

NAVIG: Navigation Assisted by Artificial Vision and GNSS

Brian FG Katz¹
¹ LIMSI-CNRS
Orsay, France
(+33) 01 69 85 80 80

brian.katz@limsi.fr

Philippe Truillet²
² IRIT
CNRS & Université Paul Sabatier
Toulouse, France
(+33) 05 61 55 74 09

{truillet/jouffrais}@irit.fr

Simon Thorpe³
³ CerCo
CNRS & Université Paul Sabatier
Toulouse, France
(+33) 05 62 17 28 00

simon.thorpe@cerco.ups-tls.fr

ABSTRACT

Finding ones way to an unknown destination, navigating complex routes, finding desired inanimate objects; these are all tasks that can be challenging for the visually impaired. The project NAVIG (Navigation Assisted by artificial VIsion and GNSS) is directed towards increasing the autonomy of visually impaired users in known and unknown environments, exterior and interior, large scale and small scale, through a combination of a Global Navigation Satellite System (GNSS) and rapid visual recognition with which the precise position of the user can be determined. Relying on geographical databases and visually identified objects, the user is guided to their desired destination through spatialized audio rendering, always maintained in the head-centered reference frame. This paper presents the overall project design and architecture of the NAVIG system.

Keywords

Assisted navigation, guidance, spatial audio, visually impaired.

1. INTRODUCTION

The report “Inequality of Chances” (*Inégalité des Chances*), a result of a national study between 2003 and 2005 led by the Canadian National Institute for the Blind, shows that at least 50% of blind people require assistance in their daily life [3]. A recent literature review of existing electronic mobility aids (2007-2008) for visually impaired by [7] identified more than 140 products, systems, and assistive devices while providing details on 21 commercially available systems. The different systems were divided into two categories: (1) obstacle detection, orientation, or micro-navigation and (2) macro-navigation. Micro-navigation or obstacle avoidance systems are primarily concerned with “where” and obstacle is, and not so much “what” it is. Macro-navigation systems are almost exclusively GPS navigation based systems which have been adapted for visually deficient users. These systems are primarily limited by the precision of the positioning system and the details in the geographical database.

No commercial products were reported which were able to detect and locate specific objects without the necessity of pre-equipping them with dedicated sensors (*e.g.* RFID tags). Some research systems are under study in this area. A visual based system, incorporating a handheld stereo camera and WiFi based tracking for indoor use has been presented by [6]. This system relies on the use of 3D models of precise objects and defined spaces in order for them to be identified, greatly limiting its use outside of the designed environment.

Direct sensory substitution systems, which directly transform data from images to auditory or tactile devices without any

interpretation, can be used for rudimentary obstacle avoidance or for research on brain plasticity and learning. But, these systems have a large learning curve and are very hard to use by the blind people in any practical sense.

Based on the needs of blind people, the NAVIG¹ project (2009-2011) aims to design an assistive device which provides assistance for two problematic situations: far-field navigation (macro-navigation combined with obstacle avoidance) and near-field guidance (object identification and grasping guidance) [4]. Audio guidance will be provided through the generation of an audio augmented reality environment via binaural rendering, allowing the full exploitation of the human perceptual and cognitive capacity for spatial hearing.

NAVIG follows a method of participatory design, and aims to permit visually deficient individuals to move about towards a desired destination in a sure and precise manner, without interfering with normal behavior or mobility. In addition to being an aid to mobility and orientation, the system permits the localization and grasping of objects without the need to pre-equip them with electronic tags. The combination of these two functions provides a power assistive device.

2. SYSTEM OVERVIEW

The different objectives of NAVIG will be attained by combining input data furnished through satellite based geolocalisation and an ultra-rapid image recognition system (with detection rates on the order of 10 ms per object possible). Guidance will be provided using spatialized audio rendering with both text-to-speech and specifically designed sonification metaphors.

The system prototype architecture is divided into several functional elements structured around a multi-agent framework using a communication based on the IVY middleware [2]. This modularity allows for rapid prototyping of different modules while still achieving a high degree of overall fidelity. With this architecture, agents are able to connect or disconnect dynamically to different data stream on the IVY bus. The general architecture of the system is shown in Figure 1.

The main operating elements of the NAVIG system can be divided into three groups: data input, user communication, and internal system control. The data input elements consist of a satellite-based geopositioning system, acceleration and orientation sensors, map databases, an ultra-rapid image recognition platform connected to a series of head mounted cameras, and a data fusion module. User communications are handled predominantly through

¹ <http://navig.irit.fr>

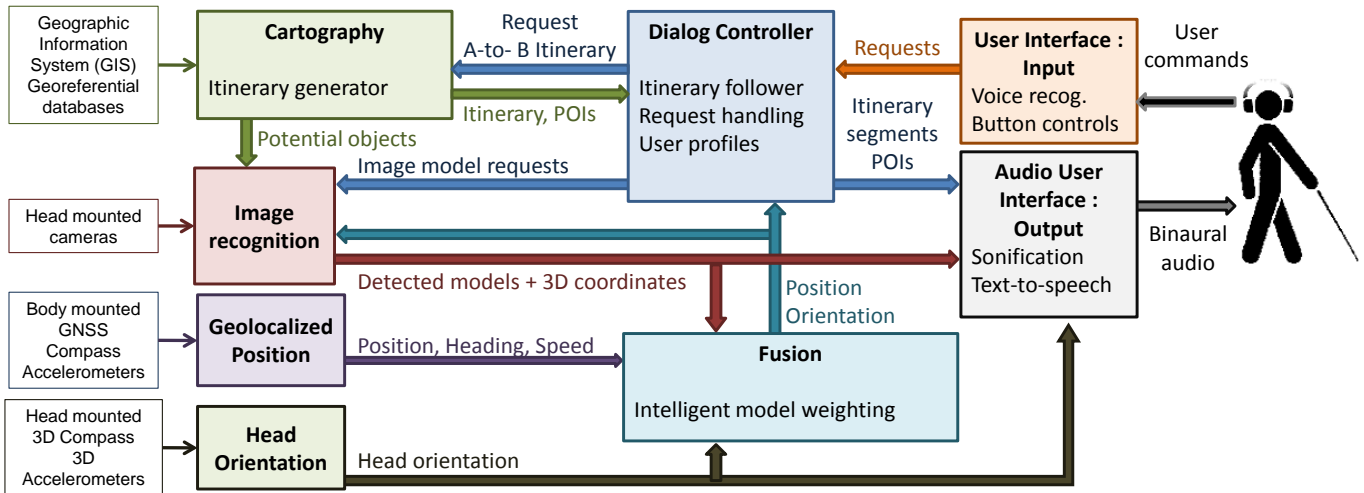


Figure 1. NAVIG system architecture overview.

a voice recognition system for input and an audio rendering engine using text-to-speech and conceptual trajectory sonification for output. Internal system control is handled by the central Dialog Controller which provides instructional information for all the separate modules based on the current situation, user needs and requests, etc.

3. FUNCTIONAL OVERVIEW

While each element can be seen separately, the operation of the system relies on the collaboration of the different distinct elements to perform several distinct functions.

3.1 Precision Geolocalisation

While high precision GNSS systems exist, they are too cumbersome and costly to be used in a portable assistive device. One of the major limitations to consumer satellite based geolocalisation systems (e.g. GPS) is their degree of precision and robustness in difficult conditions, a major criticism of most macro-navigation assistance systems. The NAVIG system aims to greatly reduce positional errors, increasing precision and stability through the use multiple data inputs (satellite, orientation, acceleration, image recognition) which are integrated with cartography and other location based databases.

The fusion of positional information from the recognition and geolocalisation systems in real-time is a novel approach which results in an improvement in positional precision of the user. The approach is to combine satellite data from the GNSS element (based on the *Angéo* system developed by the project partner NAVOCAP) with visually located landmarks. Using a detailed database and search algorithms, the position of the user can be triangulated. The integration of accelerometers provides added stability in separating tracking jitter from actual user motion. Head orientation tracking can be employed to further aid the image recognition search algorithms, as one can predict which objects will be in the field of view of which camera.

3.2 Object Identification

One of the truly novel functions of the NAVIG system is as a micro-navigation aid. The image recognition platform (based on the library developed by the partner SpikeNet Technology) is central to the functioning of this phase of navigation as it is the

only source of data input. While geolocalisation is used for arriving at a destination (i.e. the post office), it cannot be used to find the door, mailbox, etc.

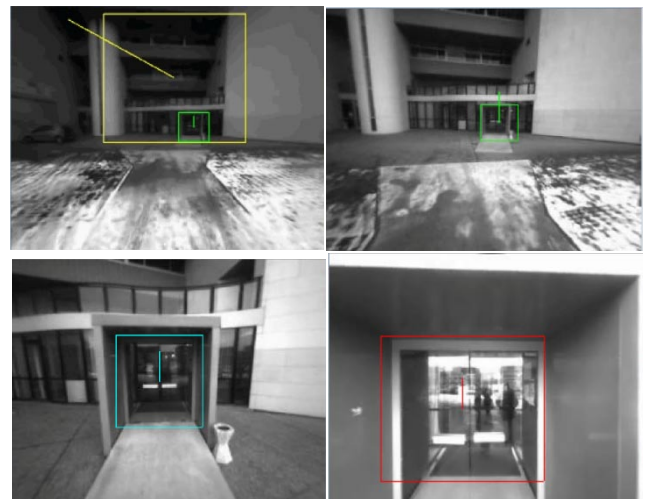


Figure 2. Multi-scale object identification.

Figure 2 presents a series identification results as the user approaches the destination (here the entrance to the IRT laboratory). The geolocalisation guidance system is used to approach the general entry to the building; subsequently the system detects the entrance location, and as the user approaches the detection is refined to locate the doorway, the door, and so on.

The rapidity of the image recognition platform is a vital part of the success of the overall system, in order to provide stable and coherent information while the user is in motion. The current software implementation is capable of detecting multiple objects in just tens of milliseconds. Figure 3 shows an example of multiple object recognition, where a water bottle, coffee cup, and webcam were are simultaneously detected and located in 14.1 ms.



Figure 3. Multiple object recognition.

Using multiple cameras it is possible to extract the spatial position of objects located in the field of view of multiple cameras using stereoscopic disparity. The current design will incorporate 6 cameras, providing a field of view on the order of 240°. As visually impaired users are accustomed to perceiving their surroundings through spatial audition, and not vision, they are not naturally limited to the same limited frontal field of perception as sight persons, and the assistive system should avoid imposing new limitations. Each camera will be equipped with its own fast Digital Signal Processor which will allow for the implementation of preprocessing parts of the vision algorithm in the embedded chips, allowing for greatly improved speed for such a large field of view. A detailed description of the image recognition system is provided in [5].

3.3 User Guidance

Once the user position has been determined, and the destination location or object identified, the primary task of the assistive system is to guide the user in a safe and reliable manner. Figure 4 presents a conceptual typical situation. The direct trajectory between user and destination, indicated by the black arrow, is not the correct trajectory. This type of error is consistent with many pedestrian guidance systems, which are more point-to-point and not true aids in urban settings. The ideal trajectory path, shown in red, provides a safe path. In addition, obstacles, such as poles and cars, as well as important landmarks such as curbs and pedestrian crossings, are well indicated. An ideal navigation aid would take into account all these elements.

3.3.1 Cartography trajectory

The primary task in trajectory determination and guidance is to establish the path to be taken. This is performed using traditional means, but including additional criteria relevant for visually impaired users, such as sidewalk width, pedestrian crossings, general points of interest (POI), and other obstacles.

During the course of navigation, the planned trajectory remains flexible, taking into account any additional information that arrives from the image recognition system. In addition to obstacles, this could include temporary route changes, such as road work sites, but can also include refinements such as precise positioning of pedestrian crossings that, while they may exist in georeferential databases, may not coincide with their true position. Integration of visual identification of such elements improves the accuracy and safety of the navigational aid.

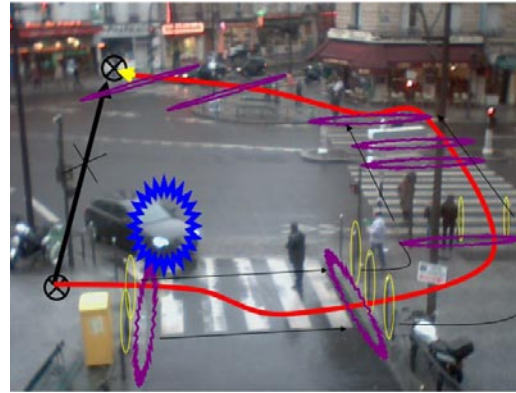


Figure 4. Conceptual pedestrian guidance trajectory.

3.3.2 Micro-navigational trajectory

Unlike pedestrian navigation, trajectory determination in the near field, or in indoor situations, is more difficult, as the only data source is that from the image recognition platform. As a first attempt, direct point-to-point guidance can be used to attain any requested object. Nevertheless, work is underway to develop intelligent navigation paths even in these situations. A contextual example of a typical situation is shown in Figure 5. The user requests the *knife*. While the object can be easily and quickly identified, and its position determined, in this context there are a number of obstacles. In addition, a *knife* has a preferred orientation for grasping, and it would preferable if the assistive device was aware of the orientation of the object and would direct the user accordingly.

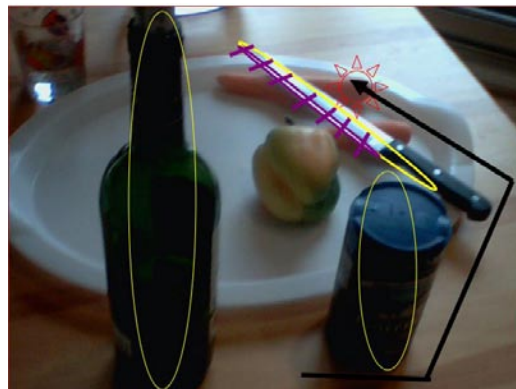


Figure 5. Conceptual grasping guidance trajectory.

3.3.3 Spatial Audio

While location of the user, of obstacles, and determination of the proper trajectory to follow to attain the intended goals is fundamental properties of the system, this information is not useful if it cannot be exploited by the user. The NAVIG system proposes to make use of the human capacity for hearing and specifically spatial audition by presenting guidance and navigational information via binaural 3D audio scenes [1].

In contrast to traditional devices which rely on turn by turn instructions, the NAVIG consortium is working towards providing spatial information to the user concerning the trajectory, their position in it, and important landmarks. With the

goal of providing the user the information necessary to construct accurate cognitive maps of the environment, the users will become more confident in their displacement.

Visually impaired persons are already exploiting their sense of hearing beyond the capacities of most sighted people. Using this same modality channel to provide additional and important information is novel, and the design of such information to minimize cognitive load and to maximize understanding is a key component of the system.

They are many instances where textural verbal communication is optimal, such as indicating street names or landmarks. At the same time, a path is not a verbal object, but a spatial object, and the exploitation of auditory trajectory rendering can be more informative and more intuitive than a list of verbal instructions. The ability to have a global sense of the trajectory is also highly desirable. In order to minimize the masking of real environmental sounds, tests are underway using bone conduction headphones and open air-tube headphones for the sound rendering, including digital filters to optimize the quality of binaural rendering on these particular devices.

3.4 Dialog Controller / User Interface

The central core of the NAVIG system is the Dialog Controller. Acting as the directing agent between the incoming data from different sensing modules, the controller must weigh the different inputs in relation to the task at hand, and the environment, in order to aid the fusion of this data. The dialog controller is also responsible for receiving the input commands of the user (via voice recognition), forwarding instructions to the proper sensing modules (such as choosing which image models should be loaded into the active search), and tailoring the auditory output to the user's preferences (sonification profile, notification rate, etc.).

During navigation, the dialog controller determines what notification messages are to be presented to the user. The different types of dialog and instructions for navigational assistance for visually impaired users, such as the choice and use of landmarks or directional orientation cues, are different from those employed in traditional navigational aids. As such, a specific guidance grammar is being defined and developed for implementation in the system.

4. PROTOTYPE V.1

At the end of the first year of the project NAVIG, the first functional prototype has been successfully tested on a simple scenario. This system, employing 2 cameras and operating on a traditional laptop, can be seen in Figure 6. The current version uses a Bumblebee stereoscopic camera system (Point Grey Research, Inc). The next version, with up to six 2-megapixel cameras, is currently under development.

5. CONCLUSION

This paper introduces the NAVIG assistance system for the visually impaired whose aim is increased autonomy and mobility in the context of both pedestrian navigation and object location and grasping. Combining satellite, image, and other sensor information, high precision geolocalisation is achieved. Exploiting a rapid image recognition platform and spatial audio rendering, detailed trajectories can be determined and presented to

the user for attaining macro- or micro-navigational destinations. An advanced dialog controller is being developed to facilitate usage and optimize performance for visually impaired users.



Figure 6. NAVIG Prototype V1.

6. ACKNOWLEDGMENTS

This work was supported by the French National Research Agency (ANR) through TecSan program (project NAVIG n° ANR-08-TECS-011) and the Midi-Pyrénées region through the APRRTT program. The NAVIG consortium includes IRIT, LIMSI, CerCo, SpikeNet Technology, NAVOCAP, CESDV - Institute for Young Blind, and the community of Grand Toulouse.

7. REFERENCES

- [1] Begault, D.R. 1994. 3-D Sound for Virtual Reality and Multimedia. Academic Press, Cambridge.
- [2] Buisson, M., Bustico, A., Chatty, S., Colin, F.-R., Jestin, Y., Maury, S., Mertz, C., and Truillet, P. 2002. "Ivy: un bus logiciel au service du développement de prototypes de systèmes interactifs," Proc. 14th French-speaking conference on Human-computer interaction (IHM '02), (Poitiers, 26 – 29 Nov-2002) IHM'02.
- [3] Canadian Institute for the Blind. 2005. Inégalité des chances : Rapport sur les besoins des personnes aveugles ou handicapées visuelles vivant au Canada. Technical Report. <http://www.cnib.ca/fr/apropos/publications/recherche>. (2-nov-2005)
- [4] Dramas, F., Oriola, B., Katz, B.F.G., Thorpe, S., Jouffrais, C. (2008) "Designing an assistive device for the blind based on object localization and augmented auditory reality," ACM Conf. Computers and Accessibility, (Halifax, 13-15 Oct 2008), ASSETS 2008.
- [5] Dramas, F., Thorpe, S., and Jouffrais, C. 2009. "Artificial Vision for the Blind: A Bio-Inspired Algorithm for Objects and Obstacles Detection." Intl. J. Image and Graphics, World Scientific, 2009 (in press).
- [6] Hub, A., Diepstraten, J., and Ertl, T. 2004. "Design and development of an indoor navigation and object identification system for the blind." ASSETS 2004, 147-152.
- [7] Roentgen, U.R., Gelderblom, G.J., Soede, M., and de Witte, L.P. 2008. "Inventory of Electronic Mobility Aids for Persons with Visual Impairments: A Literature Review." J. Visual Impairment & Blindness. (Nov. 2005), 702-724.