University, Iloilo City, Philippines. His research interests include distributed systems, cloud systems, data mining, ubiquitous computing, and context-aware systems.

변 영 철 (Yung-Cheol Byun)

Dr. Yungcheol Byun is a full professor at the Computer Engineering Dept. (CE) at Jeju National University. His research interests include the areas of Pattern Recognition, RFID & IoT Middleware, and Artificial Intelligence. Outside of his research activities, Dr. Byun has been hosting international conferences, CNSI, ICESI, and also serving as a chair in various international conferences. He received his Ph.D. from Yonsei University in 2001. Before joining the current university, he was a senior researcher of Electronics and Telecommunications Research Institute (ETRI).