Integrating Computer Vision Object Recognition with Location Based Services for the Blind

Conference Paper · June 2014

DOI: 10.1007/978-3-319-07446-7_48

5 authors, including:

Hugo Fernandes
Institute for Systems and Computer Engineering, Technology and Science (INESC …
42 PUBLICATIONS **610** CITATIONS

SEE PROFILE

Paulo Costa
Instituto Politécnico de Leiria
14 PUBLICATIONS **172** CITATIONS

SEE PROFILE

Hugo Paredes
Institute for Systems and Computer Engineering, Technology and Science (INESC …
192 PUBLICATIONS **1,090** CITATIONS

SEE PROFILE

Joao Barroso
Universidade de Trás-os-Montes e Alto Douro
287 PUBLICATIONS **1,519** CITATIONS

SEE PROFILE

Integrating Computer Vision Object Recognition with Location Based Services for the Blind

Hugo Fernandes[1], Paulo Costa[2], Hugo Paredes[1], Vítor Filipe[1], João Barroso[1]

[1]INESC-TEC and UTAD – University of Trás-os-Montes e Alto Douro, Vila Real, Portugal

{hugof, hparedes, vfilipe, jbarroso}@utad.pt

[2]School of Technology and Management, Polytechnic Institute of Leiria, Leiria, Portugal

paulo.costa@ipleiria.pt

Abstract. The task of moving from one place to another is a difficult challenge that involves obstacle avoidance, staying on street walks, finding doors, knowing the current location and keeping on track through the desired path. Nowadays, navigation systems are widely used to find the correct path, or the quickest, between two places. While assistive technology has contributed to the improvement of the quality of life of people with disabilities, people with visual impairment still face enormous limitations in terms of their mobility. In recent years, several approaches have been made to create systems that allow seamless tracking and navigation both in indoor and outdoor environments. However there is still an enormous lack of availability of information that can be used to assist the navigation of users with visual impairments as well as a lack of sufficient precision in terms of the estimation of the user's location. Blavigator is a navigation system designed to help users with visual impairments. In a known location, the use of object recognition algorithms can provide contextual feedback to the user and even serve as a validator to the positioning module and geographic information system of a navigation system for the visually impaired. This paper proposes a method where the use of computer vision algorithms validate the outputs of the positioning system of the Blavigator prototype.

Keywords: location-based, services, blind, navigation, computer vision, object recognition

1 Introduction

Moving from one place to another is a difficult challenge that involves obstacle avoidance, staying on street walks, finding doors, knowing the current location and keeping on track through the desired path, until the destination is reached. While assistive technology has contributed to the improvement of the quality of life of people with disabilities, with major advances in recent years, people with visual impairment still face enormous limitations in terms of their mobility. Nowadays, navigation systems are widely used to find the correct path, or the quickest, between two places. In

adfa, p. 1, 2011.

recent years, several approaches have been made to create systems that allow seamless tracking and navigation both in indoor and outdoor environments. However there is still an enormous lack of availability of information that can be used to assist the navigation of users with visual impairments (or other kinds of impairment), as well as a lack of sufficient precision in terms of the estimation of the user's location. All these factors combined, maintain a situation of large disparity between the availability of such technology among users who suffer from physical limitations and those who do not suffer such limitations. To address the task of finding the user location in indoor environments several techniques and technologies have been used such as sonar, radio signal triangulation, radio signal (beacon) emitters, or signal fingerprinting. All these technologies can be, and have been, used to develop systems that help enhancing the personal space range of blind or visually impaired users. In the case of outdoor environments, some hybrid systems have been proposed that use GPS as the main information source and use radiofrequency for correction and minimization of the location error. Another hybrid approach may be the use of computer vision algorithms to work together with global positioning. In a known location, the use of object recognition algorithms can provide contextual feedback to the user and even serve as a validator to the positioning and information system modules of a navigation system for the visually impaired. This paper proposes a method where the use of computer vision algorithms validate the outputs of the positioning system.

Section 2 presents related work in the field of navigation systems and the determination of the user's location, whether indoor or outdoor. Specifically some works are presented which focus on user's with visual impairment. Section 3 describes how the prototype addresses the problem of creating a navigation system for the blind. Section 4 presents some final considerations regarding the techniques used in object recognition.

2 Related work

Location and navigation systems have become very important and widely available in recent years as a tool for finding the quickest or optimal route to a specific destination or simply to retrieve contextual information about the environment and nearby points-of-interest (POI). To determine the user's location most of these systems use the Global Positioning System (GPS), but since GPS signals are greatly degraded inside of buildings they only work well in outdoor environment.

To address the task of finding the user location in indoor environments several techniques and technologies have been used such as sonar, radio signal triangulation, radio signal (beacon) emitters, or signal fingerprinting. All these technologies can be, and have been, used to develop systems that help enhancing the personal space range of blind or visually impaired users [1]. Another technology widely used in this context is Radio-Frequency Identification (RFID). RFID tags are built-in with electronic components that store an identification code that can be read by an RFID tag reader. In recent years some research teams [2][3][4] have developed navigation systems based on this technology. In the case of outdoor environments, some hybrid systems have been proposed that use GPS as the main information source and use RFID for correction and minimization of the location error. The research team at the University of Trás-os-Montes e Alto Douro (UTAD) has an extensive work in terms of accessibility and rehabilitation. In the last few years, the team has given major focus to visual impairment and on how existing technology may help in everyday life applications. From an extensive review of the state of the art and its best practices, three main projects have been developed: the SmartVision [5][6], Nav4B [7] and Blavigator [8] projects.

The main goal of the SmartVision project was to develop and integrate technology for aiding blind users and those with severe visual impairments into a small portable device that is cheap and easy to assemble using off-the-shelf components. This device should be extremely easy to carry and use, yet providing all necessary help for autonomous navigation. It should be stressed out that the device was designed to be an extension of the white cane, not a replacement, and to be "non-invasive", issuing warning signals when approaching a possible obstacle, a point-of-interest or when the footpath in front is curved and the heading direction should be adapted.

In this sense, the SmartVision prototype addressed three main applications: (1) local navigation for centering on footpaths etc. and obstacle avoidance, in the immediate surroundings, but just beyond the reach of the white cane; (2) global navigation for finding one's way; and (3) object/obstacle recognition, not only on the shelves in a pantry or supermarket, but also outdoor: bus stops, taxi stands, cash machines (ATM) and telephone booths. The Nav4B project aimed to be an extension of the work done in the SmartVision project.

The new prototype, developed by the Blavigator project is built with the same modular structure as the SmartVision project. The Blavigator project aimed at creating a small, cheap and portable device that included all the features of the SmartVision prototype, with added performance optimization.

3 Blavigator Prototype

3.1 Modular structure of the prototype

The prototype is composed by the same modules that compose the SmartVision prototype: Interface, Geographic Information System (GIS), Navigation, Location, Computer Vision/Object Recognition and the central Decision Module (Fig. 1).

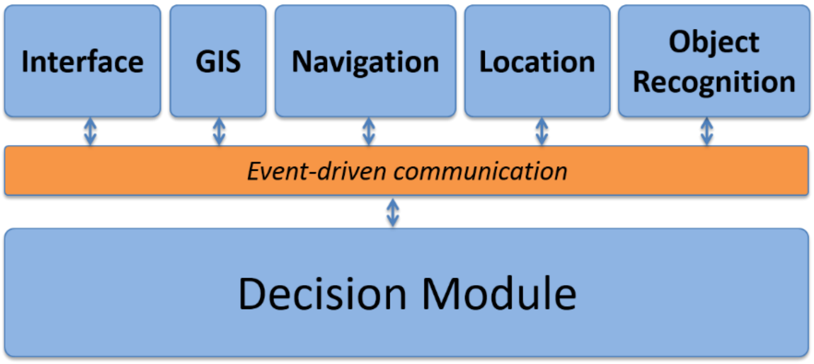

Fig. 1. Modules of the Blavigator prototype.

The Interface Module is responsible for the user interface. The system outputs information through vibration (haptic actuator present in the electronic white cane) and through text-to-speech technology (per user request or in the presence of relevant contextual information).

The Geographic Information System keeps a local representation of the data stored in a dedicated remote server supported and maintained by the Blavigator project. The Geographic Information System (GIS) in the remote server stores data about all the points-of-interest, layers, etc., in a MySQL database. All the required CRUD (create, read, update and delete) operations are available through a web application interface (Fig. 2).

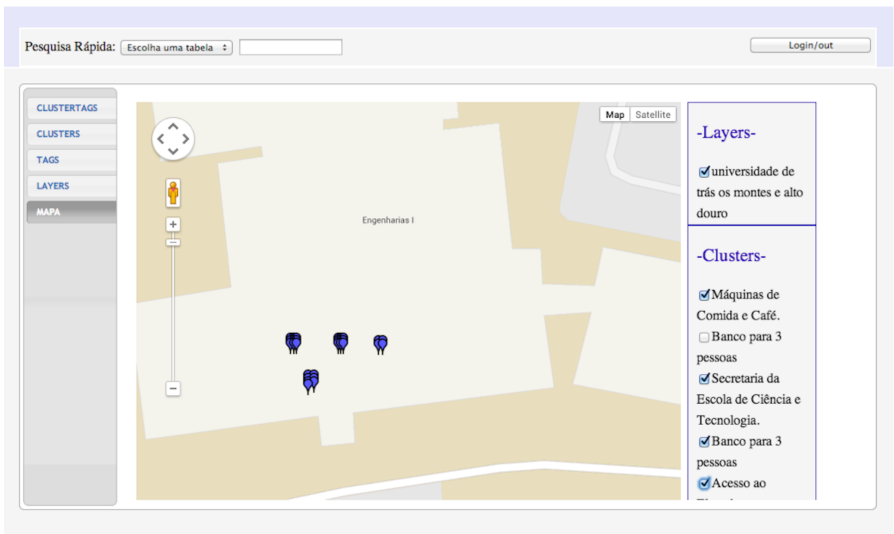

Fig. 2. Points of Interest (POI)/Objects/Layers stored in the Geographic Information System.

The Navigation Module, by request of the Decision Module (and user input), uses data from both the Location Module and the GIS to calculate the optimal route to a specific destination (point-of-interest), to guide the user through the calculated route and to provide contextual information about the user's surroundings.

The Computer Vision/Object Recognition Module is used to validate the information retrieved by the Location Module in combination with the GIS Module. In other words, when the Location Module informs the system about a new coordinate, the central Decision Module queries the GIS about which objects are expected to be found in the user's surroundings. Then, knowing which nearby objects may be found, object recognition algorithms and the device's built-in camera are used to search for these objects. Once an object is found, this information is used to validate the outputs from the Location Module and, at the same time, feed the user with contextual information, extending the reach of the traditional white cane.

3.2 Navigation software, user setup and system interface

When a new geographic coordinate in sensed by the Location Module, the mobile application (software) can handle this input in one of two ways (or modes) according to the user's requirements: Navigating or Touring. If the user wants to navigate to a specific point-of-interest, the system uses the current coordinate and the destination coordinate to calculate a route between the two points. Then, each input from the Location Module triggers the Navigation Module to keep the user on his track. On the other hand, if the user wants to stroll around, as if sightseeing, the system uses the current coordinate to query the GIS Module about relevant features (points-of-interest) in the user's surroundings. There is an 'alert' level within the software that can be modified by the user to filter the types of alerts he wants to be warned about. This is a support system that must not be intrusive neither an obstacle by itself, overriding the will of the user.

The interface between the user and the mobile software application is bidirectional (Fig. 3).

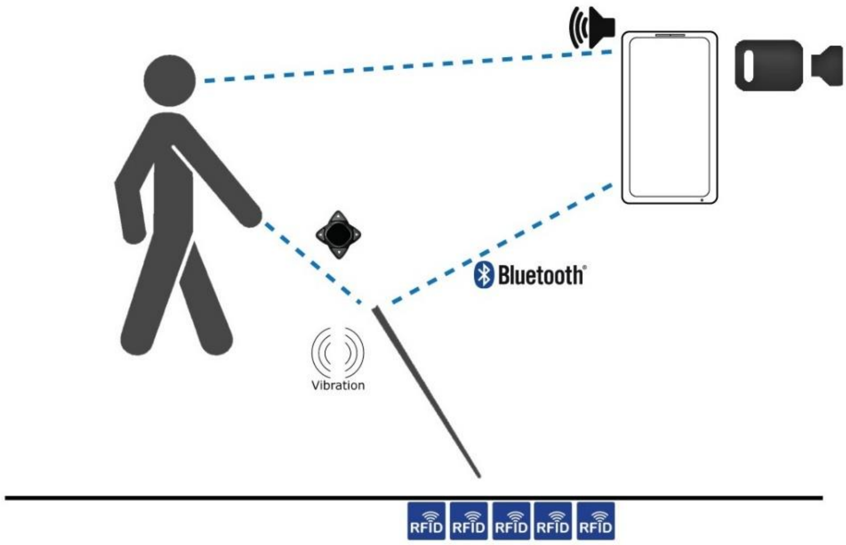

Fig. 3. User setup and system interface.

The user interfaces the software using a small joystick placed conveniently in the white cane. The joystick works in four directions and has a press button as well (a total of five switches). This joystick is used to navigate the software application and the inputs are sent via Bluetooth. When the software needs to provide feedback to the blind user, text-to-speech technology is used. This type of audio cues is used only at specific locations (such as dangerous areas or important points-of-interest) or by user request, if he feels the need to know with more detail the contents of his surroundings.

No interface is made with the mobile device itself, directly. The mobile application is designed to be interfaced exclusively using text-to-speech and the joystick.

3.3 Computer Vision Object recognition

Regarding the Computer Vision module, which is the focus of the proposed work, its function is to extract information about the surrounding environment so as to enhance the navigation of a blind person. This paper proposes the use of algorithms for recognizing certain elements often found in the environment and which contribute significantly to a secure mobility of the user. These algorithms are used to create a library of routines able to individually recognize the elements in the surrounding environment and are activated according to the elements expect to be found in the place where the user is located, at every moment. In addition, object recognition is often a computationally demanding task which gets more and more demanding as the number of elements to be detected the image gets larger. This module must be able to provide indications in real time.

To test this idea, three different objects commonly found in the Engineering building of the University of Trás-os-Montes e Alto Douro where geo-referenced and stored in the GIS of the Blavigator system. The main idea is that the Object Recognition module is activated each time the GIS module indicates that there is a geo-referenced object (or set of objects) nearby. This query is made each time the Location Module feeds the system with a new geographical coordinate.

The Blavigator prototype was developed using a smartphone, due to its high computational power, small size and low price. The smartphone runs Google's Android operating system. Due to this fact, the object recognition module was built using OpenCV (Open Source Computer Vision Library) and developed using Android Developer Tools.

Figures 4, 5 and 6 show the results of the detection of these objects using a simple template matching technique.

Fig. 4. Detection of trash can in the corridor (case #1).

Fig. 5. Detection of fire extinguisher in the corridor (case #2).

Fig. 6. Detection of ATM in the main hall (case #3).

It is important for the recognition of picture elements to be activated according to the type of elements known to be close to the blind person at each specific moment.

4 Final remarks

As there are often several important elements in the environment simultaneously is very important that algorithms are simultaneously very robust and very fast. The elements in the examples shown in Figures 4, 5 and 6 were detected using a simple template matching technique. This technique was used for its simplicity in these initial tests and to prove the concept presented in this paper. However this technique is also known to be extremely time consuming and also known to fail the detection when there is a big mismatch between the template's size and pose when compared to the actual size and of the object in the image.

To overcome this issue the extraction of this information in the image can be made using "Machine Learning" techniques, in particular through the use of Artificial Neural Networks (ANN), "Support Vector Machines" (SVM) for pattern recognition, or through the use the "Wavelet Transform." The robustness of the detection of features in the image can be increased using techniques invariant to scale, using methods such as SIFT ("Scale-invariant feature transform") or SURF ("speeded Up Robust Features").

References

1. P. Strumillo, Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation, in: 3rd Conference on Human System Interactions (HSI), pp. 17 – 24, (2010).
2. S. Chumkamon, P. Tuvaphanthaphiphat and P. Keeratiwintakorn, "A Blind Navigation System Using RFID for Indoor Environments", 5th International Conference on In Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Thailand, Vol. 2, pp. 765-768, (2008).
3. S. Willis and S. Helal, "RFID Information Grid for Blind Navigational and Wayfinding", Proceedings of the 9th IEEE International Symposium on Wearable Computers, Osaka, pp. 34-37, (2005).
4. E. D‟Atri, C. Medaglia, E. Panizzi and A. D‟Atri, "A system to aid blind people in the mobility: A usability test and its results", Proceedings of the Second International Conference on Systems, Martinique, pp. 35, (2007).
5. J.M.H. du Buf, J. Barroso, J.M.F. Rodrigues, H. Paredes, M. Farrajota, H. Fernandes, J. José, V. Teixiera, M. Saleiro "The SmartVision navigation prototype for the blind" Proc. Int. Conf. on Software Development for Enhancing Accessibility and Fighting Info-exclusion (DSAI), 2010, pp. 167-174.
6. Fernandes, H.; du Buf, J; Rodrigues, J.M.F; Barroso, J.; Paredes, H.; Farrajota, M.; José, J.. "The SmartVision Navigation Prototype for Blind Users", Journal of Digital Content Technology and its Applications 5, 5: 351 – 361, (2011).
7. Fernandes, Hugo; Faria, José; Paredes, Hugo; Barroso, João. 2011. "An integrated system for blind day-to-day life autonomy", The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility, Dundee, Scotland, UK, (2011)

8. Fernandes, H.; Adão, T.; Magalhães, L.; Paredes, H.; Barroso, J.. 2012. "Navigation Module of Blavigator Prototype", World Automation Congress 2012, In Proceedings of the World Automation Congress, Puerto Vallarta., (2012)