NAV-VIR: an audio-tactile virtual environment to assist visually impaired people*

Marc-Aurèle Rivière[1], Simon Gay[1], Katerine Romeo[1], Edwige Pissaloux[1], Michal Bujacz[2], Piotr Skulimowski[2], Pawel Strumillo[2]

Abstract— **This paper introduces the NAV-VIR system, a multimodal virtual environment to assist visually impaired people in virtually discovering and exploring unknown areas from the safety of their home. The originality of NAV-VIR resides in (1) an optimized representation of the surrounding topography, the spatial gist, based on human spatial cognition models and the sensorimotor supplementation framework, and (2) a multimodal orientation-aware immersive virtual environment relying on two synergetic interfaces: an interactive force feedback tablet, the F2T, and an immersive HRTF-based 3D audio simulation relying on binaural recordings of real environments. This paper presents NAV-VIR functionalities and its preliminary evaluation through a simple shape and movement perception task.**

I. INTRODUCTION

Independent navigation in unknown environments is one of the greatest challenges Visually Impaired People (VIP) face. Indeed, vision plays a crucial role in gathering the information necessary to establish a mental map of the environment [1]. For VIP, travelling in an unknown area involves a tedious preparation process to memorize the distances and succession of turns needed to reach a target, and the various Points of Interest (PoI) to find along the way.

Assistive devices can improve the mobility of VIP by supplementing no-longer-accessible sources of spatial information about their environment, either as direct assistance during the actual journey, or as a tool to virtually navigate a route in order to discover an area before physically visiting it. This prior exploration can help VIP build a mental representation of an area [2]. Such a representation will enable more confident mobility within the environment and lessen the cognitive load associated with a direct exploration [3]. Traditionally, journey preparation is done relying on static 3D maps (cf. fig 1). However, these maps only display a predefined area at a fixed scale, do not inform on the user's current position and cannot provide directions towards a goal.

Figure 1. Examples of metallic (left) and thermo-printed (right) maps.

NAV-VIR is a dynamic Human-Machine Interface (HMI) able to provide a comprehensive and intuitive spatial representation through audio-tactile feedback. The NAV-VIR system is comprised of two synergetic parts (cf. fig 2): (1) an interactive tactile interface, the F2T (Force Feedback Tablet [4]), to provide intuitive spatial information about possible paths, and (2) a dynamic audio environment that provides a realistic orientation-aware 3D simulation of the audio cues the VIP are likely to encounter during their actual journey.

The high level goal of this project is to facilitate autonomous navigation for VIP by investigating how to design an intuitive and dynamic interface to assist journey preparation thanks to a simulated multimodal environment. NAV-VIR attempts to realistically simulate real world navigation by enabling its users to haptically explore a simplified map while being informed of surrounding PoIs and recognizable audio cues. It provides VIP with a minimal yet sufficient representation of the explored area (spatial gist [5]), designed to minimize the cognitive load of the user while maximizing the usability of the multimodal space representation and its transferability to the real world.

Figure 2. NAV-VIR system illustration

This article is organized as follows: section 2 summarizes related work on journey preparation for VIP, section 3 describes NAV-VIR's architecture and functionalities, section 4 briefly presents the preliminary evaluations led with our current prototype and section 5 discusses about future work.

II. RELATED WORK

Traditional 3D maps used in Orientation & Mobility centers allow to integrate spatial information from haptic exploration. However, these maps have to be designed by hand, require specific printers, and do not allow for dynamic changes of scales [1]. Moreover, they mainly rely on the path integration capabilities of VIP, which become unreliable for long journeys unless known PoI are recognized along the way.

Approaches relying on ICT devices attempt to overcome those limitations by introducing an interactive & dynamic component to the exploration, enabling the user to receive auditory and/or tactile feedback while exploring the map [6].

Those assistive solutions usually rely on the sonification of the environment, through spatialized audio cues to convey topographical information [7] or information about nearby PoIs [8]. Other solutions try to combine 2D movements with audio feedback to simulate map exploration on tabletops [9], or use mobile applications to virtually discover an area through 3D audio [8], [10], [11]. Audio "serious games" were also developed to help VIP learn the general skills needed to navigate by means of audio cues [12].

However, the aforementioned solutions frequently suffer from inaccurate distance and angle estimation by VIP [11] due to the absence of proprioceptive feedback to estimate their movement patterns. Some solutions tried to compensate this lack of sensorimotor loops by incorporating tactile information, for example through touchscreens supplemented with raised-line overlays [3], or pin-matrix displays for symbolic "you are here" maps [13]. Despite being very useful to support the mobility of VIP, these systems lack important cues for autonomous map discovery, such as the ability to find one's location on the map or the capability to guide the user along the path needed to reach a target.

In contrast, NAV-VIR's tactile interface allows users to feel their movements while exploring a map, leveraging the crucial role of haptic sensorimotor loops in spatial information integration [14]. The force feedback also allows guiding of the user's finger along a path for a machine-assisted exploration in complex environments, and make it possible to "reset" their exploration and bring back their finger to their current location on the map ("you are here" function). Furthermore, by using a multimodal feedback, NAV-VIR lowers the strain on each sensory channel, allows a more intuitive interpretation of the information provided [15] and improves the rehabilitative potential of the device [16]. The use of a minimal spatial representation, the "spatial gist" [5], [17], lowers the cognitive load of such an interface by providing precisely the information required to build mental maps. Indeed, our brain possesses specialized areas that generate complex mental maps from the integration of specific multimodal cues [18]. Exploiting those processes should facilitate building a mental map from minimal sensory information.

Based on spatial cognition models [19], [20] and previous research on intuitive map representation [11], [21], a map can be represented as a graph where nodes are mobility decision points (such as intersections) and edges are walkable paths. PoIs can then be added to this navigation graph as a different type of node, meant to facilitate learning of the environment's structure.

Finally, NAV-VIR also conveys realistic spatialized audio markers during the virtual journey. These audio cues provide location information about salient elements of the environment that VIP will be able to recognize during the actual trip. These audio PoIs (e.g. fountain, elevator, café, road) will allow VIP to improve their cognitive map and recover faster from errors in path integration.

III. NAV-VIR: ARCHITECTURE AND FUNCTIONALITIES

The general use scenario of NAV-VIR can be summarized as follows: (1) the user chooses an area to explore through audio commands, (2) the map is loaded from a GIS provider and automatically converted into its equivalent topological representation, (3) known PoIs are uploaded from the GIS (e.g. cafés, shops, roads, churches, fountains) and converted into localized sound sources which will make up the audio layer of the virtual environment. This audio-enhanced navigation graph is accessed through the F2T, which allows the user to explore the map with the use of a joystick controlled by a force feedback mechanism. The spatialized audio PoIs are conveyed through a standard pair of headphones.

A. NAV-VIR Tactile Interface

1) F2T Specifications

The F2T (fig 3) provides the means to explore graphical content by re-encoding visual or spatial information as bi-dimensional force feedback. It is an interactive and dynamic 2D haptic interface that relies on force feedback to interact with the user. The displayed information can be explored by moving a joystick whose resistance levels depend on the underlying information (e.g. slowing down or blocking the user when trying to "move" over a building).
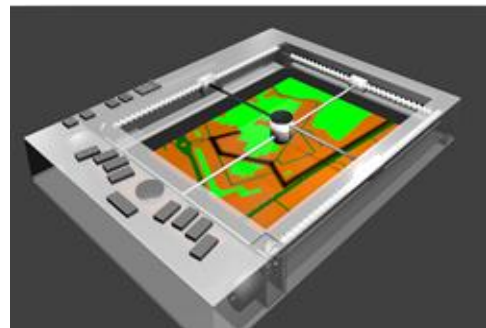

Figure 3. F2T Model

The F2T can provide both a passive (resisting or facilitating movement) or an active feedback (forced movement) during the exploration. Passive feedback is used to convey information about the map during a free exploration, while the active feedback is used to provide direct guidance along a path. Examples of simple "tactile images" can be seen in figure 4, where colors represent different types of frictions.
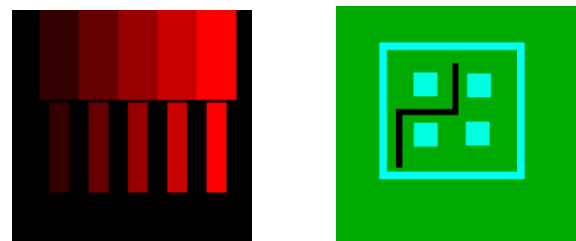

Figure 4. Examples of tactile "images": (left) continuous vs intermittent resistance gradient; (right) guided exploration on a simple map.

Based on the functional mapping between the user's actions (movement direction, speed and acceleration) and the tablet's response, we can distinguish 2 basic categories of passive feedback. The "solid" type of friction, where the user

has to push the joystick with a force above a given threshold for it to move at all. This type of friction could be used to simulate ridges or edges, or to block the user's movement if he were to try to move over a non-walkable area while exploring a map. On the other hand, a "fluid" friction represents a linear response to the user's movements: the resistance generated by the F2T is directly proportional to the force applied by the user. Different levels of fluid friction could be used to represent different types of terrains (pavement, park, …).

Furthermore, by combining different types of passive and active feedback, we can create more complex tactile textures such as "canyons", where trying to exit the "canyon" from either side will result in a fluid friction type of resistance, and the user will then be pushed back towards the bottom of the canyon. Those canyons could be used to indicate "walkable" paths or areas, allowing for a free exploration of the surroundings, while maintaining the user inside the "canyon", preventing them from getting lost on the map.
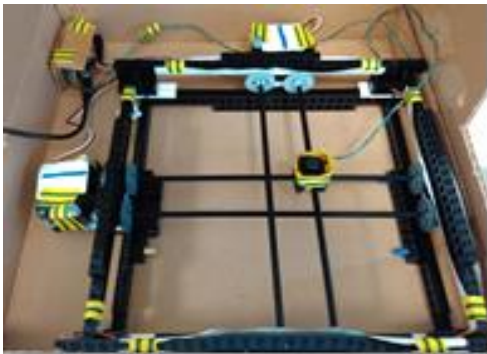
2) F2T implementation

Figure 5.   First F2T prototype (LEGO)

Our current prototype (fig. 5) is actuated by 2 servo-motors moving a flat joystick, controlled by an Arduino single-board microcontroller and a Java interface. This device can be combined with a tactile screen, allowing for a collaborative learning experience with the VIP's caretakers.

B. NAV-VIR Spatialized Audio Environment

1) Spatial Audio

The main goal of the audio interface is to aid the users in orienting themselves and understanding their environment when exploring a map, by providing realistic 3D audio cues for the PoIs they will encounter during their physical journey.

Stereo panning is the most basic technique used for sound spatialization, however its main issue is the in-head localization. More advanced techniques such the Head Related Transfer Functions (HRTF) enable the generation of 3D externalized sounds by means of directional filter banks. HRTFs can be generic or personalized for a listener, however, the latter require measurements in an anechoic chamber [22]. Binaural recordings, although much less flexible than HRTF-based spatial audio, are also a viable solution which has been shown to be more effective than synthetic auditory environments in recreating perception mechanisms such as echolocation [23].

2) Environment Sonification Implementation

The NAV-VIR system uses a combined approach to audio synthesis. Binaural recordings from real environments are combined with spatially filtered sound events, direction sonification and spoken instructions. The spatial filtering is performed by Resonance Audio, which is packaged as "Google VR Audio" in the Google VR SDK for Android.

Furthermore, simple sonification is used to signal the direction of the North, and of the target destination of the journey (or the nearest point along the path). The device's user interface will be vocalized using the default installed screen reader, such as TalkBack.

IV.   PRELIMINARY EVALUATIONS

A. Materials & Method

Preliminary evaluations were led to assess how simple shapes and directions were perceived when provided through the F2T. They involved a panel of 14 participants, 8 woman and 6 men, from both France and Poland. 7 were congenitally blind, 3 were late blind and 4 were blindfolded sighted. 8 were in the 20-40 age bracket, and the remaining 6 in the 40-60.

Those preliminary evaluations consisted of 2 main blocs. The first one aimed to assess if the participants were able to perceive the direction of simple movements generated by the F2T (active feedback). The second one aimed to assess the perception of simple geometrical shapes, such as square, circles and triangles, also provided through active feedback.

The experiments were organized as follow: (1) first, the experiment was explained to the participants which then proceeded to a guided "discovery" of the F2T, followed by (2) a description of the experimental conditions and what kind of questions would be asked of them. For the (3) first bloc of the experiments (perception of simple movements), the subjects were instructed to provide the perceived direction of the movement according to cardinal directions (e.g. North, South, South-East, …). For the (4) second bloc, the subjects were simply instructed to orally name the perceived shape. Each shape or movement was displayed a maximum of two times.

B. Preliminary results

Our participants were mostly able to correctly identify the direction of movements generated by the device ($\mu = 85{,}7\%$, $\sigma = 9\%$), and to recognize simple geometric shapes ($\mu = 94{,}6\%$, $\sigma = 6{,}8\%$) (fig. 6).
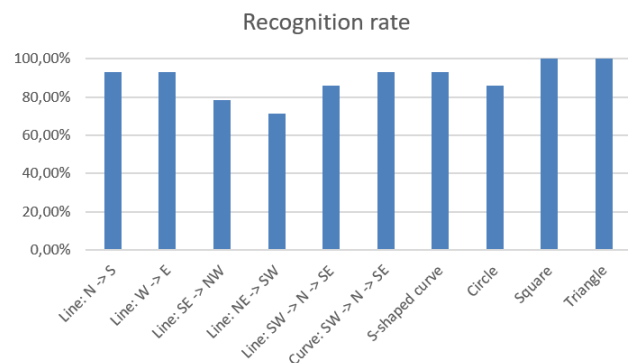

Figure 6.   Average recognition rate for each movement and shape

V. DISCUSSION AND FUTURE WORK

This paper introduced the NAV-VIR system, an audio-tactile virtual environment to assist VIP in exploring and planning a journey in an unknown area. This sensory supplementation device aims to provide a simplified topological representation of the environment that the user can interact with through a multimodal interface. NAV-VIR relies on the combination of a novel force-feedback tactile tablet used to provide both easily interpretable spatial information and spatialized audio cues to indicate relevant PoIs along a path. This two-modality representation of space enables the user to build a more complete mental map of its surroundings.

The purpose of this preliminary study was to assess the understandability of the spatial information provided through NAV-VIR. Our preliminary evaluation results seem to indicate that NAV-VIR is a suitable medium to generate interpretable dynamic tactile stimuli, allowing to reliably communicate movements, directions and shapes to the user. Indeed, all the subjects were able to recognize basic shapes and movements with this device, which will allow us to build and convey more complex representations in the future.

One limit of our current prototype is the slight delay in the joystick position estimation, resulting in its movements being quite jerky, making straight lines feel like steps or curves, thus making finer shape recognition more complex. This technological limitation has been solved with the second prototype currently developed, by using more accurate motors to reduce the positioning latency and a 3D printed casing for more stability. A more in-depth evaluation will later be conducted with this second prototype, to assess sound-augmented map comprehension and the transferability of the acquired spatial representation to real-world navigation tasks.

By adopting a multimodal approach and helping VIP make use of deficient sensorimotor loops, NAV-VIR should also help with the rehabilitation of their spatial cognition processes. Indeed, based on this early study, we hypothesize that this approach will result in a better comprehension of spatial relations by VIP. Finally, looking form a wider perspective, NAV-VIR could prove a valuable navigation-training tool leading to the improvement of the quality of life, safety, security and autonomy of VIP, as well as their integration in our society.

REFERENCES

[1] W. R. Wiener, R. L. Welsh, and B. B. Blasch, Eds., *Foundations of orientation and mobility*, 3rd ed. New York: AFB Press, 2010.

[2] R. L. Klatzky, J. M. Loomis, R. G. Golledge, J. G. Cicinelli, S. Doherty, and J. W. Pellegrino, "Acquisition of Route and Survey Knowledge in the Absence of Vision," *J. Mot. Behav.*, vol. 22, no. 1, pp. 19–43, Mar. 1990.

[3] A. M. Brock, P. Truillet, B. Oriola, D. Picard, and C. Jouffrais, "Interactivity improves usability of geographic maps for visually impaired people," *Human–Computer Interact.*, vol. 30, no. 2, pp. 156–194, 2015.

[4] S. Gay, M.-A. Rivière, and E. Pissaloux, "Towards Haptic Surface Devices with Force Feedback for Visually Impaired People," in *Computers Helping People with Special Needs*, vol. 10897, K. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 258–266.

[5] M.-A. Riviere, S. Gay, and E. Pissaloux, "TactiBelt: Integrating Spatial Cognition and Mobility Theories into the Design of a Novel Orientation and Mobility Assistive Device for the Blind," in *Computers Helping People with Special Needs*, vol. 10897, K.

[6] Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 110–113.

[6] E. Pissaloux and R. Velázquez, *Mobility of visually impaired people: fundamentals and ICT assistive technologies.* Springer International Publishing, 2018.

[7] J. Su, A. Rosenzweig, A. Goel, E. de Lara, and K. N. Truong, "Timbremap: enabling the visually-impaired to use maps on touch-enabled devices," in *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*, Lisbon, Portugal, 2010, p. 17.

[8] P. Skulimowski, P. Korbel, and P. Wawrzyniak, "POI Explorer – A Sonified Mobile Application Aiding the Visually Impaired in Urban Navigation," presented at the 2014 Federated Conference on Computer Science and Information Systems, 2014, pp. 969–976.

[9] T. Guerreiro, K. Montague, J. Guerreiro, R. Nunes, H. Nicolau, and D. J. V. Gonçalves, "Blind People Interacting with Large Touch Surfaces: Strategies for One-handed and Two-handed Exploration," in *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces - ITS '15*, Madeira, Portugal, 2015, pp. 25–34.

[10] K. Yatani, N. Banovic, and K. Truong, "SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback," in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, Austin, Texas, USA, 2012, p. 415.

[11] J. Guerreiro, D. Ahmetovic, K. M. Kitani, and C. Asakawa, "Virtual Navigation for Blind People: Building Sequential Representations of the Real-World," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '17*, Baltimore, Maryland, USA, 2017, pp. 280–289.

[12] K. Allain *et al.*, "An audio game for training navigation skills of blind children," in *Sonic Interactions for Virtual Environments (SIVE), 2015 IEEE 2nd VR Workshop on*, 2015, pp. 1–4.

[13] L. Zeng and G. Weber, "Exploration of Location-Aware You-Are-Here Maps on a Pin-Matrix Display," *IEEE Trans. Hum.-Mach. Syst.*, vol. 46, no. 1, pp. 88–100, Feb. 2016.

[14] C. Lenay, O. Gapenne, S. Hanneton, C. Marque, and C. Genouëlle, "Sensory substitution: Limits and perspectives," *Touching Knowing*, pp. 275–292, 2003.

[15] A. Kristjánsson *et al.*, "Designing sensory-substitution devices: Principles, pitfalls and potentials," *Restor. Neurol. Neurosci.*, vol. 34, no. 5, pp. 769–787, Sep. 2016.

[16] A. Isaiah, T. Vongpaisal, A. J. King, and D. E. H. Hartley, "Multisensory Training Improves Auditory Spatial Processing following Bilateral Cochlear Implantation," *J. Neurosci.*, vol. 34, no. 33, pp. 11119–11130, Aug. 2014.

[17] E. E. Pissaloux, R. Velazquez, and F. Maingreaud, "A New Framework for Cognitive Mobility of Visually Impaired Users in Using Tactile Device," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 6, pp. 1040–1051, Dec. 2017.

[18] J. M. Loomis, R. L. Klatzky, and N. A. Giudice, "Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language," in *Multisensory imagery*, Springer, 2013, pp. 131–155.

[19] R. F. Wang and E. S. Spelke, "Human spatial representation: Insights from animals," *Trends Cogn. Sci.*, vol. 6, no. 9, pp. 376–382, 2002.

[20] S. A. Lee and E. S. Spelke, "Two systems of spatial representation underlying navigation," *Exp. Brain Res.*, vol. 206, no. 2, pp. 179–188, Oct. 2010.

[21] L. Yang and M. Worboys, "Generation of navigation graphs for indoor space," *Int. J. Geogr. Inf. Sci.*, vol. 29, no. 10, pp. 1737–1756, Oct. 2015.

[22] M. Bujacz, P. Skulimowski, and P. Strumillo, "Naviton—A Prototype Mobility Aid for Auditory Presentation of Three-Dimensional Scenes to the Visually Impaired," *J Audio Eng Soc*, vol. 60, no. 9, pp. 696–708, 2012.

[23] M. Bujacz *et al.*, "EchoVis: Training Echolocation Using Binaural Recordings – Initial Benchmark Results," in *Computers Helping People with Special Needs*, vol. 10897, K. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 102–109.