NAVIG: Guidance system for the visually impaired using virtual augmented reality

Brian F.G. Katz[a,*], Florian Dramas[b,c], Gaëtan Parseihian[a], Olivier Gutierrez[b], Slim Kammoun[b], Adrien Brilhault[b,c], Lucie Brunet[a], Mathieu Gallay[a], Bernard Oriola[b], Malika Auvray[a], Philippe Truillet[b], Michel Denis[a], Simon Thorpe[c] and Christophe Jouffrais[b]

[a]*LIMSI-CNRS, Université Paris Sud, Orsay, France*
[b]*IRIT, CNRS & Université Paul Sabatier, Toulouse, France*
[c]*CerCo, CNRS & Université Paul Sabatier, Toulouse, France*

Abstract. Finding ones way to an unknown destination, navigating complex routes, finding 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 his or her desired destination through spatialized semantic audio rendering, always maintained in the head-centered reference frame. This paper presents the overall project design and architecture of the NAVIG system. In addition, details of the new type of detection and localization device are presented in relation to guidance directives developed through participative design with potential users and educators for the visually impaired. A fundamental concept in this project is the belief that this type of assistive device is able to solve one of the major problems faced by the visually impaired: their difficulty in localizing specific objects.

Keywords: Assisted navigation, guidance, spatial audio, visually impaired assistive device, need analysis

1. Introduction

At least 50% of visually impaired people require assistance in their daily life, according to the report *Inégalité des Chances* ("Inequality of Opportunities"), a national study between 2003 and 2005 led by the Canadian National Institute for the Blind [1]. A recent literature review of existing electronic mobility aids (2007–2008) for visually impaired by [2] 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 or micro-navigation and (2) macro-navigation. Micro-navigation or obstacle avoidance systems are primarily concerned with "where" an obstacle is, and not so much with "what" it is. Macro-navigation systems are almost exclusively Global Positioning System (GPS) navigation based systems which have been adapted for visually impaired 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 [3]. This system relies on the use of 3D models of precise objects and predefined spaces in order for them to be identified.

*Corresponding author: LIMSI-CNRS, Université Paris Sud, 91403 Orsay, France. Tel.: +33 1 69 85 80 80; Fax: +33 1 69 85 80 88; E-mail: brian.katz@limsi.fr.

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 detection and avoidance, or for research on brain plasticity and learning [4,5]. But, these systems have a large learning curve and are very hard to use by visually impaired persons in any practical sense.

The NAVIG project (2009–2011) [6] aims to design an assistive device which provides aid in two problematic situations: far-field navigation (macro-navigation) and near-field guidance (object identification and grasping guidance). Information for the user 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 [7]. The incorporation of artificial vision enables the system to locate objects without the need to pre-equip them with electronic tags. The combination of both near-field and far-field functions provides a powerful assistive device for navigation guidance, situation awareness, and object identification and localization. The system should permit visually impaired individuals to move about towards a desired destination in a sure and precise manner, without interfering with normal behavior or mobility. The approach used has been based on a needs analysis of the visually impaired [8].

The aim of this paper is to present the NAVIG system architecture and specifics related to macro and micro navigational use. NAVIG follows a method of participatory design and some aspects of this process are presented here. A presentation of the first prototype is also shown, with a discussion of developments underway. Portions of this work have been previously presented [9,10].

2. User needs and participatory design

One of the first points to emphasize is that the NAVIG system is not intended to replace the white cane or the guide dog. It is rather to be considered as a complement to these devices and general Orientation and Mobility (O&M) training sessions. Therefore, the primary objective for the device is to help visually impaired people improve their daily autonomy and ability to mentally represent their environment beyond what is possible with the use of traditional assistive devices. The visually impaired most often employ egocentric

spatial representation strategies [8,11]. This implies that the various paths they know throughout their environment are well learned. The integration of these different paths into a general or global representation of the city, which is necessary for route variations such as detours, shortcuts, and reorganizations of the planned journey, often requires a significant level of additional effort on the part of the individual.

Designing an assistive device for visually impaired users implies the need to sufficiently describe the problems these individuals habitually face, and hence the needs that the device should respond to. In order to provide ongoing assistance throughout the project, the NAVIG team included the Institute for the Young Blind, Toulouse, enabling a participatory design strategy with potential users of the system. A panel of four O&M instructors and 24 visually impaired participants were identified and included in brainstorming and participatory design sessions. This work led to adapting the traditional participatory design method, which has been primarily centered on visual support, to visually impaired users [12]. In addition to participatory design sessions, the panel of visually impaired participants maintained their involvement through the different psychological and ergonomic experiments and prototype evaluations. A result of the various design sessions with O&M instructors and visually impaired volunteers has been the construction of a number of guidelines for the development of the assistive device. See [13] for further details on the participatory design phases, including a detailed task analysis of navigation by the visually impaired under various conditions.

2.1. Route planning study

A brainstorming session with six visually impaired participants (3 early blind and 3 late blind) was conducted to address the needs for journey planning for an autonomous pedestrian itinerary. Results indicate that a detailed preliminary planning phase improves cognitive representations of the environment.

The need for a customizable preliminary planning phase has been confirmed by an empirical study focusing on the interconnection between the preparation and execution phases of navigation and on the internal and external factors underlying this activity. Six visually impaired (4 early blind and 2 late blind) participants were followed by experimenters while preparing and completing a two-kilometer unfamiliar route in a residential area of Paris (see Fig. 1). During the execution of the journey, the technique of information on-demand
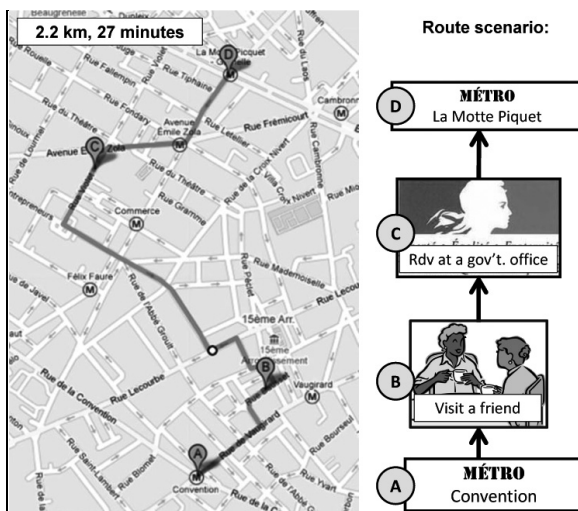
Fig. 1. Scenario used for route planning study.

Table 1
Group descriptions and group means of the number of questions posed during the *information on-demand* experiment

Group	Preparation condition	Number of questions		
		Prep.	Route	Total
G1	Without human assistance	40	146	186
G2	With human assistance	77	156	233
G3	No preparation	0	207	207

(adapted from [14]) was adopted in order to assess each participant's needs in terms of guidance. The principle of the method is to require an expert subject to perform a certain task, with a known outcome, in a given situation. In this case, the visually impaired person must perform a journey which is previously unknown to him or her. Instead of allowing subjects to act as usual (i.e. alone), they were offered the opportunity to request desired information instead of searching for the information themselves, such that the experimenter simulated a navigation assistive device. The initial situation must present a minimum level of information. During the course of the journey, the subject is able to ask questions to obtain information deemed necessary in order to make the trip as efficient and pleasant as possible (this will depend on the individual's expectations for these constraints). Analysis of the results provides information concerning the cognitive processes employed by the expert.

Participants were divided into three groups according to the level of preparatory planning performed. Along the route, each individual was followed by an experimenter, who responded to all questions. All questions were recorded and later transcribed for analysis. A description of the different groups and the number of questions posed by each group is given in Table 1 for both the preparatory and route guidance phases.

Analysis of the transcribed questions allowed for the classification of the different questions posed by all participants. Several categories of questions were created which allowed for a classification according to their relation to the route in progress. The instances of the different categories of questions with respect to the

different groups highlights the relation between route planning and the type or amount of required information for an efficient and pleasant trip. The category descriptions and a summary of the results are shown in Table 2.

This design study has resulted in the definition of several elements for integration into the NAVIG system. Detailed analysis of the types of questions posed has resulted in guidelines for the information to be provided to the user during the course of navigation, taking into account the presence, or absence, of preparatory route planning. The advantages observed for users who performed preparatory route planning have lead to the definition of an overview function for the guidance system. Users of the NAVIG system should be able to prepare and preview their upcoming journey at home, accessing an overview or a more detailed description of the path to follow, receiving knowledge about the different points of interest along the path including appropriate spatial cues to locate their specific path with respect to the general surrounding environment, and being able to preview the actual route guidance.

Information concerning the route segment, and the timing of when to present this information, is a component of the route guidance system. A certain degree of flexibility in these rules should be designed into the system. The adjustment of these parameters will be the subject of further testing with the actual system.

2.2. Route difficulty assessment

Instructors and potential users have specifically insisted on the fact that an assistance device must be highly customizable. Concerning the routes offered by the system (see Section 4.6), this implies that the most independent and confident users can request the shortest path, even if it involves various difficulties, whereas the most conservative users can request a longer but more cautious or prudent path route.

A means of identifying or classifying route points in order to privilege or avoid them in route planning was seen as necessary. A two-hour brainstorming session was conducted with six participants. A consensus was

Table 2

Classification of questions posed during the route guidance phase of the *informa-tion on-demand* experiment and the proportion of each category as a function of experimental group

Category	Question pertaining to...	*G1*	*G2*	*G3*	Overall
T-1	previous route segment	3%	3%	3%	3%
T0	a precise point	34%	24%	27%	28%
T1	current route segment, next step	48%	32%	38%	38%
T2	the next 2 route segments	11%	9%	12%	10%
T3-4	the next 3 to 4 route segments	4%	11%	14%	10%
T5	more than 4 segments in advance, global route planning (global route concept, not precise instructions)	0%	21%	6%	11%

Table 3

Predominant pedestrian obstacles encountered with associated mean difficulty scores with standard deviation (rated on a scale from 1 to 5)

Obstacle	Difficulty score
Narrow sidewalk	4.67 ± 0.8
Large spaces, parks	4.33 ± 0.5
High obstacle	3.33 ± 1.4
Roundabout	3.17 ± 0.4
Intersection without traffic lights	2.83 ± 0.4
Wide sidewalk	2.67 ± 0.5
Sidewalk Terrace	2.17 ± 0.4
Stairway	2.00 ± 1.1

achieved for a set of route elements to privilege: simple crossings, roads with lighter traffic, larger walkways to allow for faster walking, and the shortest route if time is a chosen consideration. In addition, a set of elements to avoid was also established: plazas, roundabouts, large open areas, very large walkways due the increased presence of obstacles, very narrow walkways, shared pedestrian/automobile areas, and areas with many poles or fences.

In the interest of refining these results, and in or-der to be able to apply them to automated route se-lection, a method to attribute difficulty scores for the most common elements in a city was established. Par-ticipants were first asked to cite the three types of events/obstacles they find the most problematic in a pedestrian journey. A list of items was derived from their answers. Participants were then asked to rate the difficulty of each element on a scale of 1 to 5 (see Table 3).

In order to address specifically the issues relating to the difficulties of intersections and crossroads, a list of the different types of crossroads that can be encoun-tered was also created. The easiest type of crossroad for a visually impaired is a three-branch crossroad (a ⊤-form crossroad), associated with a difficulty coef-ficient of 1. A 4-branch crossroad (×-form) was at-tributed a difficulty coefficient of 2, with 1 unit added

for each additional branch. Roundabouts and special crossroads were attributed a coefficient corresponding to the number of branches plus 1 for each additional difficulty element, such as a median strip or bus lane. Each itinerary generated by the system (see Section 4.6) can therefore have an associated overall difficulty score computed from the various crossroads and obstacles present. Those scores can be used to create a list of proposed paths which the user will be able to select from, based on their own criteria of confidence, time available, or the acceptable level of difficulty.

2.3. Ideal guidance information

Discussions concerning the audio guidance provided by an ideal system emphasized the need to control the amount of information presented, and that the degree of information should take into account the current con-ditions and individual user needs and preferences. In general, the amount of information provided should be minimal without excess, presenting only what is neces-sary and sufficient to aid the user. The information pro-vided should be highly efficient and minimally intru-sive. These guidelines have been considered in the de-sign of the audio feedback for the user (see Sections 4.6 and 4.5).

Overall, it has been assumed that a device fitting the specific needs of visually impaired users in preparing pedestrian itineraries will contribute to the construction of a more reliable cognitive spatial representation as well as a reduction in the level of anxiety during pedes-trian journeys in unfamiliar environments, which was a common concern expressed during interviews with O&M instructors and visually impaired individuals.

3. System overview

The different objectives of NAVIG will be attained by combining input data furnished through satellite
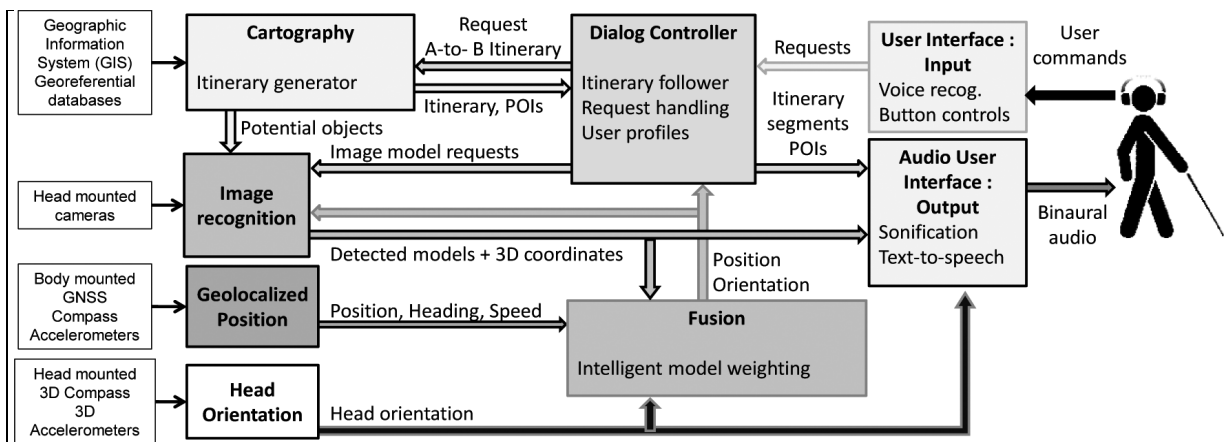
Fig. 2. NAVIG system architecture overview.

based geolocalisation and an ultra-rapid image recognition system. Guidance will be provided using spatialized audio rendering with both text-to-speech and specifically designed semantic sonification metaphors.

The system prototype architecture is divided into several functional elements structured around a multi-agent framework employing a communication protocol based on the IVY middleware [15]. With this architecture, agents are able to connect or disconnect dynamically to different data streams on the IVY bus. The general architecture of the system is shown in Fig. 2. 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, Geographic Information System (GIS) map databases, an ultra-rapid module processing images from a series of head mounted cameras, and a data fusion module. User communications are handled predominantly through 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 the different modules based on the current situation such as user needs and requests.

4. Functional overview

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

4.1. Dialog controller / user interface

The central node of the NAVIG system is the Dialog Controller, acting as the directing agent between the incoming data from different sensing, processing, and output modules. The dialog controller is also responsible for receiving input commands from the user (via voice recognition and a simple keypad interface), forwarding instructions to the proper sensing modules (such as selecting which objects should be detected by the vision module), and tailoring the auditory output to the user's preferences (sonification profile, notification rate, etc.). The interaction with the dialog controller is done via voice recognition input (Dragon Naturally Speaking) and synthetic speech output (SRA – Speech Recognition Agent, compatible MS-SAPI). A voice menu was implemented which represented the different navigation possibilities mentioned during the brainstorming sessions with potential users. To improve the performance of speech recognition, voice recognition was based on grammar and specific dictionaries handled dynamically by SRA and the dialog controller. The usability of this interaction module has been evaluated for a group of eight users [16].

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. One important issue that must be considered is the process of selection of objects and landmarks that the system must render in situations
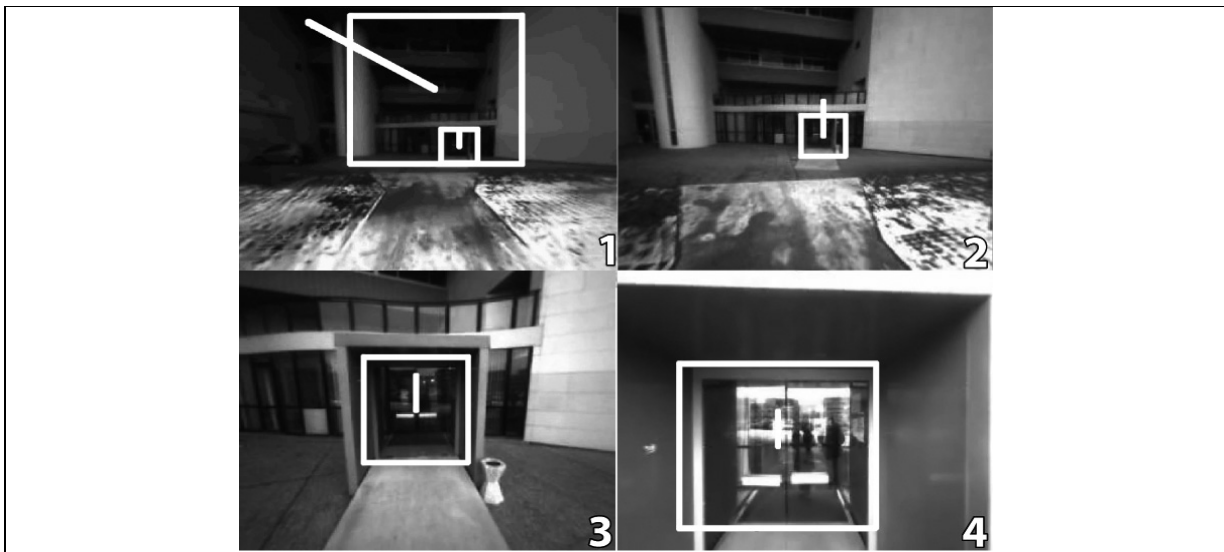
Fig. 3. Multi-scale object identification highlighting global model from a distance to precise detection of the doorway upon arrival.

with a large degree of information. This point will be systematically explored using an experimental platform which includes the different modules of the NAVIG device. Different scenarios and types of guidance during navigation tasks will be evaluated in this virtual environment platform (see [17]). This experimental platform will help to design the best output interaction relative to guidance and spatial cognition, without impairing regular O&M skills. The results will then be evaluated in real scenarios with the NAVIG device.

4.2. Object identification and localization

One of the truly novel functions of the NAVIG system is as a micro-navigation aid. Research on biological vision has shown that the visual system is able to detect (specific) objects extremely rapidly [18]. Computational modeling of this process has led to the development of a biologically inspired image processing library (developed by the project partner SpikeNet Technology [19]) for object detection and localization. The image recognition platform is central to the system functions during micro-navigation as it is the only source of data input. While geopositioning systems are used for arriving at a destination (i.e. the post office), it cannot be used to find the door, mailbox, etc.

As an example, a location-based guidance system could first be used to approach the general entry to a building; subsequently the vision unit would detect the entrance location, and as the user approaches the entrance the detection would be refined to locate the

doorway, the door, and so on. Figure 3 presents a series of identification results as a user approaches the destination (here the entrance to the IRIT laboratory).

The SpikeNet image recognition algorithm provides high robustness to noise and light conditions, providing ultra-rapid object localization in two dimensions. Tolerances for a given model for variations in rotation are on the order of ± 8 to $10°$ and in size of roughly $\pm 10\%$. Employing dynamical model modifications that change the current set of active models to adapt to the current target object means that only tens of models may be required for a given identification task. 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. For a given image size, processing time is strongly dependent on the size of the visual forms that the system is trying to find. Essentially, the recognition engine will be very fast for finding shapes that are large in size, and progressively slower for smaller ones. Processing time also depends directly on the number of different forms that are being tested simultaneously. As an illustration, a performance test of the recognition engine on a PC notebook (Intel i7 820QM processor (1.73 Ghz) and 4 GB memory) achieved a stable analysis speed of 15 fps on a 320×240 px image stream while concurrently searching for 750 visual shapes of size 120×120 px. The number of forms that can be tested at 15 fps drops to 250 for shapes 60×60 px pixels across, and 65 when the dimensions reach 30×30 px. A detailed description of the image recognition system is provided in [20].

Using multiple cameras, it is possible to extract the spatial position of objects simultaneously located in the field of view of more than one camera using stereoscopic algorithms. As visually impaired users are accustomed to perceiving their surroundings through spatial audition, and not vision, they are not naturally restricted to the same limited frontal field of perception as sighted persons. It is clear that the assistive system should avoid imposing new limitations if at all possible.

Micro-guidance navigation, based on target recognition, as well as object grasping, relying on the same mechanisms, constitutes the first function of the visual component in the NAVIG project. This role is essential as it restores a functional visual-motor loop allowing a visually impaired user to move the body or the hand to targets of interest. A second function, detailed in [21], uses visual targets with known GPS coordinates as visual landmarks to refine the satellite base geopositioning. The approach is to estimate the user's position from the detection of a visual target and to fuse it with the other sensors as described in the next section. If, for example, the storefront of a store with known GPS coordinates is in the camera's field of view, it is then possible, using the 3D position of the object in the camera reference frame (obtained through stereo vision), and the orientation of the head (acquired by an Inertial Measurement Unit), to estimate the geolocation of the user. This positioning method can provide an estimate relying exclusively on vision, in case of GPS signal loss, or can be integrated to a larger fusion scheme. The visual landmarks used for positioning are termed Visual reference Points (*VP*) and are stored in the Geographic Information System (*GIS*) with their coordinates. They consist of static elements of the environment with known geographic position which aid in accurately positioning the user when they are detected by the vision unit. These points can be real objects, such as road-signs, shop fronts, etc., as well as visual features without semantic meaning: a part of a wall, a distinctive configuration of buildings, etc. With a multiple camera system the visual models needed to detect objects do not necessarily have to include all possible viewing angles: assuming the user walks along the street, an orthogonal view of a shop front might be sufficient. But if needed, or desired with the aim of increasing the probability of visual detection, the construction of SpikeNet models from a single image can be accomplished through automatic perspective transformations.

Regarding the creation of models, at present this has been done manually from recorded videos of the evaluation test site. Development is underway for creating a tool which will allow a semi-supervised generation of models. A future step will be the fully automatic construction of models in a new area, based on services similar to Google Street View [22]. For instance, the Topographic Department of the City of Toulouse has provided the project with 3D visual recordings of all the streets of the city, combining eight different views taken from car-mounted cameras at one meter intervals, as well as laser data which allows one to obtain the GPS coordinates of any point in these images. With this database, it could be possible to randomly search for patterns throughout the city streets which are distinctive enough so as not to trigger false detections in neighboring streets, and to automatically store them as Visual Reference Points with their associated coordinates.

4.3. Multi-sensor data fusion

The measurement of physical quantities such as position, orientation, and acceleration rely on sensors which inherently report approximated values. This fact, in addition to occasional sensor dropout or failure, results in a given system receiving inaccurate or incomplete information to some degree. As such, the NAVIG system employs a collection of different sensors to obtain the same information, e.g. position. These different position approximations must then be combined to provide the best estimate through a data fusion model. There are three main issues identified in sensor data fusion:

- Interpretation and representation: Typically handled with probabilistic descriptions [23].
- Fusion and estimation: Methods such as Bayesian estimation [24] and Kalman filtering [25] are widely used.
- Sensor management: Solutions are based either on a centralized or a decentralized sensors architecture [26].

Centralizing the fusion process combines all of the raw data from all the sensors in one main processing module. In principle, this is the best way to realize the data fusion as all the information is still present. In practice, centralized fusion frequently overloads the processing unit with large amounts of data. Preprocessing the data at each sensor drastically reduces the required data flow, and in practice, the optimal setup is usually a hybrid of these two types. Care must be taken in fusing different types of data, ensuring that transformations are performed to provide a unified coordinate system before the data fusion process.

It is important to note that different sensor systems operate with different and sometimes variable refresh rates. The sensor fusion strategy takes into account the amount of time from the last received data to automatically adjust the weights (i.e. estimated accuracy) attributed to each sensor. For example, a sudden dropout in GPS signal for more than a few seconds (in an urban canyon) would gradually reduce the weight attributed to the GPS data. Similar corrections would be applied to the other sensors, depending on the specific time characteristics of each of these sensors.

4.3.1. Pedestrian navigation

Several GPS systems equipped with different sensors have been developed to increase navigation accuracy in vehicles [27]. In these systems, the inertia of the vehicle is important and dead reckoning strategies are appropriate for predicting the position at the next time step (p_{t+1}). Furthermore, velocity and trajectory of a vehicle exhibit smooth and relatively slow variations. Finally, there is a high probability that the vehicle follows the sense of traffic known for the given side of the road. All these elements make accurate position estimation possible for vehicles, but they are not applicable in the case of pedestrian navigation.

The aim of the current fusion algorithm is to manage these two issues. First, taking into account pedestrian ways of moving. Secondly, employing user mounted cameras which will recognize natural objects in urban scenes and allow for a precise estimate of the position of the user. This solution avoids the time and expense of equipping the environment with specific instrumentation as mentioned in [28,29].

4.3.2. Multi-sensor fusion in the NAVIG system

The estimation of positional information from the image recognition and geolocation systems in real-time is a novel approach which results in an improvement in determining the precise location of the user. The approach combines satellite data from the GNSS element (based on the Angeo system [30] developed by the project partner NAVOCAP) and position estimations based on the Visual reference Points (VP) with known geographical coordinates (see Fig. 4). Using a detailed database which includes the coordinates of these landmarks, the position of the user can be geometrically estimated to a high degree of precision. The integration of accelerometers provides added stability in separating tracking jitter from actual user motion.

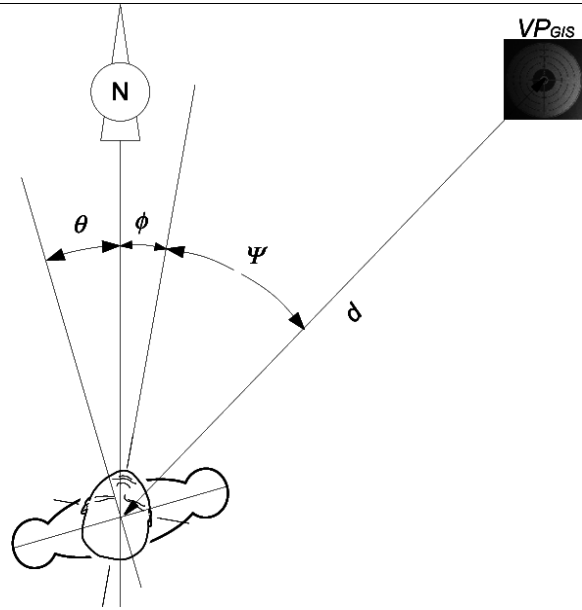The fusion algorithm uses three different inputs. First, an enhanced GPS sensor assisted by an inertial

Fig. 4. Geometrical data available for estimating user position from the visual identification of a geolocalized target: user heading, θ, head orientation, ϕ, estimated position (angle, Ψ, and distance, d) of the VP in the head-centered reference frame, and the VP coordinates in the GIS database, VP_{GIS}.

system provides accurate coordinates. Secondly, a GIS is used to verify that positions are coherent with map constraints (for example, not being in a building). Finally, the vision system provides information on any geographically located objects detected.

4.3.3. Temporal data fusion

Consider the user position estimated for a given fusion frame. In the case where one or more objects are recognized, the resulting position estimation given by the object localization, $\widehat{VP}$, is determined by the barycenter of all the coordinates of the resulting positions as described in Eq. (1) If a given object is detected multiple times in the same frame, each detection event is given equal weight, therefore attributing a greater weight to the position determined from said object.

$$\widehat{VP} = \frac{\sum_{i=1}^{n}\left(\sum_{j=1}^{m_i}(VP(i,j)|GIS)\right)}{\sum_{i=1}^{n} m_i} \qquad (1)$$

– i: n recognized VP objects
– j: m detection events for VP_i since the last frame
– $VP(i,j)$: position estimate based on a detected VP

– GIS: GIS coherency tests

Finally, the best position estimation at the current time, $\hat{p}_t$, is determined by traversing a set of points x of an area bounded by GPS, $\widehat{VP}$, $\hat{p}_{t-1}$ and choosing the one that has the highest probability, as described in Eq. (2)

$$\hat{p}_t = \overset{\text{argmax}}{x} \left(\mathbf{P}(x|GPS, \widehat{VP}, \hat{p}_{t-1}, GIS)\right) \quad (2)$$

where:

– x: current point being tested
– GPS: GPS provided position and degree of confidence
– $\widehat{VP}$: position estimate based on Eq. (1)
– $\hat{p}_{t-1}$: previous fusion process position estimate
– GIS: GIS coherency tests
– $\mathbf{P}(|)$: probability in x given the different variables.

4.4. Geographical Information System (GIS)

Accuracy of basic GPS systems is on the order of 10–15 m (RMS error). This is far from being sufficient in order to guide visually impaired users in urban environments. To determinate a user's location with higher precision [31], suggest using Differential GPS. With this technique, the absolute positional accuracy increases to around 1 m. Unfortunately, appropriate signals from landmark stations and satellites are not always available. Another technique to increase position accuracy, particularly in urban canyons with low GPS signals or dropout areas such as tunnels, relies on a dead reckoning process [32] to estimate current position based upon previously determined positions. These techniques allow a considerable improvement in positioning, but are still insufficient in order to guide a visually impaired person in unknown environments.

The GIS component is a tool for storing, editing, analyzing, manipulating, and integrating geographical information. Querying the spatial database is achieved using a Structured Query Language (SQL) to extract data information of a spatial nature. The GIS is an important component in guidance assistive devices and includes a digital database as well as software tools for route selection, tracking the traveler's path, and extraction of information regarding local features destined to aid the user.

On the user's side [33], have shown that blind users benefit from an accurate representation of the environment during navigational tasks. For this, it is proposed here to enhance the GIS database, such as by adding classes of important objects, in order to provide the visually impaired users with important specific information concerning the itinerary and surroundings, which may serve to build sparse but useful representation of the environment.

Currently, commercial GIS systems have been almost exclusively developed for car travel. Most spatial databases used in GIS have been developed without considering the needs of pedestrians in general, let alone those of the visually impaired [34,35]. Following the approach of participatory design, a series of brainstorming sessions and interviews were conducted with potential visually impaired users and orientation and mobility (O&M) instructors. The results produced a set of classified objects that should be included as part of a GIS which could be used by an assistive device. Four important classes of objects were identified:

– Walking Areas (WA): All possible pedestrian paths as defined in [36] (e.g. sidewalks, and pedestrian crossings).
– Landmarks (LM): Places or objects that can be detected by the user in order to confirm his own position along the itinerary (e.g. changes in texture of the ground, telephone poles, or traffic lights).
– Points of Interest (POI): Places or objects that are potential destinations. These are useful or interesting places allowing a better understanding of the environment while traveling (e.g. public buildings, shops, etc.).
– Visual reference Points (VP): Geolocated objects used by the vision module to refine the user position estimate.

For each object in the database, multiple classifications are possible. For instance, a bus stop is tagged as a LM because it can be detected by the user. It is also tagged as a POI if it is a potential destination or as a VP if it can be detected by the artificial vision module.

In addition, the user has the possibility to add specific locations (such as home, work, or sidewalks that become impractical in rainy situations) which are then integrated into a specific GIS corresponding to the user. The class of these objects is called Favorite Point (FP). Each point will be associated with a specific tag added by the user. Although the NAVIG system may operate as a regular GPS based guidance device, it will greatly benefit from the existence of these different classes of points in the database. Concerning WA, LM, and POI, GIS databases are becoming more and more complete and accurate. However, more effort is still required in considering the needs of people with disabilities. Of course, one can note that local authorities are enhancing and publicly providing their own GIS

database (this is already the case in some French cities.[1] On the other hand, very interesting and complementary approaches are based on social cooperation [37] proposed that annotation of GIS databases may rely on data collected "on the move" by users themselves combined with information gathered by internal sensors of the device (compass and GPS). To increase data sources and facilitate sharing between users, the authors proposed a client-server architecture with a database remotely stored on a server. The database is constantly updated and anonymously shared among users.

4.5. Spatial audio

While location of the user, of obstacles, and determination of the proper trajectory to follow to attain the intended goals are 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 [38]. The 3D sound module provides binaural rendering over headphones using a high performance spatialization engine developed at LIMSI. This engine was developed under the Max/MSP programming environment [39].

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. The goal is to provide users with the information necessary to construct accurate cognitive maps of the environment, making them more confident in their locomotion.

Visually impaired persons are already exploiting their sense of hearing beyond the capacities of most sighted people. Previous studies have shown the capacity of visually impaired individuals to create architectural maps of interior spaces through both real and virtual environment navigation using only their sense of hearing [40]. Employing the auditory modality channel to provide additional and important information, while novel, must be done cautiously. To function correctly and to be accepted by potential users, such auditory information must be designed to minimize cognitive load and to maximize understanding.

Although the use of stereo headphone is required to produce binaural 3D sound, wearing traditional head-

phones results in a certain degree of masking of real sounds which is problematic for visually impaired people. Instead, a solution employing bonephones, headphones that work via the transmission of vibrations against the side of head, transmitting the sound via bone conduction, was adopted. Previous studies have demonstrated the efficient use of bonephones within a virtual 3D audio orientation context [41]. These particular headphones, situated just in front of the ears without any obstruction of the ear canal or pinna, permit the use of 3D audio without any masking of the real acoustic environment. Because of their complex frequency response, tailored equalization is necessary in order to properly render all the spectral cues of the Head Related Transfer Function.

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 contrast to previous works of sensory substitution, where images captured by a camera are directly transformed into sound, the aim of the 3D sound module is to generate informational auditory content at the spatial position which directly coincides with that of a specific target. Various methods for semantic or informative spatial sonification has been shown to be effective in spatial manipulations of virtual objects [42] and scientific exploration and navigation tasks within large abstract datasets [43] in multi-modal virtual environments. Spatial sonification for spatial data exploration and guidance [44] and target acquisition [45] in audio and audio-haptic virtual environments without visual renderings have also been shown to be effective in previous studies.

The proposed approach attempts to exploit textural description when it is most effective, and sonification and spatilisation for conveying other infromation which is more complicated or cumbersome to convey via textural presentation. For example, the description of an intersection crossroad can be done in several ways:

– For a global description, the intersection type (X, T, or more complex) and the number of branches are provided via simple TTS, whereas the description of each branch (street name and street circulation direction) is given using spatialized TTS. For each branch, the system renders street name with spatialized speech traveling along the street indicating the direction of circulation;
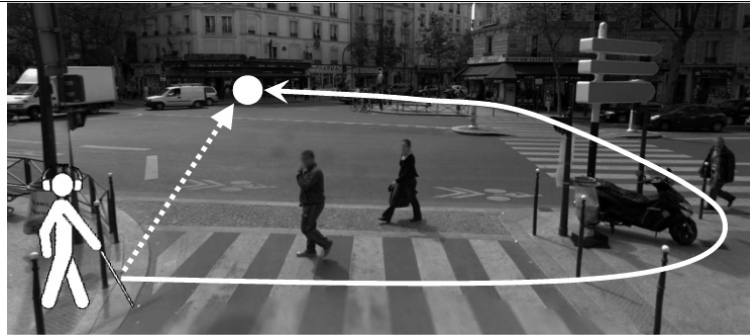
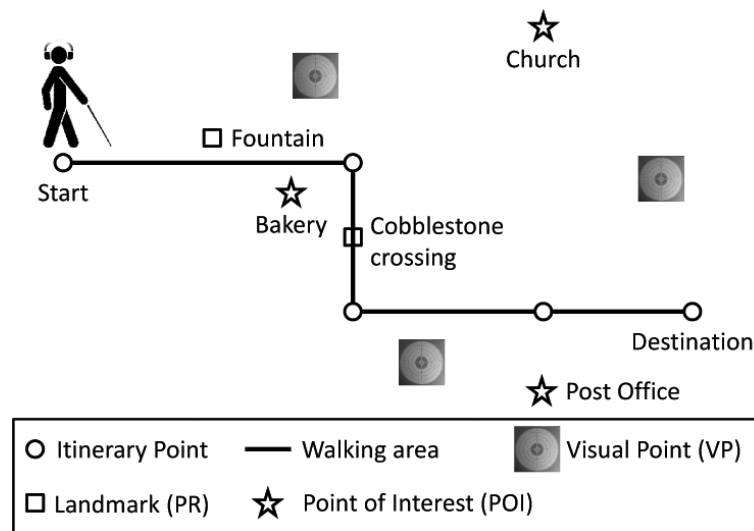Fig. 5. Conceptual pedestrian guidance trajectory.

Fig. 6. An illustration of a route.

– For guiding the user through the intersection, the system will render a spatial sound along the trajectory, adding other spatial sounds to indicate positions related to changes of ground surface and TTS to describe the crossing (number of streets, presence of traffic light or pedestrian crossing, etc.).

4.6. User guidance

Once the user position has been determined, and the destination location or object identified, the primary task for the assistive system is to guide the user in a safe and reliable manner. Figure 5 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 set-

tings. The ideal trajectory path provides a safe path, using pedestrian crossing. In addition, obstacles, such as poles and cars, as well as important objects such as curbs and pedestrian crossings, are well indicated. An ideal navigation aid should take into account all of these elements.

Traditionally, route or path selection is assumed to be the result of minimization procedures such as selecting the shortest or the quickest path. For visually impaired users, a longer route may be more convenient than a shorter route, in order to avoid various obstacles or other difficulties. These route optimization rules can vary between individual users, due to mobility training experience and other personal factors (see Section 2.2). An adapted routing algorithm for visually impaired pedestrians has been proposed to improve path choice [46]. The aim is to find the most preferred route that connects origin and destination points. The selected path can be represented as a road map contain-
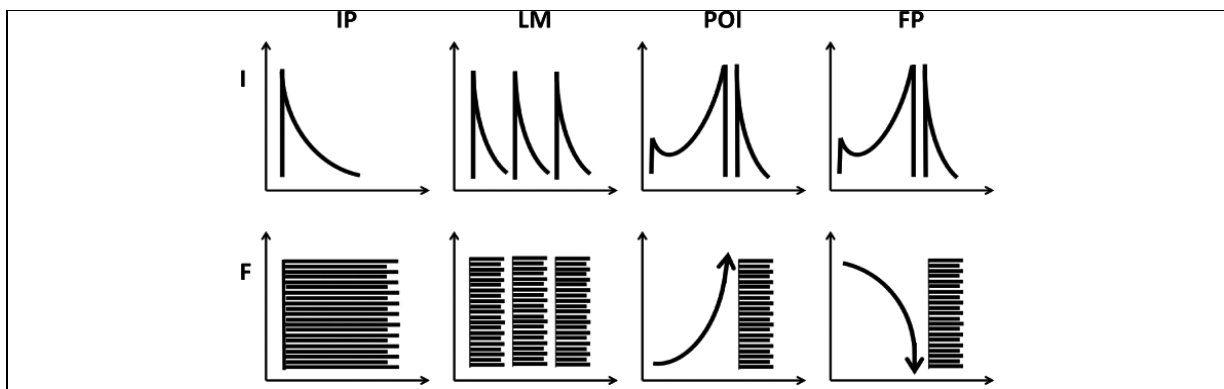
Fig. 7. Illustration of semantic sound grammar indicating (upper) intensity and (lower) frequency profiles of each element of the palette.

ing a succession of Itinerary Points (IP) linked by WA as well as a collection of nearby POI, FP, LM, and VP as defined in Section 4.4 (see example in Fig. 6).

As a function of their preferences and knowledge of the system, users have the possibility to choose different levels of detail of information and the way that this information will be presented to them. A series of brainstorming sessions and interviews with potential blind users identified at least two types of navigation which should be considered [13]. First, the *normal* mode is used for point-to-point guidance along a calculated trajectory from point A to a point B. In this mode, the user needs just a minimum of information to understand and perform the navigation task, with only the IP and LM elements being necessary. In contrast, in *exploration* mode, the user is interested in exploring a neighborhood or a specific trajectory. It would therefore be desirable to provide additional information, such as the presence of nearby bakeries, municipal buildings, or bus stops. This mode requires the presentation of IP, LM, FP, and POI. At each use, the user can personalize the presented information by selecting certain types of POI or LM that are of personal interest. To facilitate this categorical presentation, each object class is divided in several categories that can be used to filter the information (7 categories of POI, 4 LM, and 2 FP).

Different levels of verbalization are provided in the NAVIG system depending on user needs (see Section 2). If the IP are always rendered by placing a virtual 3D sound object at the next point to reach, the POI, FP and LM can be rendered using Text-To-Speech (TTS) or semantic sounds. All presented information is spatialized so that the user hears the description of each object coming from its position. Users can choose to use only TTS, a mix of spatial TTS and semantic sounds, or only semantic sound. A semantic

sound grammar has been developed to allow the user to rapidly identify and differentiate between each class of objects (IP, POI, FP, and LM) and to be informed about the subcategories within each class. An overall sound design grammar has been established so that each sound can be easily localized (i.e. large spectrum, sharp attack), to minimize the possibility of confusion between classes, and to take into consideration the superposition of the virtual soundscape and the real acoustic world. Contrary to commercially available navigation systems, the semantic sounds are not fixed. They are constructed according to the newly developed concept of morphocons (acoustics parameter motif) which are an extension of earcons (musical note motif) defined by Blattner [47]. The morphocons (a contraction of morphological earcons) allow for the construction of a hierarchical sound grammar based on temporal variations of several acoustical parameters (such as frequency, envelope, or rhythm). The resulting semantic sound grammar is illustrated with elementary temporal functions in Fig. 7 and is described as follows;

- IP: a brief sound
- LM: a rhythmic pattern of three short sounds. Rhythmic variations of this pattern allow to differentiate the LM type.
- POI: a sound whose frequency increases steadily followed by a short sound. The first sound is common to all categories of POI while the short sound can differentiate them.
- FP: a sound whose frequency decreases steadily followed by a short sound. The first sound is common to all categories of FP while the short sound can differentiate them.

Sound durations are between 0.2 and 1.5 sec. This common grammar allows for the realization of a variety of sound palettes (i.e. birds, water, musical, or
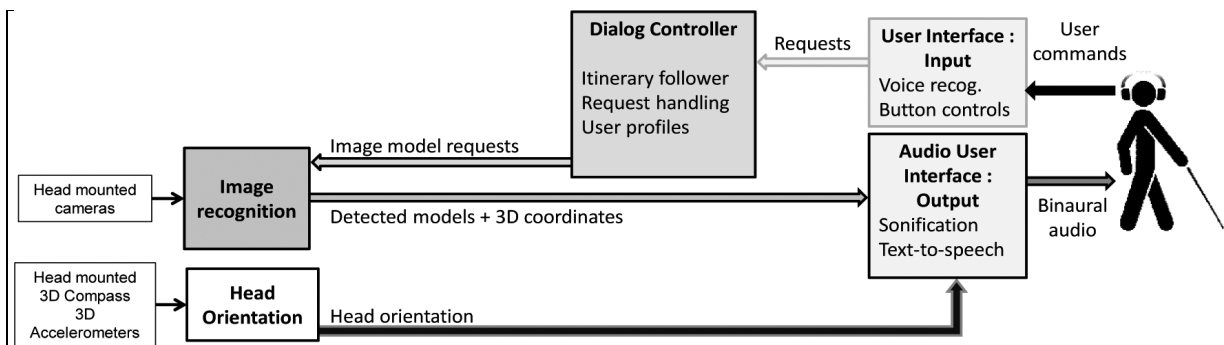
Fig. 8. NAVIG system architecture overview during near field guidance.

videogame) satisfying individual user preferences in terms of sound aesthetic while keeping the coherency with the objects to display. As such, switching between palettes should not imply any significant change in cognitive load or learning period.

5. Near field assistive mode

In its simplest form, direct point-to-point guidance can be used to attain any requested object. The task of object grasping then consists of a direct loop between the recognition algorithm detecting the target and the sound spatialization engine attributing and rendering a sound object at the physical location of the actual object. As such, the architecture for near field guidance is dynamically simplified in an attempt to optimize performance and minimize system latencies (see Fig. 8).

The user interacts with the system via voice recognition, providing the dialog module with the target name to locate, which then activates the research of the corresponding models in the artificial vision module. If detected, the position of the target is directly sent to the sonification module. Rapid image recognition of objects in the camera's field of view provides head-centered coordinates for detected object, offering built-in head tracking. For robustness in case of lost identification or objects drifting out of the field of view, a 3D head tracking device is also included to interpolate object positions, insuring fluidity and maintaining a refresh latency of no more than 10 ms with respect to head movements.

To improve the micro-navigation task, route selection should also be addressed. Unlike in the macro-scale pedestrian navigation, trajectory determination in the near field situations is more difficult as the only data source is that from the image recognition platform.

Nevertheless, intelligent navigation paths could be developed even in these situations. A contextual example of a typical situation would be to find a knife on a cluttered kitchen counter top. The user requests the knife. While the object can be easily and quickly identified, and its position determined, in this context there can be a number of obstacles, which could be in the direct path to the knife, such as seasoning bottles. In addition, a knife has a preferred orientation for grasping, and it would advantageous to the user if the assistive device was aware of the orientation of the object and would direct him or her accordingly to the handle, and not the blade.

In addition to micro-navigation, this assistive device may also serve for object recognition. Indeed, during the user-centered design sessions, visually impaired participants mentioned the recurrent problem of distinguishing among similar objects (e.g. canned foods, currency bills). The NAVIG prototype has been used in a study where the participants had to classify different European currency bills [48]. As there are few mobile systems that are able to satisfactorily recognize currency bills (see [49]), the aim was to evaluate this sub function of the vision module of the device. Due to the high-speed and robustness of the SpikeNet artificial recognition algorithm, users were able to identify 100% of the bills that were presented and performed the sorting task flawlessly. Measured response times (including bill manipulation, recognition, and classification tasks) was slightly above ten seconds per bill on average. Users were in agreement that the usability of the system was good.

6. Prototype V.1

At the end of the first year of the project NAVIG, the first functional prototype (shown in Fig. 9) has been

Fig. 9. NAVIG Prototype V1.

successfully tested on a simple scenario. This system operates on a laptop. The artificial vision module currently uses video streams from two head mounted cameras (320×240 px at 48 Hz). The prototype employs a stereo camera pair with an approximately $100°$ viewing angle allowing for object distance computations to be based on stereoscopic disparity and the calibration matrix of the lenses. The prototype hardware is based on a BumbleBee stereoscopic camera system (Point Grey Research, Inc), an XSens orientation tracker, headphones, microphone, and a notebook equipped with an Intel i7 820QM processor (1.73 Ghz) and 4 GB memory. Preliminary experiments consisted in testing outdoor guidance on the University of Toulouse campus. They involved numerous blindfolded users and one blind user that were guided on a 500 m itinerary. The aim was to test the functioning of the whole system (integration of the different modules) during real time guidance (including turn-by-turn instructions and rendering of different POI and LM). At this point, evaluation has not included usability, mounting, portability, or appearance of the NAVIG system. However, these preliminary experiments have shown that it is possible to design a wearable device that can provide fully analyzed information to the user.

The design of an assistive device for visually impaired users must take into account users' needs as well as their behavioral and cognitive abilities in spatial and navigational tasks. This first prototype will be eval-

uated in the spring of 2012 by a panel of 20 visually impaired participants involved in the project, with the headphones replaced by bonephones. 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.

7. Conclusion

This paper has introduced 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. Of course, this kind of assistive device, or *electronic orientation aid*, does not replace traditional mobility aids such as the cane or a guide-dog. It must be considered as an additional device that provides the user with important information for spatial cognition including landmarks (e.g. decision or confirmation), routes to follow (guidance), and spatial descriptions. Indeed, we suggest that a spatial environment description mode, based on 3D synthesis of the relative location of important points in the surrounding area, may help visually impaired users generate a sparsely populated yet functional mental map of the environment. This capacity of the system will be evaluated as part of ongoing ergonomic evaluations of the NAVIG system.

It is noted that the NAVIG device can also assist in indoor navigation situations. It could rely on a map of the building, including models of stationary localized targets, which could include models of signs (most of the building have specific ones), machines (elevators, vending machines), etc. Indoor guidance could then rely on embedded vision only, or be completed with dead reckoning and/or indoor localization techniques (e.g. Wi-Fi triangulation). In addition, it is suggested that the NAVIG device would be of great interest to localize many kinds of objects that can be found in a specific building (in the near-field assistive mode). For instance, furniture (phones, chairs, windows, etc.) are

usually standardized. These mobile objects could be detected in the near space around the user and may greatly improve autonomy in known and unknown local environments.

Acknowledgments

References

[1] Inégalité des chances: Rapport sur les besoins des personnes aveugles ou handicapées visuelles vivant au Canada. Canadian Institute for the Blind; 2005. Available from: www.cnib.ca/fr/apropos/publications/recherche.

[2] Roentgen UR, Gelderblom GJ, Soede M, de Witte LP. Inventory of Electronic Mobility Aids for Persons with Visual Impairments: A Literature Review. Journal of Visual Impairment & Blindness. 2008;102(11):702–724.

[3] Hub A, Diepstraten J, Ertl T. Design and development of an indoor navigation and object identification system for the blind. ACM SIGACCESS Accessibility and Computing. 2004;77–78:147–152.

[4] Auvray M, Myin E. Perception with compensatory devices. From sensory substitution to sensorimotor extension. Cognitive Science. 2009;33:1036–1058.

[5] Auvray M, Hanneton S, O'Regan JK. Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice. Perception. 2007;36(3):416–430.

[6] Project NAVIG. Available from: navig.irit.fr.

[7] Dramas F, Oriola B, Katz BFG, Thorpe S, Jouffrais C. Designing an assistive device for the blind based on object localization and augmented auditory reality. In: ACM Conf. Computers and Accessibility (ASSETS). Halifax, Canada; 2008. p. 263–264.

[8] Gallay M, Denis M, Auvray M. Navigation assistance for blind pedestrians: Guidelines for the design of devices and implications for spatial cognition. In: Representing space in cognition: Interrelations of behaviour, language, and formal models. Tenbrink T, Wiener J, Claramunt C, (Eds.). Oxford University Press (UK); (accepted).

[9] Katz BFG, Truillet P, Thorpe S, Jouffrais C. NAVIG: Navigation assisted by artificial vision and GNSS. In: Workshop on Multimodal Location Based Techniques for Extreme Navigation (Pervasive 2010). Helsinki, Finland; 2010. p. 1–4.

[10] Parseihian G, Brilhault A, Dramas F. NAVIG: An object localization system for the blind. In: Workshop on Multimodal Location Based Techniques for Extreme Navigation (Pervasive 2010). Helsinki, Finland; 2010.

[11] Noordzij ML, Zuidhoek S, Postma A. The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. Cognition. 2006;100:321–342.

[12] Brock A, Vinot JL, Oriola B, Kammoun S, Truillet P, Jouffrais C. Méthodes et outils de conception participative avec des utilisateurs non-voyants. In: 22nd French-speaking conference on HCI. Luxembourg; 2010. p. 65–72.

[13] Brunet L. Étude des besoins et des stratégies des personnes non-voyantes lors de la navigation pour la conception d'un dispositif d'aide performant et accepté (*Needs and strategy study of blind people during navigation for the design of a functional and accepted aid device*). Department of Ergonomics, Université Paris-Sud, Orsay, France; 2010.

[14] Bisseret A, Sebillote S, Falzon P. Techniques pratiques pour l'étude des activités expertes. Toulouse: Octarès Editions, 1999.

[15] Buisson M, Bustico A, Chatty S, Colin FR, Jestin Y, Maury S, et al. Ivy: un bus logiciel au service du développement de prototypes de systèmes interactifs. In: 14th French-speaking conference on Human-computer interaction (IHM '02); 2002. p. 223–226.

[16] Kammoun S, Oriola B, Truillet P, Jouffrais C. Conception et Evaluation d'un Système de Dialogue pour un dispositif de Suppléance pour Déficients Visuels. In : Ubiquité et Mobilité. Toulouse; 2011. p. 34–37.

[17] Kammoun S, Macé M, Oriola B, Jouffrais J. Designing a Virtual Environment Framework for Improving Guidance for the Visually Impaired. In: Designing Inclusive Systems (CWUAAT). Langdon, Clarkson, Robinson, Lazar, Heylighen (Eds.), Springer-Verlag, 2012 (In press).

[18] Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. Nature. 1996;381(6582):520–522.

[19] SpikeNet Technology;. Available from: www.spikenet-technology.com.

[20] Dramas F, Thorpe S, Jouffrais C. Artificial Vision for the Blind: A Bio-Inspired Algorithm for Objects and Obstacles Detection. International Journal of Image and Graphics. 2010;10(4):531–544.

[21] Brilhault A, Kammoun S, Gutierrez O, Truillet P, Jouffrais C. Fusion of Artificial Vision and GPS to Improve Blind Pedestrian Positioning. In: Intl. Conf. on New Technologies, Mobility and Security, IEEE. France; 2011.

[22] Google Street View;. Available from: maps.google.com/help/maps/streetview/.

[23] Durrant-Whyte HF. Sensor models and multisensory integration. International Journal of Robotics Research. 1988;7(6):97–113.

[24] Berger JO. Statistical Decision Theory and Bayesian Analysis. Second edition. Springer; 1985.

[25] Bar-Shalom Y. Tracking and Data Association. Academic Press Professional; 1987.

[26] Mitchell HB. Multi-Sensor Data Fusion: An introduction. Springer; 2007.

[27] Cappelle C, El Najjar ME, Pomorski D, Charpillet F. Multi-sensors data fusion using dynamic bayesian network for robotised vehicle geo-localisation. In: Intl. Conf. on Information Fusion. Cologne; 2010. p. 1–8.

[28] Park SK, Young Soo Suh TND. The pedestrian navigation system using inertial sensors. In: ICROS-SICE International Joint Conference. Fukuoka, Japan; 2009. p. 3970–3974.

[29] Bentzen B, Mitchell P. Audible signage as a wayfinding aid: Comparison of Verbal Landmarks® and Talking Signs. J of Visual Impairment and Blindness. 1995;89. p. 494–505.

[30] Angeo Technology;. Available from: www.angeo.fr.

[31] Loomis JM, Golledge R, Klatzlty RL, Speigle J, Tietz J. Personal guidance system for visually impaired. In: First annual ACM conf. on Assistive technologies (Assets '94); 1994. p. 85–90.

[32] Mayerhofer B, Pressl B, Wieser M. ODILIA-A Mobility Concept for the Visually Impaired. In: Computers Helping People with Special Needs. Miesenberger K, Klaus J, Zagler W, Karshmer A (Eds.) Springer; LNCS 2008;5105(2008). p. 1109–1116.

[33] Fletcher JF. Spatial representation in blind children. Development compared to sighted children. J of Visual Impairment and Blindness. 1980;74(10):318–385.

[34] Golledge R, Klatzky RL, Loomis JM, Speigle J, Tietz J. A geographical information system for a GPS-based personal guidance system. Intl J of Geographical Information Science. 1998;p. 727–749.

[35] Völkel T, Kühn R, Weber G. Mobility Impaired Pedestrians Are Not Cars: Requirements for the Annotation of Geographical Data In: Computers Helping People with Special Needs, LNCS 2008;5105(2008), p. 1085-1092.

[36] Zheng J, Winstanley A, Pan Z, Coveney S. Spatial Characteristics of Walking Areas for Pedestrian Navigation. In: Third Intl. Conf. on Multimedia and Ubiquitous Engineering, IEEE. China; 2009. p. 452–458.

[37] Völkel T, Weber G. A New Approach for Pedestrian Navigation for Mobility Impaired Users Based on Multimodal Annotation of Geographical Data In: Universal Access in Human-Computer Interaction. Ambient Interaction, LNCS 2007; 4555(2007), p. 575–584.

[38] Begault DR. 3-D Sound for Virtual Reality and Multimedia. New York: Academic Press; 1994.

[39] Max visual programming language for music and multimedia;. Available from: cycling74.com/products/maxmspjitter/.

[40] Katz BFG, Picinali L. Spatial Audio Applied to Research with the Blind. In: Strumillo P, editor. Advances in Sound Localization. InTech; 2011. p. 225–250.

[41] Walker BN, Lindsay J. Navigation Performance in a Virtual Environment with Bonephones. In: Intl. Conf. on Auditory Display. Limerick, Ireland; 2005. p. 260–263.

[42] Férey N, Nelson J, Martin C, Picinali L, Bouyer G, Tek A, et al. Multisensory VR interaction for protein-docking in the CoRSAIRe project. Virtual Reality. 2009;13:273–293.

[43] Vézien JM, Ménélas B, Nelson J, Picinali L, Bourdot P, Ammi M, et al. Multisensory VR exploration for computer fluid dynamics in the CoRSAIRe project. Virtual Reality. 2009;13:257–271.

[44] Katz BFG, Rio E, Picinali L, Warusfel O. The Effect of Spatialization in a Data Sonification Exploration Task. In: 14th Meeting of the Intl. Conf. on Auditory Display (ICAD). Paris; 2008. p. 1–7.

[45] Ménélas B, Picinali L, Katz BFG, Bourdot P. Audio haptic feedbacks in a task of targets acquisition. In: IEEE Symposium on 3D User Interfaces (3DUI 2010). Waltham, USA; 2010. p. 51–54.

[46] Kammoun S, Dramas F, Oriola B, Jouffrais C. Route Selection Algorithm for Blind Pedestrian. In: Intl. Conf. on Control, Automation and Systems, IEEE, KINTEX. Gyeonggi-do, Korea; 2010. p. 2223–2228.

[47] Blattner MM, Sumikawa DA, Greenberg RM. Earcons and icons: their structure and common design principles. In: Journal of Human Computer Interaction. 1989;4(1), 11–44.

[48] Parlouar R, Dramas F, Macé M, Jouffrais C. Assistive device for the blind based on object recognition: an application to identify currency bills. In: ACM Conference on Computers and Accessibility (ASSETS 2009). Pittsburgh, PA, USA; 2009. p. 227–228.

[49] Liu X. A camera phone based currency reader for the visually impaired. In: 10th international ACM SIGACCESS conference on Computers and accessibility. Halifax, Nova Scotia, Canada; 2008. p. 305–306.