AAPG2021	SAM-Guide		PRC
Coordinated by:	Christian GRAFF	48 Months	609k €
CES 33 - Interaction, Robotics			

SAM-guide:

Spatial Awareness from Multimodal guidance

Site	Partner	Name	Current position	Role & responsibilities in the project	PM	Total
Grenoble (UGA)	Laboratoire de Psychologie et Neuro-Cognition **(LPNC)** CNRS / Université Grenoble-Alpes	**GRAFF Christian**	**Maître de conférences** *HDR CE Neuroscience*	**Project Coordinator** Neuroscience of Perception; Psychophysics; Serious games; Experimental design	21	**67 PM**
		To be recruited	PhD Student (50% LPNC, 50% GIPSA)	*Human-Machine Interfaces; Psychophysics; Multimodal integration*	*36*	
			2 Interns	*Experimental evaluations; Psychophysics*	*10*	
	Grenoble Image Parole Signal Automatique **(GIPSA-Lab)** CNRS / Université Grenoble-Alpes	**HUET Sylvain**	**Maître de conférences** *Engineering Sciences*	**Scientific leader for GIPSA-Lab** HMI design; Augmented Reality; Motion capture	21	**41 PM**
		PELLERIN Denis	Professor *Engineering Sciences*	Auditory and Visual saliency; Computer Vision	4	
		JACCAZ Rémy	Assistant engineer	Prototype replication & maintenance	1	
		To be recruited	*3 Interns*	*Engineering; HMI; Augmented Reality*	*15*	
Normandy (NU)	Laboratoire d'Informatique du Traitement de l'Information et des Systèmes **(LITIS)** Normandy University (Rouen)	**PISSALOUX Edwige**	**Professor** *Engineering Sciences*	**Scientific leader for LITIS** Computer vision; Perception; ICT for mobility and health	15	**63 PM**
		To be recruited	*PhD Student*	*AR game development; Multimodal HMI; Localization and tracking*	*36*	
			Ingénieur d'Étude	*HMI design; Experimental design*	*12*	
	Centre d'étude sport et actions motrices **(CesamS)** Normandy University (Caen)	**FAUGLOIRE Elise**	**Maître de conférences** *Ergonomics*	**Scientific leader for CesamS** Perception-action coupling; Psychophysics; Activity and user-centered ergonomics; Experimental design & evaluations	18	**70 PM**
		MANTEL Bruno	Maître de conférences *Ergonomics*	HMI ergonomics; Affordance perception; Multisensory integration; Augmented Reality	18	
		To be recruited	*Post-Doc*	*Multimodal HMI; Spatial Cognition*	*24*	
			2 Interns	*Ergonomics & experimental evaluation of HMI*	*10*	
Palaiseau (X)	Centre de Mathématiques Appliquées **(CMAP)** CNRS / Ecole Polytechnique	**ALOUGES François**	**Professor** *Applied Mathematics*	**Scientific leader for CMAP** Audio processing; Mathematical sound modelling	9	**49 PM**
		FERRAND Sylvain	Research Engineer *Engineering Sciences*	Audio processing; Real time localization; HMI design; Electrical engineering	12	
		To be recruited	*Research Engineer*	*Human-Machine Interfaces; Audio processing; Real-time localization*	*18*	
			2 Interns	*Environmental sensing & floor texture detection*	*10*	
TOTAL						**290 PM**

This consortium comprises **five partners** split between **three sites**: Palaiseau with the CMAp lab (from Ecole Polytechnique), University Grenoble-Alpes with GIPSA-Lab & LPNC, and Normandy University with the LITIS and Cesams labs. The consortium wished to report that the project's estimated budget had to be increased by 13.9% since the pre-proposal, mostly to account for the recent 20% salary increase required for the two PhD students we plan to recruit, and the 4% increase in administrative management costs. There was also a change of Scientific Leader for the CMAp partner: Professor François Alouges will be the Scientific Leader of CMAP as Sylvain Ferrand has taken on new responsibilities in his laboratory. Finally, as per ANR recommendations, the number of PMs indicated for researchers with teaching duties is scaled to their entire activity. However, since they devote 50 % of their time to teaching, their monthly salaries have been divided by two in the financial data.

I. Proposal's context, positioning and objective(s):

a. Objectives and research hypotheses:

SAM-Guide's high level objective is to **efficiently assist Visually Impaired People** (VIP) in **tasks that require interactions with space.** To accomplish this, we will study and model how to **optimally supplement vision with both auditory and tactile feedback**. Spatial interactions encompass a wide range of tasks at different scales, which we subdivided in three subtypes based on VIP's needs: **navigating and orienting** oneself, **finding and reaching an object**, and being able to partake into **sport activities** such as **track-running or shooting**. For each of these subtasks, we will design, study and fine-tine **transcoding schemes,** which will require investigating both the best way to represent the interaction (providing the minimal yet sufficient information) and to convert that representation into auditory and tactile codes. Those transcoding schemes will be based on a **common "grammar"** and **common theoretical principles** to ensure their **interoperability**.

We will also establish **metrics and protocols** to properly assess the adequacy between the provided feedback and the type of spatial interaction investigated. Feedback will be provided through the partners existing interfaces, and the evaluations done through **Augmented Reality (AR) serious games** relying on **motion capture platforms** to track the user's movements. Finally, we will integrate all transcoding schemes into a **coherent holistic system**, and progressively **transpose it to real-world conditions** in order to ensure its **usefulness and acceptance by VIP**.

Our project will be guided by a set of **nested high-level principles**, based on current **neurocognitive models of human perception and spatial cognition** and **ergonomic considerations on HMI design**:

1. **Task-focused**: limit the scope of the information to provide by focusing on <u>supplementing spatial interaction</u> rather than trying to substitute vision in its entirety.
2. **Congruency**: communicate using a "code" <u>similar to known human sensorimotor laws</u>, thus <u>leveraging existing specialized brain networks</u> to process this information. This implies favoring <u>tight temporal coupling</u> between the user's <u>actions and their perceptual consequences</u>, which facilitates the <u>embodiment of the device</u> and <u>emergence of new sensorimotor control laws</u>.
3. **Parsimony**: focus on providing the <u>minimal yet necessary information</u> to enable autonomous interactions with space (i.e. finding an <u>efficient lower-dimensional representation</u>).
4. **Multimodality**: <u>leverage sensory associative areas</u> by combining multiple modalities through redundant or complementary signals, <u>facilitating their integration</u> at higher levels.
5. **Common transcoding "grammar"**: use common symbols to represent each cue we communicate through our interfaces. Here, we use <u>beacons</u>, 3D spheres characterized by an <u>ego-centered distance and orientation</u>, which can correspond to real elements of the environment (object to reach, target to aim for, landmarks) or abstract locations (path intersection, room center), be fixed or mobile, and have different scales of presence (local or global cue). This allows us to frame every type of spatial interaction as following or moving through a network of interconnected beacons.

We expect that following these general principles will make the SAM-Guide's system **more intuitive**, **salient** and **rich**, facilitate the **transition between scales and tasks**, require a **lower cognitive load**, minimize the **risk of interference** (with natural uses of a sensory modality), accelerate the **embodiment** of the device, improve its **usefulness** for VIP, and increase its **rehabilitation potential**.

Three types of tasks were chosen to best represent the various aspects of spatial cognition (and the impact of visual deficiency on it). Those tasks correspond to different difficulties encountered by VIP, as well as to **different scales of interaction with space** (personal, peri-personal, distant). For each of those subtasks, we will investigate **two interconnected levels of hypotheses**, guided by our general principles:

(1) What information to convey: amongst everything that vision captures, selecting the <u>relevant information</u> for each type of spatial interaction is crucial to keep the provided feedback parsimonious. The aim is to define an <u>optimal lower-dimensional representation</u> that preserves only the information necessary to accomplish that task. This includes considerations such as : can one represent a room by its center and encompassing radius instead of the distance to each wall without impacting one's ability to navigate in said room ? Can one understand the layout of a floor plan solely from perceiving a set of beacons representing points of interest (e.g. entrance, elevator, office) and corridor intersections ? Can one represent walkable space through local path affordances without impeding locomotion ? Should distance and angle to a target be provided in a single frame of reference, or should this frame change when one gets close enough to said target (e.g. peri-personal space) ?

(2) How to encode the elements of this representation, which includes several levels of considerations such as: how to best distribute the information between audition and touch, based on the inherent properties of these sensory channels (e.g. temporal and spatial discrmination), and on ergonomic and environmental constraints placed on these modalities (e.g. audition required to detect dangers while walking outdoors) ? Which information should be made multimodal, and should it be split or repeated between modalities, and in which circumstances ? Which information about the environment should be attributed to which parameter(s) of the output modality (e.g. distance to the frequency of successive vibrational pulses) ? Which functional form should be used to link the variations of the two (e.g. pulse frequency evolves as a continuous power law of the distance) ? Should higher-order information be used, such as using time-to-contact as a proxy variable for distance in order to have a progressive extinction of the signal when the user stops moving in relation to the target ?

Answering those questions involves exploring a **large hypothesis space**, and is heavily influenced by the task within which they are tested. To reach our goals in a timely manner, our consortium will rely on an innovative methodology: after filtering out the hypotheses that do not match our general guiding principles, we will explore the remaining ones using a **three step approach**:

(1) We will first evaluate our candidate hypotheses for each spatial interaction subtask in isolation, in a **controlled environment** using **motion capture** to precisely measure the participant's movements and **AR serious games** to create virtual worlds around them. This methodology makes it possible to **easily create varied testing environments and tasks**, and to **quickly prototype different types of sensors** (e.g. placing a virtual camera on the torso of the participant's game avatar). Furthermore, progress harmonization between sites becomes a simple matter of code sharing, using a common methodology and tools to ensure the compatibility of our results.

(2) Secondly, we will progressively **evaluate and improve the interoperability** of our best transcoding schemes in **more complex scenarios**, combining or chaining subtasks in more noisy environments (e.g. navigating to a target area and reaching a specific object in it, ignoring distractors).

(3) Finally, we will test our system **"out of the lab"** by extending it to **wider indoor and outdoor environments**, culminating with a **real-world testing scenario** encompassing each of the three spatial interaction subtasks: a **game of laser-run**.

Our methodology will adopt a **user-centered "living-lab" approach**, involving VIP and mobility instructors at every step, ensuring that our developments are in **adequation with their needs**.

SAM-Guide will provide the scientific community with a better understanding of the underlying principles of **spatial cognition**, and the role played by multisensory integration in the emergence of **spatial awareness**. Moreover, **guidelines** for the efficient representation of spatial interactions in HMI will be formulated (for VIP or sighted users), and illustrated by our own **modular sensor-agnostic and interface-agnostic transcoding librairies**. These could contribute to many fields or industries relying on Human-Machine interactions, such as cobotics, smart prosthesis & wheelchairs, or the development of more immersive games in AR or VR. Furthermore, our **prototype of wearable multimodal HMI** for visionless indoor and outdoor spatial interactions will be provided, which will include a prototype laser-run system for VIP. Finally, a set of **AR audio-tactile serious games** for VIP will be provided, which could be used either for novel **mobility training strategies,** or for recreational purposes. Overall, our primary hope is to see our efforts have a real impact on VIP's quality of life, by empowering them through more autonomy, access to new forms of leisure, and an easier participation in our society.

b. Position of the project as it relates to the state of the art:

Fifty years ago, **Sensory Substitution (SS)** emerged as a **new paradigm** promising fast and dynamic multi-purpose human-machine-environment interactions based on abstract low-level signals [1]. With enough training, those signals were supposed to be processed naturally and unconsciously by the user, leading to the embodiment of the device, and resulting in an intuitive perception akin to a new sensory experience [2]. This specific type of HMI relies on two phenomena: the brain's **cross-modal plasticity**, which occurs when loss of input in one sensory modality leads to reorganization in brain representation of other sensory modalities, and the **modality independent** organisation of the "sensory" cortices sub-networks, which reflects a **task-specific repartition** rather than sensory specific [3], [4]. Both phenomena have been observed in brain-imaging studies through the re-organisation of cortical networks in participants trained with SSDs [5]–[7], with congenitally blind participants being able to achieve a representation of space equivalent to that of sighted people after training [8].

One of the most well-known auditory SSD is *The vOICe* [9], which converts the **visual field's brightness** into an array of sounds of varying pitch and loudness. Long-term users of this device have shown the emergence of phenomenal consciousness of spatial awareness [10]. It can use both a computer or smartphone's camera, and has been extensively used in research on multisensory integration. Since then, many more devices have been developed (see [11] for a review), often relying on smartphone sensors to assist navigation [12], [13], map exploration [14], [15], or target-reaching [16].

UGA's SSD, "AdViS" is based on principles similar to The vOICe, relying on a depth camera to provide distance and orientation information of relevant spatial cues, transcoded through stereophony, pitch and loudness, with instantaneous feedback [17]. Their more recent work has focused on target-reaching tasks, exploring the interplay of **frames of references** [18], using not only **visual inputs** but also **body segment proprioception** to converge towards a more efficient reduction of finger-to-target distance [19], [20]. Indeed, the frame-of-reference choice in SSD design will either favor access to a **vision-like** representation (in the scope of exploiting "visual" processing brain areas), or to some particular **task-oriented constructs** based on hearing and touch. Most SSDs, both audio and tactile, consider either **absolute** XYZ coordinates with an arbitrary Cartesian origin [21], or remain attached to the **sensor providing the** spatial information, i.e. the camera (e.g. The vOICe).

Taking advantage of AR environments and motion capture systems, they have studied novel **embodied representations** using **natural pointers** (such as the fingertip) and references frames (such as the torso) instead of camera-centric ones, as well as the efficiency of deviation-minimization feedback for target reaching [20]. They also studied the possibility of using eye-movements to control a foveal area of higher density of information collection for late blind individuals, mimicking the natural separation of the visual processing stream during scene understanding [22]. They are currently investigating different strategies to map spatial parameters onto the **functional wealth** of auditory parameters [23], such as dissociating the pitch and yaw when projecting the angle between a pointer and a target, as well as the proper articulation of transitions between approach and grasping phases during target-reaching.

CMAp's focus has been on enabling **autonomous practice of sports for VIP**, using **spatialized audio SSDs**. Despite existing commercial solutions relying on **smartphone** sensors to assist VIP in those activities, such as the Navi-rando [24] app which guides people on hiking trails by combining GPS and IMU data for reliable localisation, many VIP rely are constrained to participate in collective outdoors activities (trekking, skiing, cycling, ...) where they follow a **sighted guide** providing voice indications, or simply relying on the sound of the guide's movements to follow the track. Mimicking this behavior, CMAp proposed a **virtual audio guide** in the form of a **3D "fairy"** which moves along the user, constantly staying in front, and emitting sound cues [25]. This mode of guidance allows the VIP to gradually adjust their orientation by running towards the perceived position of this guide, and relies on the fact that, just like the visual system, the auditory system actively tracks the movements and velocity of a sound source [26]. This approach requires both the **fast and accurate real-time** localization of the user (in order to update the simulated position of the guide, and the **faithful representation** of this position to the user. This is ensured through the use of binaural sounds with **HRTF** (head-related transfer function) sound convolution, in which sound parameters are mathematically defined according to the hearer's **anatomy and position** relative to the sound **source** [27]. HRTF enables the **precise and externalized perception** of the **3D position of a sound source**, enabling faster reactions from the user. Their current works focus on improving their localization algorithms, exploring algorithms to modulate the signal along cumulative errors in order to compensate for natural **veering** [28], as well as the accuracy of their 3D sound modelling [29].

A more recent focus of CMAp has been facilitating paralympic games for VIP, which comprises activities such as **blind shooting** with sound-emitting targets [30]. The laser gun appears as a novel exciting **tool** to study the transcoding of remote spatial parameters in a playful context, and would make an excellent support to study and develop the interoperability of SSDs designed for different scales and tasks (navigation, object-reaching and shooting). CMAP has developed an initial prototype of audio shooting device for VIP, which has been demonstrated to the Féderation Francaise de Pentathlon Moderne (FFPM), who expressed their interest in participating in further developments.

NU's focus has been on the development of tactile SSDs for **autonomous orientation and navigation**, **map comprehension**, **access to art and Web content,** as well as the **optimization of the tactile encoding** of remote cues (using motion capture systems and AR). Current vibrotactile system developments have long abandoned the idea of transcoding the whole landscape on a cumbersome matrix fixed on the subject's back, but still struggle to match the wide-spread adoption of minimalist devices such as the electronic white cane [31]. These wearable devices have been adapted to many different types of interfaces, including glasses [32], portable taxel-surface displays [33], bracelets [34], gloves [35], soles [36], or directly exploiting smartphones vibrotactile feedback [37]. The most studied, and probably the most efficient for navigational tasks, are **vibrotactile belts** [38], [39]. They have been used to dynamically indicate a reference **direction** such as the magnetic north [40] or a functional guiding direction [41], [42]. Vibrotactile belts combine the advantages of **ego-centered directional** information with the innate spatialization and externalization of touch, and have been shown to be more efficient and intuitive than traditional speech-based instructions for conveying locomotion instructions [41]. Their research also investigated **ecological modes of responses**, and **whole-body movements affordance-based HMI design** [43], [44].

Recently, NU explored the **efficiency of different combinations of tactile stimuli** (intensity, duration, number pulses, actuator position on the body, ...) in dynamic pattern recognitions tasks [45], and is currently comparing the relevance of different modes of transcoding for distance and orientation information when navigating between beacons, and how to best combine tactile feedback with spatialized sound guidance for situations requiring more precise movements. Another current research focus of theirs is **autonomous orientation and navigation** [46], and **non-visual map comprehension**, both to formulate models and guidelines on VIP's mobility assistance [47]–[49], as well as to design novel assistive devices such as an interactive and dynamic 2D interface for graphical content exploration, based on force-feedback [50]–[52].

Recent **advancements in ICT** (Information and Communication Technologies) combined with the democratization of wearable computing and IoT (Internet of Things) devices have further increased the scope of possibilities for VIP assistive devices (c.f. [53] for a review). Those assistive systems rely on the fusion of multiple embedded sensors to extract detailed and precisely located information from the environment, often relying on computer vision technology [54]. However, despite the theoretical possibilities of SS and the ubiquity of technical solutions to implement them, very few SSDs are actually in use by VIP [55], [56]. As of today, the most popular assistive device remains the simple electronic white cane, which warns of the presence of remote obstacles in front of the user [31], [57], [58], or specific smartphone applications such as *Microsoft's SeeingAI* (which can read and vocalize grocery tags and other natural text), and unobtrusive accessories dedicated to simple and **specific tasks** such as revealing clothing colors. This lack of adoption has been linked to many **pitfalls** of SSD design trends [49], [59], [60], such as their **interference with other activities** or other sensory channels (e.g. relying on audition to detect dangers outdoors) [56], being too **cumbersome** or too **conspicuous**, requiring **too**

much training to be practical or trying to **convey too much information** with the misguided aim of replacing the lost sensory modality [61].

Following those observations, SAM-Guide's philosophy is not to substitute vision as a whole, but rather to focus on one type of tasks, **spatial interactions**, which are a key difficulty often expressed by VIP, even after extensive Orientation & Mobility training. Spatial interactions involve numerous interrelated physical and cognitive processes, and rely on the integration of multisensory egocentric and allocentric information over time to build and update a holistic mental representation of the environment. [62]. We subdivided those processes in three categories corresponding to different scales of interactions (personal, peripersonal, distant) and corresponding to different needs expressed by VIP: **navigating and orienting** oneself, **finding and reaching an object**, and being able to partake into **sport activities** such as **track-running or shooting**. Multiple SSDs have been developed for each of these tasks [53], but often do so using different frameworks and principles, resulting in several disparate systems relying on different interfaces and transcoding schemes. This limited scope and lack of interoperability has been a limiting factor in their adoption, requiring a more complex learning process and complicating the transition between tasks [60].

SAM-Guide's originality will come from a set of **theoretical principles** meant to ensure that our developed transcoding schemes are intuitive, efficient, and match VIP's actual needs and habits. Firstly, we will devise a **common syntax** across tasks and scales, reframing spatial interactions as **target-reaching affordances,** and symbolizing spatial properties by **3D ego-centered beacons**. Our system will thus function using **scale and task-agnostic control laws** that recode the environment through simple orientation and distance cues, which were shown to be sufficient to **elicit spatial awareness** [63]. Representing spatial interactions using this **common "grammar" and symbolism** will unify the multiple domains of **spatial cognition** and ease the **transition** from one task's space to the next, accelerating the learning process and limiting the number of devices to wear and master.

Secondly, we plan to **combine both auditory and tactile feedback**, taking advantage of the strengths and weaknesses of both modalities, as well as the unique advantages that come from their interactions. Indeed, multimodal interfaces have been shown to improve the richness of the feedback, **improve rehabilitation** [64], makes the provided information **more salient** and **less prone to interference** [65], and reduce the cognitive load and training time [60]. Combining multiple modalities through **redundant** or **complementary** signals will leverage **sensory associative areas** and facilitate their **integration at higher levels** [8], resulting in better performances and a more intuitive use [66].

Thirdly, we will focus on providing **parsimonious feedback**, relying on neurocognitive models of … to define the minimal yet necessary information to enable autonomous interactions with space (i.e. finding an efficient lower-dimensional representation). When it comes to spatial representations, many studies on **spatial cognition,** which models how the brain extracts higher-level descriptors to summarize spatial information efficiently and to build cognitive maps of the surroundings [67], have yielded clues on "invariants" of spatial processing in mammals, both in functions [68], [69] and underlying brain structures [70]. For example, recent studies in neuroscience have shown that rodents keep a constant "pointer" towards the center of the room they are currently in to facilitate spatial awareness [71], or that our brains extracts the "skeletal structure" of objects [72] and of paths in a scene [73] in order to recognize and manipulate them. Based on these models, we will devise a **non-visual spatial presentation** which will promote **spatial awareness** and allow for **efficient mobility** whilst only requiring a low cognitive load.

Finally, we will ensure that our transcoding schemes are **congruent** with existing sensorimotor laws, thus leveraging existing specialized brain networks to process this information. A major aspect of this approach is the necessity of **action-perception coupling**, a circular causality between the device stimulation and the user's action when updated at sufficient rate, to elicit **embodied perception** [74]–[77] and facilitate the acquisition of new **sensorimotor control laws** [40], [43], [44]. Another aspect to consider is the use of **ego-centered feedback** with a **drift-correction approach**, which references the **ego-centered deviations** between the user and its target instead of the absolute position of both, and corresponds to a more natural scheme of perception [78]. Finally, the classical work of Weber's and Steven on power laws in psychophysics has shown that human perception naturally maps perceptual changes to the logarithm of environmental parameters.

Another innovative aspect of SAM-Guide is the methodology it will adopt for the technical development and evaluation of its transcoding schemes. SSD development is traditionally a slow endeavor due to (1) the necessity to develop a **complete prototype** to enable experimental evaluations, (2) being limited in the information you can provide to the user by the sensors and processing algorithms included in this prototype, (3) having to design and randomize complex evaluation environments (e.g. a maze of cardboards), or being limited to **simple and often not ecologically plausible** that limit the validity and generalisability of the results obtained. Some of these problems have been solved by using virtual avatars that users can move in 3D environments [79], which allows the **virtual prototyping** of different SSD sensors, drastically reducing development costs and time. However, without physical movements within these environments, the users lose access to proprioceptive feedback [80], which limits the conclusions drawn from such studies. Combining virtual environments with motion capture systems preserves most of the natural action-perception loops, while still. This method has recently shown an increase in adoption [81], [82] and both NU [41] and UGA [20], [83] have fully adopted this methodology in their current endeavors.

Our assessments will rely on a combination of metrics such as **psychometric methods** (e.g. **psychophysical curves)** to evaluate the amount of information that can be conveyed, **behavioral measures** (e.g. time-to-event or mental rotations tasks) to quantify the **efficiency** of the transcoding schemes employed for a given task, **ergonomic analyses** to assess the adequacy between the task and the exploration strategies allowed by the information provided, and qualitative **user-feedback** from VIP and mobility instructors. Finally, we will ensure the adequacy of our developed transcoding schemes to **real-world conditions** by expanding them to **more complex scenarios**, combining or chaining tasks in more ecological environments where participants have to maintain concurrent objectives relating to tasks at different scales, whilst evolving in a noisy environment.

c. Methodology and risk management

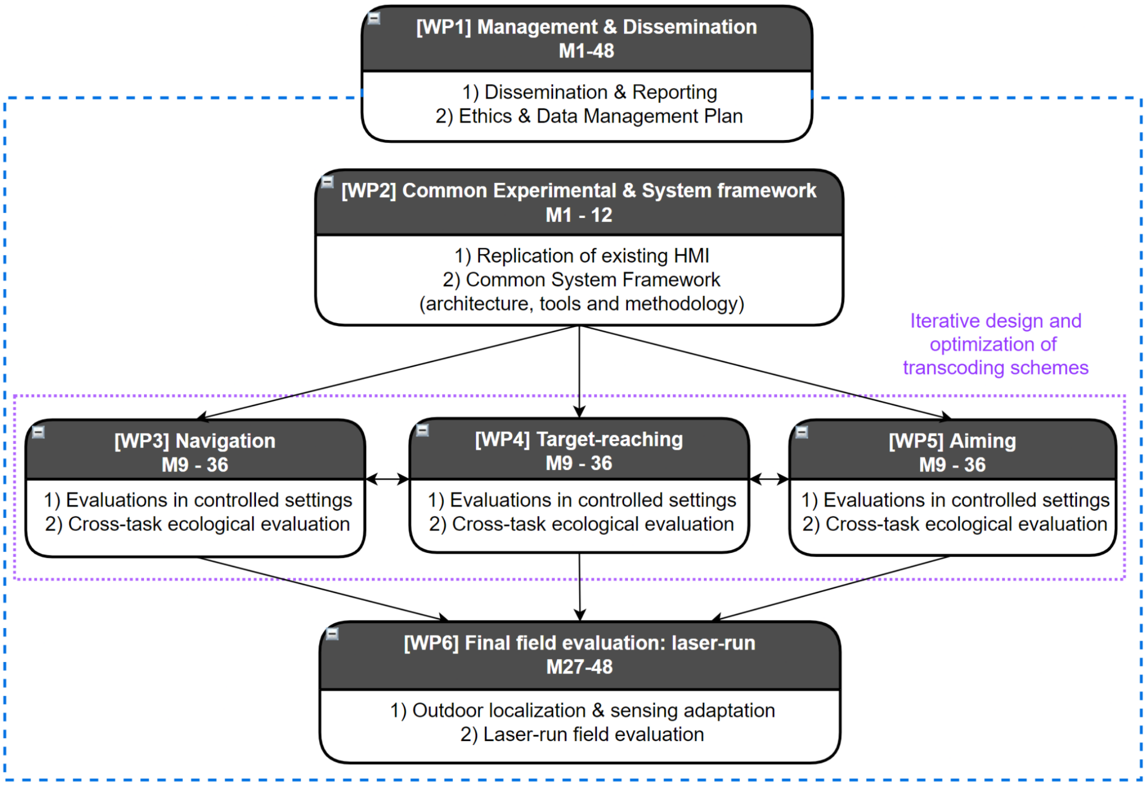
SAM-Guide aims to develop **efficient transcoding schemes for spatial interactions**, based on biologically-derived ergonomic principles and a common "grammar". Those transcoding schemes will take the form of **modular libraries** designed to adapt to a wide range of interfaces (vibrotactile or auditive) and sensors. They will be **evaluated within specifically designed testing environments**, both in <u>controlled</u> (serious AR games with motion capture) and <u>real-world</u> (indoor and outdoor) settings.
To achieve these goals, we divided SAM-Guide in **six interconnected WP** (leading partner highlighted):

WP	Name	Duration	PM LPNC	PM GIPSA	PM LITIS	PM CesamS	PM CMAP	PM Total
1	Management & Dissemination	M1 - 48	**7**	2	6	4	2	**21**
2	Common Experimental & System framework	M1 - 12	2	5	**7**	5	3	**22**
3	Navigation	M9 - 36	8	4	20	**21**	8	**61**
4	Target Reaching	M9 - 36	18	**21**	8	10	3	**60**
5	Aiming	M9 - 36	8	4	8	16	**21**	**57**
6	Final field evaluation: laser-run	M27 - 48	**24**	5	14	14	12	**69**

In a nutshell, [WP1] aims to set up an <u>efficient collaboration</u> within the consortium and ensure the proper <u>dissemination and reporting</u> of the projects' progress. [WP2] ensures each partner has access to the others' existing prototypes, and specifies the <u>initial overall system's architecture</u> **{Milestone 1}**, which is based on a modular core application and a common set of technologies and experimental protocols, to ensure the interconnectedness of the following WPs. [WP3-5] are the <u>main experimental WPs</u>, each focusing on one type of spatial interaction task (long-range navigation, close-range target-reaching, and medium-range aiming, respectively). Their primary goal is to <u>design and iteratively</u>

improve multimodal transcoding schemes for their respective task, and to develop ecological scenarios and measures to evaluate the transcodings efficiency, using the partner's interfaces and motion capture platforms to develop AR serious games. Their secondary goal is to progressively interconnect the three tasks and their respective transcoding schemes into a holistic transcoding scheme for visionless spatial interactions **{Milestone 2}**. [WP6] will adapt this global system to the real-world by using various sensors and algorithms to localize the user and capture the relevant information within the environment. The resulting system will then be iteratively evaluated and improved through an outdoor game of laser-run, which combines all three task types and their transcoding schemes, leading to the final specifications of our system **{Milestone 3}**.

The following figures and tables summarize this organization:



		Semester I		Semester II		Semester III		Semester IV		Semester V		Semester VI		Semester VII		Semester VIII	
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q14	Q15
Management & Dissemination	**WP1**																D1.5
Reporting & Dissemination	T1.1				D1.1				D1.2				D1.3				
Ethics & DMP	T1.2		D1.4				D1.4										D1.4
Common Experimental & System framework	**WP2**				D2.2												
Replication of existing HMI	T2.1				M1												
Common system framework (architecture, tools, methodology)	T2.2																
Navigation	**WP3**												D3.3				
Design and evaluation in controlled setting (protocols & games)	T3.1										D3.1						
Cross-tasks ecological evaluations	T3.2												D3.2				
Target-reaching	**WP4**												D4.3				
Design and evaluation in controlled setting (protocols & games)	T4.1										D4.1						
Cross-tasks ecological evaluations	T4.2												D4.2				
Aiming	**WP5**												D5.3				
Design and evaluation in controlled setting (protocols & games)	T5.1										D5.1		M2				
Cross-tasks ecological evaluations	T5.2												D5.2				
Final field evaluations: laser-run	**WP6**																D6.3
Outdoor adaptation (localization & environment sensing)	T6.1														D6.1		M3
Laser-run evaluations	T6.2																D6.2

The following tables will provide the details of each WP, including their breakdown in several **tasks {T}**, **deliverables {D}** and the **milestones {Mi}**. Deliverable confidentiality is coded as: **Confidential** (**Co**) for documents only accessible to the consortium and the ANR evaluation committee, **Semi-Confidential** (**SCo**) which also grants access to external partners such as VIP associations, and **Public** (**Pu**).

AAPG2021	SAM-Guide		PRC
Coordinated by:	Christian GRAFF	48 Months	609k €
CES 33 - Interaction, Robotics			

WP1	Management & Dissemination					M1 - 48
Partners [#PM]	**LPNC [7]**	GIPSA [2]	LITIS [6]	CesamS [4]	CMAP [2]	**21 PM**

Objectives: efficient remote collaboration, organizing kick-off and regular consortium meetings, ensuring the proper dissemination and visibility of the project, recruiting the non-permanents and planning short-term inter-sites exchanges for them.

{T1.1} (M1-M48) **Reporting & Dissemination:** yearly reports of the project's overall progress and enactment of the proposed dissemination plan **[LPNC, LITIS, CesamS]** (c.f. section III for more details).

- **{D1.1, D1.2, D1.3}** (M12, M24, M36 - Reports - Pu): Annual reports.
- **{D1.5}** (M48 - Report - Pu): Final report.

{T1.2} (M1-M48) **Ethics & DMP:** set up a Data Management Plan (DMP) compliant with the RGPD and Loi Jardé. We will rely on structures such as GRICAD (UGA) which freely assists research projects in the management, design and engineering of their data **[LPNC, GIPSA]**. We will also ensure our experiments comply with ethical guidelines, ethical concerns being of high importance to our consortium, the project coordinator being a founding member of the first ethical committee for behavioral experiments on humans in France (at Grenoble University) **[LPNC, LITIS, CesamS]**.

- **{D1.4}** (M9, M24, M48 - Reports - Pu): DMP initial strategy and follow up reports.

WP2	Common Experimental & System framework					M1 - 12
Partners [#PM]	LPNC [2]	GIPSA [5]	**LITIS [7]**	CesamS [5]	CMAP [3]	**22 PM**

Objectives: replicate each partners' existing interfaces and define the initial specifications of SAM-Guide's full system (which includes a common set of technologies, practices and experimental methodology). This common framework will ensure the intercompatibility of the code produced by each subsequent WP (for either the transcoding libraries or the serious games used to evaluate them), and the comparability of the experimental results obtained, leading to better modeling power and generalizability.

{T2.1} (M1-M6) **Replication of the partners' current interfaces**, allowing each partner to experiment with and combine these various interfaces during the subsequent WPs **[GIPSA]**. Those interfaces are already designed and documented, and will not be modified as they are versatile enough to accommodate the variety of transcoding schemes we will be evaluating.

{T2.2} (M1-M12) **Common system framework:** setting up the core of the system (which will incorporate the partner's current transcoding code as independent modules), and defining the standards for future software development, emphasizing modularity and generalizability (i.e. sensor and interface agnostic transcoding libraries) **[LITIS, GISPA, CMAP]**.

We will also define a common experimental methodology to guide the exploration of the myriad of possible encoding schemes that fit SAM-Guide's principles, and to guide the development of a shared set of AR serious games to evaluate those schemes. This methodology will revolve around an inclusive user-centered "living-lab" approach, involving both VIP and mobility instructors at every step of the evaluations. Sighted participants will also be solicited for most of the pre-testing, but a special emphasis will be given to the results and feedback provided by VIP and their caretakers. We will adapt our protocols to the constraints of the current pandemic, relying on the experience each partner has accumulated during the last year (e.g. one participant at a time, experiment piloted remotely, proper sanitization of the equipment, and relying on online virtual testing when possible). Finally, our experimental methodology will adhere to the open-science philosophy to best ensure the replicability of our results (i.e. pre-registration, open-sourcing the anonymized data and the statistical analysis scripts) **[LITIS, CesamS, LPNC]**.

{Mi1} The end of this WP will mark the **first milestone** of SAM-Guide, with the initial specifications of our complete system and our experimental methodology; which will be detailed in **{D2.1}** (M9 - Report - Pu).

[WP3-5] are focused on the **iterative evaluation and optimization of our transcoding schemes** through three ecologically-relevant types of spatial interaction tasks, involving target-reaching at various scales: navigating towards a distant target (while keeping track of the environment's layout) [WP3], reaching an object in a room [WP4], and aiming towards a remote target [WP5]. Each WP will focus on one type of spatial interaction task (based on prior experience & research interests) while relying on a common system architecture and methodology (cf. [WP2]). Feedback will be provided using the partners' existing HMI, which include audio-substitution headphones were spatial parameters are encoded by specific audio qualities (e.g. deviations in elevation linked to pitch) [17], vibrotactile belts using multiple layers of vibrators spread around the waist [41], [84], and 3D binaural audio headphones using HRTF [85].

Experimental evaluations will be done in **two phases** in order to progressively improve our transcoding schemes and to assess their relevance to more ecologically relevant environments:

(1) The first phase will rely on specifically designed **Augmented Reality (AR) serious games**, using the partners' existing **motion capture** systems to link the users real-life movements (within the motion capture environment) to their game avatar movements. Those serious games will act as playful mediums to evaluate our transcoding schemes, and improve the participants' commitment.

This methodology allows participants to **move and interact with their environment ecologically**, while allowing researchers to gather precise behavioral data in tightly controlled and customizable settings [41], [83]. The targets, landmarks, and obstacles might be virtual, but the user's movements (and their perceptual consequences - through the interfaces) will be ecological. With this approach, most of the necessary **prototyping can be done through software development** by virtually emulating various sensors. This will avoid time-consuming hardware development and facilitate replication among sites through the centralization of the developed AR games and transcoding libraries.

Furthermore, to address the issue of participant availability due to COVID, some **preliminary testing will partly be done online** using a simplified version of our serious games. Those tests will rely on the mouse & keyboard for movements, and the users' headphones for feedback (and, when available, the users' VR headsets for more ecologically-grounded movements & feedback based on head rotation). Recent studies indicate that this should only have a minor impact on our ability to study the participants' ability to integrate the information provided, despite the lack of actual movements [4]. NU and UGA are currently collaborating on this approach, using the Unity game development framework.

(2) In the second phase, our transcoding schemes will be **evaluated in more ecological environments** (i.e. more complex, less controlled and larger-scale), in order to assess their relevance and generalizability outside of the tightly controlled motion capture environment and tasks they were developed in. Those evaluations will rely on less accurate but wider-scale localization tools, such as the Pozyx system (indoor), or a differential GPS (outdoor). This second phase will ensure a smoother transition to [WP6], which will combine the developments of [WP3-5] in a playful outdoor sport activity encompassing all 3 types of spatial interactions: laser-run.

Our evaluations will rely on behavioral performance metrics (e.g. deviation from an optimal trajectory, time-to-event, ...) which will be provided either by the motion capture or localization system, combined with ergonomics assessments (e.g. cognitive load, task analysis) and qualitative user-feedback from VIP.

WP3	Navigation & Spatial Cognition					M9 - 36
Partners [#PM]	LPNC [8]	GIPSA [4]	LITIS [20]	**CesamS [21]**	CMAP [8]	**61 PM**

Objectives: design and iterative improvement of multimodal transcoding schemes to facilitate tasks of navigation and large-scale orientation in the absence of visual information. This WP will investigate both the **optimal way to represent navigation and orientation tasks,** intuitively through vibro-tactile and auditory feedback. Based on **neurocognitive models of mammal's navigation**, the representation will symbolize important locations of the environment (landmarks, pathway intersections, center of the room) using **beacons**. We will explore different combinations of **vibrotactile feedback** for long-range environmental cues (such as landmarks) and **3D audio feedback** (in the form of a **moving "fairy" guide** moving in front of the user [85]) for more complex navigation tasks (i.e. going around a corner, or through a door).

Hypotheses: more efficient navigation using a two-point path integration strategy; ability to understand and use topological information;

{T3.1} (M9-M30) **Design and evaluation of transcoding schemes for non-visual spatial cognition in small-scale motion capture environments**: designing the **metrics**, **tasks**, and the **serious games** encapsulating them. This first phase will evaluate the participants on simple and small-scale path-following and maze-solving tasks. It'll use the motion capture system to get fine-grained movement measures in order to test the various potential task representation and transcoding schemes accurately. The metrics and tasks developed will evaluate both the **ability to efficiently follow a path** (e.g. time-to-target, deviation from the optimal path, estimating the position of a vanishing target, ...) and to **understand the layout of the environment** while doing so (e.g. maze solving, boundary rotation, relative direction judgements, ...) **[CesamS, LITIS, LPNC, GIPSA]**.

- **{D3.1}** (M30 - Report - Pu): Report on spatial cognition evaluation protocols, tasks and games.

{T3.2} (M21-M36) **Experimental evaluation of transcoding schemes for non-visual spatial cognition in larger-scale indoor and outdoor environments**. This second phase of the evaluations will rely on the metrics and tasks developed during the previous task, but **applied in a wider and more ecological setting**. The participants will be evaluated in **more complex and larger scale navigation tasks**, both **indoor** (with actual corridors, intersections, obstacles and outside interferences) using the Pozyx large-scale indoor localisation system, and in a **controlled outdoor setting** using a differential GPS combined with CMAP localization algorithms. **[LITIS, CesamS, CMAP]**.

- **{D3.2}** (M36 - Report - SCo): Report on both phases' experimental results (data, analysis and conclusions).
- **{D3.3}** (M36 - Code - Pu): Proposed transcoding library and AR games for orientation & navigation tasks.

WP4	Target-reaching					M9 - 36
Partners [#PM]	LPNC [18]	GIPSA [21]	LITIS [8]	CesamS [10]	CMAP [3]	60 PM

Objectives: **design and iterative improvement of multimodal transcoding schemes to facilitate tasks of target-reaching** in the absence of visual information. This WP will investigate both the **optimal way to represent target reaching tasks**, and how to most efficiently communicate this representation through vibro-tactile and auditory feedback. The representation of the target reaching task including both **approaching and grasping a fixed object** in 3D space) will be based on classic neurocognitive theories linking frames-of-reference with ongoing actions to define **body-referenced geometrics** [86]. Drifts from the target in distance and angle are defined within with a bodily reference origin on the user (e.g. fingertip, torso, ...) instead of an allocentric optical sensor's fovea. Such **embodied frames-of-reference** will dissociate verticality (gravity, relative elevation), horizontality (panning, in angle) and reachability (distance from finger). Sonic continua will be chosen as most intuitive (e.g. stereophony for horizontality, intensity for distance).

Hypotheses: higher speed, faster learning, more comfort with bodily auto-referenced drift transcoding than allocentric.

{T4.1} (M9-M30) **Design and evaluation of transcoding schemes for non-visual target-reaching in small-scale motion capture environments**. This WP will focus on interactions in the peripersonal space (simple small scale tasks in controlled settings) in order to develop **metrics**, **tasks** (and the **serious games** encapsulating them) relevant to both subtasks of the target-reaching process. Both subtasks will focus on different types of feedback: reaching the target itself (reaching the peripersonal space of the target) is a medium-scale navigation task requiring low accuracy spatial information, making it suited for vibrotactile signals. However, grasping movements require precise 3D gestures and extremely fast action-perception loops for constant approach corrections, making substituted audio feedback more suited for it. Grasping transcoding schemes will thus focus on fast and informative audio signals aiming to minimize the deviations between the user's frame of reference and the target **[GIPSA, LPNC, CesamS]**.

- **{D4.1}** (M30 - Report - Pu): Report on target-reaching evaluation protocols, tasks and games.

{T4.2} (M21-M36) **Experimental evaluation of transcoding schemes for non-visual target-reaching in larger-scale indoor and outdoor environments**. This second phase of the evaluations will rely on the metrics and tasks developed during the previous task, but **applied in a wider and more ecological setting**. The participants will be evaluated in **more complex and larger-scale target reaching tasks** (e.g. going towards a room or specific location, locating a specific object in the area, going towards and grabbing it), both indoors (with the Pozyx system) and outdoors (using a differential GPS with CMAP localization algorithms) **[GIPSA, LPNC, LITIS, CesamS, CMAP]**.

- **{D4.2}** (M36 - Report - SCo): Report on both phases' experimental results (data, analysis and conclusions).
- **{D4.3}** (M36 - Code - Pu): Proposed transcoding library and AR games for target-reaching tasks.

WP5	Aiming					M9 - 36
Partners [#PM]	LPNC [8]	GIPSA [4]	LITIS [8]	CesamS [16]	CMAP [21]	57 PM

Objectives: **design multimodal transcoding schemes to facilitate ranged aiming tasks** in the absence of visual information. This WP will investigate both the **optimal way to represent aiming tasks,** and how to best convey this representation intuitively through vibro-tactile and auditory feedback.

Hypotheses: faster and more precise shooting with sonic transcodings inspired from target-reaching H/V dissociation (WP4), and with combined tactile and sonic feedback;

{T5.1} (M9-M30) **Design and evaluation of transcoding schemes for visionless shooting in small-scale motion capture environments**: designing the **metrics**, **tasks**, and the **serious games** encapsulating them. This first phase will evaluate the participants on simple and small-scale aiming tasks. It will use motion capture to locate the target in relation to the user, and the laser-gun software itself to estimate where the user is aiming in relation to the bullseye. The provided representation will provide both of those information through both vibro-tactile (for the location of the target) and auditory (for the aiming error) feedback. Since aiming requires **extremely fast and precise error-correction** (in order to be competitive), those transcoding schemes will focus on fast **error minimization on the 2D plane of the target**. This WP will be done in interaction with the Paralympic societies to ensure the compatibility of our technical improvements with the emerging rules of this novel sport **[CMAP, LPNC, CesamS]**.

- **{D5.1}** (M30 - Report - Pu): Report on aiming evaluation protocols, tasks and games.

{T5.2} (M21-M36) **Experimental evaluation of transcoding schemes for non-visual aiming in larger-scale indoor and outdoor environments**. This second phase of the evaluations will rely on the metrics and tasks developed during the previous task, but **applied in a wider and more ecological setting**. The participants will be evaluated in **more complex and larger-scale aiming tasks**, which will include several sub-tasks relevant to the laser-run game (i.e. reaching the firing stand and finding the gun itself), combining the pure aiming task with target reaching and navigation ones **[CMAP, GIPSA, LITIS, CesamS]**.

- **{D5.2}** (M36 - Report - SCo): Report on both phases' experimental results (data, analysis and conclusions).
- **{D5.3}** (M36 - Code - Pu): Proposed transcoding library and AR games for aiming tasks.

The end of [WP3-5] will mark the **second milestone** of the project, where our transcoding libraries will be finalised after being tested both in isolation (in controlled motion capture settings), and in conjunction (in wider ecological indoor and outdoor settings), resulting in a **scale-agnostic beacon-based spatial interaction paradigm** **{Mi2}**. [WP6] will consist of an outdoor sports activity encompassing the three types of tasks evaluated in [WP3-5]: **laser-run**. This activity combines multiple trials of running to a shooting stand, grabbing a laser-gun, and firing at a remote target. All partners will be regrouped on a single WP: some pre-testing will be done on each site, and the laser-run evaluations on the campus of Ecole Polytechnique.

WP6	Final field evaluation: laser-run				M27 - 48
Partners [#PM]	**LPNC [24]**	GIPSA [5]	LITIS [14]	CesamS [14]	CMAP [12]	**69 PM**

Objectives: **transposing SAM-Guide's final system to real-world sensing and localization**, and **evaluating it through a laser-run activity**, which **combines navigation** (while running, to get to the firing range), **target-reaching** (to find the gun and the firing stand), and **aiming** (to hit the target).

Hypotheses: ability to execute multiple tasks simultaneously; continuity of representations during successive tasks at different scales; VIPs' increased engagement due to the playful nature of the evaluations;

{T6.1} (M27-M42) **Adapting the final system to outdoor environment sensing & localization**. The localization will be done through a combination of wearable (e.g. IMU & cameras) or ambient (e.g. differential GPS and beacons) sensors, based on past and current work by the consortium partners' on real-time localization [17], [25], [33], [85]. Target reaching (to find the gun and the shooting target) will rely on a combination of beacons and vision-based object recognition. The information gathered by those sensors will be conveyed using the existing interfaces in order to couple the system's feedback to the user's position in the environment, and will rely on the transcoding schemes developed in [WP3-5]. The system will ensure the participants' security during evaluations, with protocols alerting the user in case of danger (e.g. collision warnings) **[CMAP, LITIS, GIPSA]**.

- **{D6.1}** (M36 - Report - Pu): Specifications of the "outdoor-ready" version of SAM-Guide.

{T6.2} (M36-M48) **Laser-run field evaluations**: those evaluations will be mainly conducted on campus of Ecole Polytechnique due to their existing contacts with HandiSport associations and with VIP used to practicing laser-run. Those evaluations will be the occasion to promote and disseminate our research in a playful environment. **[LPNC, CesamS, CMAP]**.

The end of this WP will mark the **last milestone** of this project **{Mi3}**, with our full system evaluated with VIPs, both in controlled (motion capture + AR serious games) and real-world settings (laser-running). Each piece of the system (navigation, target-reaching and aiming) will have been tested and optimized independently or in conjunction, and will be available as standalone libraries to facilitate the development of future applications relying on our research.

- **{D6.2}** (M48 - Code - Pu): Final transcoding library and localization algorithms.
- **{D6.3}** (M48 - Report - SCo): Final report on the laser-run (collected data, analysis and conclusions).

Estimated risks and proposed remediation avenues:

Risk Type	Description	Likelihood	Remediation
Functional	Difficulty to recruit participants due to COVID	Low - Medium	* When possible, we will conduct pre-testing online using simplified versions of our games. * We will adapt the design of our experiments to the current health recommendations (cf. [WP2]). * We will also recruit sighted participants for pre-testing.
Technical	Issues with hardware duplication & maintenance	Low	* Prototypes are already built, tested and documented. * Each site has skills in electronics, in addition to the GIPSA having a Research Engineer whose sole role in this project is to handle the replication and maintenance of the prototypes.
Methodological	Delays due to synchronisation issues between partners during [WP3 to 5]	Low	* [WP3-5] can progress independently from each other and do not entail any hardware modifications, the development of the transcoding libraries being done through code only. * Each site will rely on a common modular core of code which can be indiscriminately interfaced with different sensors, interfaces and transcoding libraries. This architecture will allow the instant adaptation of each site's progress (cf. [WP2]). * The use of common experimental protocols and measures will produce comparable & reproducible results between sites. * The laser-run evaluation of [WP6] will still be feasible even if one of the [WP3-5] hasn't achieved optimal results when it begins. Furthermore, [WP6] includes 6 months of technical development during which [WP3-5] could still progress.
	Partner leaving the consortium	Low	The objectives of this project are in direct alignment with each partners' (and their teams') current research endeavors.

II. Organisation and implementation of the project:

a. Scientific coordinator and its consortium / its team:

The **Project Coordinator**, Christian GRAFF (UGA, LPNC, *Body and Space* team), is a senior lecturer in behavioral biology (Maître de Conférences CE en Neurosciences, HDR) at Univ. Grenoble-Alpes, department of psychology. He holds a PhD in Life Sciences from Univ. Paris-Sud, initially specialized in **neuroethology** (especially weakly-electric electrolocation [87] and dolphin echolocation [88], two non-human senses used to navigate without vision), and is currently **interested in human perception and Sensory Substitution**. He is dedicated to theoretical and applied conversions between scales of physical continua, in time and space, combining perception and movement. He has contributed to the **development of SSDs** with several publications on the topic, and more recently on embodiment in virtual reality (kinesthesia) [76]. He also developed and commercialized a device converting fish low electric signals into melody. He has a long experience of **multidisciplinary interaction**, with supervision and co-supervision of trainees and publications in a variety of disciplines such as biology, psychology, computer sciences, electronics, philosophy, mathematics and HMI. He has a diversity of **coordination experience** within his department (department dean, international exchanges coordinator, …) and his University (ethical committee & sustainability coordinator).

The **consortium** comprises **five partner laboratories** organized in **three geographical sites**: (1) **Univ. Grenoble-Alpes** (UGA) and **Grenoble "Institut National Polytechnique"** through the **LPNC** (C. Graff) and **GIPSA-Lab** (S. Huet, D. Pellerin), (2) the **Ecole Polytechnique** (X) at **Palaiseau** through the **CMAP** (S. Ferrand, F. Alouges), and (3) the **Normandy University** (NU) through the **LITIS** (E. Pissaloux) and **CesamS** (E. Faugloire, B. Mantel). The laboratories belonging to the same sites have been sharing multidisciplinary collaborations for years, with inter-site collaborations starting more recently (e.g. a previous intern of UGA's team being currently co-supervised as a PhD student by both NU's partners).

UGA's **"AdViS"** (Adaptive Visual Substitution) multidisciplinary team consists of two specialists in **computer science, signal processing** and **electronics** (S. Huet and D. Pellerin), and one specialist in **biology, cognitive neurosciences and psychophysics** (C. Graff). They have been working together for several years on a modular audio-visual SSD [17], [83]. Their current endeavor revolves around the **virtual prototyping of SSDs** using motion capture (VICON) and AR to easily emulate both the testing environment, the SSD components, and implement various spatial-to-sound transcoding solutions [20]. In addition, AdViS' team has ongoing collaborations with a professional statistician for quantitative data modelling, and research engineers that will handle the (three-fold) replication of each team's prototypes (tactile belt and HRTF-enabled headphones) for this project.

At **Paris'** CMAP, the **X-audio team** has developed state-of-the-art numerical algorithms and a complete software suite for the **numerical simulation of acoustic scattering**, **binaural sound, reverberation and real-time rendering** [89]. The **guidance of VIP using 3D sounds during sports** (such as running or roller skating) is currently studied with encouraging results. This system relies on a virtual mobile guiding beacon, moving in front of the user, providing **spatialized audio cues** about its position [85]. Those audio cues include a virtual guide (similar to *Legend of Zelda* guiding fairy) which VIPs follow while running in order to stay on track. They have also developed expertise in **sensor network tracking** and **data fusion** for **robust real-time positioning**, and are currently conducting tests to apply real-time audio guidance to **laser-gun aiming**, for which they have developed a working prototype.

NU's team comprises a specialist in **biomedical engineering** & **electronics** (E. Pissaloux), and two in **cognitive ergonomics** & **human movement sciences** (E. Faugloire, B. Mantel). Both partners have worked for many years on assistive devices for VIP, and currently focus on **supplementing the spatial cognition capabilities of VIP** through the development of tactile interfaces for **autonomous orientation and navigation** [41], [84], **map comprehension** [52], but also **access to art** [90]. NU is currently working with a **vibrotactile belt** that provides ego-centered orientation information on relevant environmental cues (allowing VIP to localize and orient themselves in autonomy), and a **Force-Feedback Tablet** that allows the **exploration of 2D maps & images** [50]. Other foci of their research are the **optimization of the tactile encoding** of remote cues [45] using motion capture systems (i.e. Polhemus Fastrak) in AR, ecological modes of responses, and whole-body movements affordance-based HMI design [43], [44].

In addition to **NU's specialty in tactile assistive devices**, both **UGA and X have expertise on auditory encoding**, but on **complementary aspects**: spatialized sound generation and environment sonification for X, and vision-to-audition substitution for UGA. Both NU and UGA have strong competencies in **cognitive neurosciences, ergonomics and experimental evaluations**, and all three sites possess skills in **computer science and engineering**, to quickly prototype devices and implement new transcoding schemes. Those common skills will accelerate inter-site exchanges & protocols standardization.

All partners have **ongoing collaborations with VIP associations**, making it possible to involve them at every step of the project. **UGA** can lean on the CTRDV (Centre Technique Régional de la Déficience Visuelle), a public organization promoting VIP autonomy and education. **CMAP** has regular contacts with *MixHandi Cap sur la Vie* (MCV - a sports association for adapted physical activities, shared sports, culture, and entertainment), and the Institut National des Jeunes Aveugles (INJA). **NU** has been collaborating with both local (e.g. Normandy-Lorraine, "Espace Handicap'' of Rouen University) and national (e.g. *Fédération des Aveugles de France, Association Valentin Hauy*) VIP associations over the past years, as well as a late-blind researcher within their group (Dr. F. Serin), who is involved in the design and early testing of most of their ideas. All three sites have contacts with industrial partners for low-vision assistance technology, such as the *Panda-Guide* startup who expects advantage from our libraries, and offers access to potential end-users together with VIP's associations.

As a final word, in addition to its complementary skills, this consortium is united by a strong common conviction to improve VIP's quality of life, which motivated the birth of this project.

Implication of the scientific coordinator and partner's scientific leaders in on-going project(s):

Name of the researcher	PM	Call, funding agency, grant allocated	Project's title	Name of the scientific coordinator	Start - End
Christian GRAFF	4	Univ. Savoie Mont-Blanc	Réalité virtuelle et perception du mouvement	Michel GUERRAZ	2018 - 2021
Edwige PISSALOUX	24	ANR (AAPG20)	Inclusive Museal Guide	Edwige PISSALOUX & Abderrahim EL MOATAZ	2020 - 2025

b. Implemented and requested resources to reach the objectives:

Requested means by item of expenditure and by partner

		LPNC (UGA)	*GIPSA (UGA)*	*CMAp (Ecole Polytechnique)*	*LITIS (NU)*	*CesamS (NU)*
Staff expenses		121.7 k€	9 k€	78 k€	156.6 k€	92.8 k€
Instruments and material costs (including consumables)		3.5 k€	13 k€	7 k€	5.7 k€	5.7 k€
Building and ground costs		0	0	0	0	0
Outsourcing / subcontracting		3 k€	3 k€	3 k€	3 k€	3 k€
General and administrative costs & other operating expenses	Travel costs	7 k€	7 k€	8 k€	7 k€	7 k€
	Administrative management & structure costs	16.2 k€	3.8 k€	11.5 k€	20.7 k€	13 k€
Sub-total		**151.5 k€**	**35.8 k€**	**107.5 k€**	**193 k€**	**121.6 k€**
Requested funding		**609.4 k€**				

Staff expenses:

Site	Items	Cost
LPNC	• PhD student (36PM, M12-48, LPNC + GIPSA co-supervision): investigate multimodal transcoding schemes for target-reaching [WP4] and medium-range navigation [WP3 & 5]. Will participate in the dissemination of the project [WP1] and the laser-run evaluations [WP6]. • 2 Master-level interns: experimental evaluations of target-reaching transcoding schemes [WP4].	115.7 k€ 6 k€
GIPSA	• 3 Master-level interns: AR game development [WP3-5].	9 k€
CMAP	• Research Engineer (18PM, M24-42): improving the performance and accuracy of guidance strategies for aiming and long-distance navigation, focused on closed-loop control and detection of route events (e.g. curvatures, obstacles, ...) [WP3, 5 & 6].	78 k€
LITIS	• PhD student (36PM, M12-48): developing multimodal transcoding schemes for indoor and outdoor navigation using AR games and localization algorithms [WP3 & 6]. Will participate in the dissemination of the project [WP1] and the laser-run evaluations [WP6]. • Ingénieur d'Étude (12PM, M1-12): system architecture, tools and experimental protocols [WP2].	115.7 k€ 41 k€
CesamS	• Post-doc (24PM, M15-39): investigate the impact of multimodal feedback on efficient spatial cognition within AR games [WP3-6]. Participate in the laser-run evaluations [WP6] • 2 Master-level interns: Ergonomics & experimental evaluation of HMI.	86.8 k€ 6 k€
Total:		**458.2 k€**

SAM-Guide relies heavily on complex AR or real-world experimental evaluations. Most of the expensive equipment needed to carry them (e.g. motion tracking platforms) are provided by the partners' laboratories, greatly reducing material costs. We thus prioritized recruiting PhD students, postdocs and research engineers to assist the consortium in the many time consuming experimental evaluations that are the crux of this project, and further foster collaborations between partners.

Instruments and material costs:

Site	Items	Cost
LPNC	• One work computer (PhD student)	1.5 k€
GIPSA	• Small electronic hardware for the replication of each partner's existing prototypes [WP3-6] • Pozyx system: large-scale indoor localisation [WP4] • Hardware and software updates for Optical Motion Capture (for the VICON) [WP4-5] • Differential GPS: outdoor localisation [WP6] • Wireless audio digital receiver/transmitter [WP4-5] • Laser-run gun and accessories [WP5-6] • One work computer (interns)	1.0 k€ 1.0 k€ 5.5 k€ 0.5 k€ 0.8 k€ 0.7 k€ 1.5 k€
CMAP	• Pozyx system: large-scale indoor localisation [WP3] • Two work computers (Interns & RI) • Laser-run gun and accessories [WP5-6] • Inertial Motion Unit for localization & tracking [WP3-6]	1.0 k€ 3.0 k€ 0.7 k€ 0.3 k€
LITIS	• Pozyx system: large-scale indoor localisation [WP3] • Differential GPS: outdoor localisation [WP3 & 6] • Laser-run gun and accessories [WP6] • One work computer (PhD student)	1.0 k€ 0.5 k€ 0.7 k€ 1.5 k€
CesamS	• Pozyx system: large-scale indoor localisation [WP3] • Laser-run gun and accessories [WP6] • Differential GPS: outdoor localisation [WP3 & 6] • One work computer (Post-doc)	1.0 k€ 0.7 k€ 0.5 k€ 1.5 k€
	Total:	**46.4 k€**

N.B.: we added an overhead of 2k€ for each partner (10k€ in total) in order to buy miscellaneous consumables such as books and small electronic hardware.

Outsourcing / subcontracting:

Combined outsourcing expenses for all partners:

Items	Cost
• Gold-standard open-access publishing costs for 5 journal articles. • Financial compensation for participants, occupational therapists and other accompanying professionals.	10k€ 5k€
Total:	**15 k€**

General and administrative costs & other operating expenses:

Operating expenses set at 12%, following the newest public Universities' standards.

Combined travel expenses for all partners:

Items	Cost
• Missions & conferences. • Consortium meetings & inter-site exchanges for PhD students and Postdocs. • Organizing seminars, workshops, participating in non-academic expositions and other networking events.	20k€ 10k€ 6k€
Total:	**36 k€**

III. Impact and benefits of the project:

. Dissemination:

(1) The academic research community: The first result of our project will be sharing and pooling resources and expertise between five different labs. We will further exchange by presenting our results among our networks (GdR TACT, Vision, Sport; IFRATH), in recurring conferences we have previously published in, whether national (e.g. IHM, HANDICAP) or international (e.g. IROS, CHI, ICCHP), publish using open-access plans (in journals such as J. on Multimodal User Interfaces, Int. J. of Human-Computer Studies, PlosOne, Frontiers in Human Neuroscience), and make anonymized data and analyses available on open-science platforms (e.g. OSF).

(2) The general public and VIP associations: the active involvement of VIP at every step of our project, facilitated by our current close collaborations with VIP charities, Orientation & Mobility schools, and handicap-related networks, will allow us to make our project known within their communities.

Information will also be relayed through the project's website, which will act as an information hub on our current research and ongoing experiments, and provide regular updates intended for VIP associations' news feed (e.g., on new technologies @ UNADEV). It will also host the project's code and serious AR games, allowing willing parties to connect, contribute, or suggest improvements.

We also plan to promote our research in various science popularization actions coordinated at national (e.g. Semaine du Cerveau, Experimentarium) and regional levels (i.e. Science Action [Normandy], Centre Culturel Scientifique et Technique [Grenoble]), with whom we have previously collaborated. The developed serious games will be demo-ed during those events (e.g. visionless First Person Shooters or maze-solving games). We will contact both specialized (e.g. Acuité, Oxytude, …) and mainstream media we have previously collaborated with to cover such events.

(3) The industry and other economic actors: we will participate in international (e.g. SightCity) as well as national and regional exhibitions (Festival de l'Excellence Normande, Autonomic Paris). We will also organize several seminars and workshops bringing together actors from the academic and private sectors, from the various fields related to this project. The goal of those seminars will be to promote this area of research in France, as well as to reach out to bordering fields that might provide or gain new insights from our research, adding more interdisciplinary emulation to our project.

b. Outcomes:

(1) The academic research community: the main contribution of SAM-Guide will be to provide modular libraries for principled and biologically-congruent multimodal transcoding schemes optimized for spatial interactions at various scales. Those transcoding schemes will improve the richness and possibilities of existing HMI, and will thus benefit other experimental research fields relying on HMI and VR/AR. SAM-Guide will also offer standardized guidelines and ecologically-relevant experimental protocols to better design and evaluate multimodal HMI for non-visual spatial interactions. The project's results will also feed into the various courses taught by the consortium members and some of the SSD will have a second life in the hands of students during practicum.

Furthermore, SAM-Guide will help further our understanding of the neurocognitive underpinnings of human spatial cognition, how the loss of vision affects them, and how multimodal training games can help remedy or rehabilitate these losses.

Overall, the Neuroscience, Psychophysics, Ergonomics and HMI scientific communities should strongly benefit from the results of this interdisciplinary project, and its theoretical results could later open new research avenues for artificial agents learning to interact with space (e.g. for autonomous vehicles).

(2) The general public and VIP associations: our experimental results, transcoding libraries, serious games and evaluation protocols developed during the project will be made freely available to the public. They could provide O&M instructors with novel strategies & techniques for mobility training exercises, backed by spatial cognition theories and experimentally validated, which will facilitate learning and rehabilitation. Furthermore, those results could contribute to the development of new standards for the practice of laser-run in handisport.

(3) The industry and other economic actors: we will ensure the accessibility of our system's design instructions (i.e. mechanical and electronic schematics of our sensory effectors and other instrumentation devices) and code (i.e. transmodal conversion libraries, motion capture platforms interfaces libraries, AR evaluation games) to both VIP associations and low-vision assistance industries and startups (such as Panda-Guide, a French start-up developing assistive devices for VIP with whom several members of the consortium already have contacts, and who expressed interest in this project).

SAM-Guide's outcomes could result in the commercialization of new products that would benefit VIP and the wider public. For example, embedding our transcoding libraries into a smartphone application, which, combined with simple headphones and/or a miniaturized tactile belt, would make a smart wearable navigation assistant that does not rely on sight (e.g. for motorcyclists). Our transcoding schemes could easily be combined to screen readers, and allow users to browse the web as if they were navigating between the nodes of the page's structure. Finally, the developed AR serious games could be published for both VIP and sighted people, offering an immersive way to experience the game's content.

In the long term, SAM-Guide could facilitate the development of novel multimodal and inclusive HMI that allow users to interact intuitively with complex environments without relying on sight. Direct transpositions could be made to smart prosthetics and wheelchairs, or situations where (1) visibility is deteriorated (e.g. firefighters, telesurgery), (2) the visual input channel is overloaded or has to be dedicated to other tasks (e.g. exploration of complex virtual visual spaces, driving, drone piloting), and/or (3) where a rapid and intuitive communication of intent is needed (e.g. human-robot interaction in manufacturing). Finally, our HMI principles could also be generalized to a wider (sighted) audience by providing an intuitive way to perceive and react to information from remote sensors, providing new ways for people to communicate between themselves or to interface with the growing number of wearables and IoT devices in our society.

References:

[1] P. Bach-Y-Rita, C. C. Collins, F. A. Saunders, B. White, and L. Scadden, "Vision Substitution by Tactile Image Projection," *Nature*, vol. 221, no. 5184, pp. 963–964, Mar. 1969, doi: 10.1038/221963a0.

[2] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," *Behav. Brain Sci.*, vol. 24, no. 5, pp. 939-973, 2001.

[3] A. Amedi, S. Hofstetter, S. Maidenbaum, and B. Heimler, "Task Selectivity as a Comprehensive Principle for Brain Organization," *Trends Cogn. Sci.*, vol. 21, no. 5, pp. 307–310, May 2017, doi: 10.1016/j.tics.2017.03.007.

[4] D. J. Huffman and A. D. Ekstrom, "A Modality-Independent Network Underlies the Retrieval of Large-Scale Spatial Environments in the Human Brain," *Neuron*, vol. 104, no. 3, pp. 611-622.e7, Nov. 2019, doi: 10.1016/j.neuron.2019.08.012.

[5] A. Amedi *et al.*, "Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex," *Nat. Neurosci.*, vol. 10, no. 6, pp. 687–689, Jun. 2007, doi: 10.1038/nn1912.

[6] S. Dehaene and L. Cohen, "Cultural Recycling of Cortical Maps," *Neuron*, vol. 56, no. 2, pp. 384–398, Oct. 2007.

[7] M. Ptito, M. Bleau, I. Djerourou, S. Paré, F. C. Schneider, and D.-R. Chebat, "Brain-Machine Interfaces to Assist the Blind," *Front. Hum. Neurosci.*, vol. 15, 2021, doi: 10.3389/fnhum.2021.638887.

[8] D.-R. Chebat, V. Harrar, R. Kupers, S. Maidenbaum, A. Amedi, and M. Ptito, "Sensory Substitution and the Neural Correlates of Navigation in Blindness," in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velazquez, Eds. Cham: Springer International Publishing, 2018, pp. 167–200.

[9] P. B. Meijer, "An experimental system for auditory image representations," *Biomed. Eng. IEEE Trans. On*, vol. 39, no. 2, pp. 112–121, 1992.

[10] J. Ward and P. Meijer, "Visual experiences in the blind induced by an auditory sensory substitution device," *Conscious. Cogn.*, vol. 19, no. 1, pp. 492–500, Mar. 2010, doi: 10.1016/j.concog.2009.10.006.

[11] M. Bujacz and P. Strumiłło, "Sonification: Review of Auditory Display Solutions in Electronic Travel Aids for the Blind," *Arch. Acoust.*, vol. 41, no. 3, Jan. 2016, doi: 10.1515/aoa-2016-0040.

[12] B. F. G. Katz *et al.*, "NAVIG: augmented reality guidance system for the visually impaired: Combining object localization, GNSS, and spatial audio," *Virtual Real.*, vol. 16, no. 4, pp. 253–269, Nov. 2012, doi: 10.1007/s10055-012-0213-6.

[13] S. Paré, M. Bleau, I. Djerourou, V. Malotaux, R. Kupers, and M. Ptito, "Spatial navigation with horizontally spatialized sounds in early and late blind individuals," *PloS One*, 2021, doi: 10.1371/journal.pone.0247448.

[14] 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, doi: 10.1145/3132525.3132545.

[15] 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, doi: 10.1109/THMS.2015.2477999.

[16] J. M. Coughlan, B. Biggs, M.-A. Rivière, and H. Shen, "An Audio-Based 3D Spatial Guidance AR System for Blind Users," in *Computers Helping People with Special Needs*, vol. 12376, K. Miesenberger, R. Manduchi, M. Covarrubias Rodriguez, and P. Peňáz, Eds. Cham: Springer International Publishing, 2020, pp. 475–484.

[17] C. Stoll, R. Palluel-Germain, V. Fristot, D. **Pellerin**, D. Alleysson, and C. **Graff**, "Navigating from a Depth Image Converted into Sound," *Appl. Bionics Biomech.*, vol. 2015, pp. 1–9, 2015, doi: 10.1155/2015/543492.

[18] P. Vindras and P. Viviani, "Frames of reference and control parameters in visuomanual pointing.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 24, no. 2, pp. 569–591, 1998, doi: 10.1037/0096-1523.24.2.569.

[19] C. Fricaud, "Substituer la vue par l'ouie : comparaison d'encodages sonores pour atteindre une cible invisible en environnement virtuel, encadrement : C. Graff, S. Huet, D. Pellerin," Master 2 Sciences Cognitives Université Lumière Lyon 2, Master's Thesis, 2019.

[20] M. Giroux, J. Barra, C. **Graff**, and M. Guerraz, "Multisensory integration of visual cues from first- to third-person perspective avatars in the perception of self-motion," *Atten. Percept. Psychophys.*, Apr. 2021, doi: 10.3758/s13414-021-02276-3.

[21] D. Camors, D. Appert, J.-B. Durand, and C. Jouffrais, "Tactile Cues for Improving Target Localization in Subjects with Tunnel Vision," *Multimodal Technol. Interact.*, vol. 3, no. 2, p. 26, Apr. 2019, doi: 10.3390/mti3020026.

[22] A. M. Larson and L. C. Loschky, "The contributions of central versus peripheral vision to scene gist recognition," *J. Vis.*, vol. 9, no. 10, pp. 6–6, Sep. 2009, doi: 10.1167/9.10.6.

[23] C. **Graff**, “Comparing sensory capacity: an estimate of functional wealth,” in *Proceedings of the 34th Annual Meeting of the International Society for Psychophysics*, Lüneburg, Germany, 2018, pp. 146–147.
[24] J. Z. Flores, L. Rasseneur, C. Gass, and R. Farcy, “Navi Rando - GPS-IMU Smart Phone Application for Helping Visually Impaired People Practicing Hiking,” 2016.
[25] S. **Ferrand**, F. **Alouges**, and M. Aussal, “An electronic travel aid device to help blind people playing sport,” *IEEE Instrum. Meas. Mag.*, vol. 23, no. 4, pp. 14–21, Jun. 2020, doi: 10.1109/MIM.2020.9126047.
[26] J. A. García-Uceda Calvo, M. M. van Wanrooij, and A. J. Van Opstal, “Adaptive response behavior in the pursuit of unpredictably moving sounds,” *eneuro*, p. ENEURO.0556-20.2021, Apr. 2021, doi: 10.1523/ENEURO.0556-20.2021.
[27] S. **Ferrand**, F. **Alouges**, and M. Aussal, “An Augmented Reality Audio Device Helping Blind People Navigation,” in *Computers Helping People with Special Needs*, vol. 10897, K. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 28–35.
[28] M. Mandanici, A. Rodà, S. Canazza, and G. Cavagnoli, “Following the Cuckoo Sound: A Responsive Floor to Train Blind Children to Avoid Veering,” in *Smart Objects and Technologies for Social Good*, vol. 233, B. Guidi, L. Ricci, C. Calafate, O. Gaggi, and J. Marquez-Barja, Eds. Cham: Springer International Publishing, 2018, pp. 11–20.
[29] M. Aussal, “Méthodes numériques pour la spatialisation sonore, de la simulation à la synthèse binaurale,” 2014.
[30] T. Sasayama, I. Oka, S. Ata, T. T. Zin, H. Watanabe, and H. Sasano, “Archery sight-system by magnetic sensors for visually impaired persons,” in *2013 International Conference on Advanced Technologies for Communications (ATC 2013)*, Ho Chi Minh, Vietnam, Oct. 2013, pp. 559–562, doi: 10.1109/ATC.2013.6698178.
[31] G. Motta *et al.*, “Overview of Smart White Canes: Connected Smart Cane from Front End to Back End,” in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velazquez, Eds. Cham: Springer International Publishing, 2018, pp. 469–535.
[32] G. Korres, A. El Issawi, and M. Eid, “TActile Glasses (TAG) for Obstacle Avoidance,” in *International Conference on Universal Access in Human-Computer Interaction*, 2014, pp. 741–749, Accessed: Mar. 26, 2017. [Online]. Available: http://link.springer.com/10.1007/978-3-319-07446-7_70.
[33] E. **Pissaloux**, R. Velazquez, and F. Maingreaud, “Intelligent glasses: A multimodal interface for data communication to the visually impaired,” Aug. 2008, pp. 120–124, doi: 10.1109/MFI.2008.4648120.
[34] J. R. Marston, J. M. Loomis, R. L. Klatzky, and R. G. Golledge, “Nonvisual Route following with Guidance from a Simple Haptic or Auditory Display,” *J. Vis. Impair. Blind.*, vol. 101, no. 4, pp. 203–211, Apr. 2007, doi: 10.1177/0145482X0710100403.
[35] S. Meers and K. Ward, “A substitute vision system for providing 3D perception and GPS navigation via electro-tactile stimulation,” 2005, Accessed: Sep. 24, 2016. [Online]. Available: http://ro.uow.edu.au/infopapers/435/.
[36] R. Velázquez, E. **Pissaloux**, and A. Lay-Ekuakille, “Tactile-Foot Stimulation Can Assist the Navigation of People with Visual Impairment,” *Appl. Bionics Biomech.*, vol. 2015, pp. 1–9, 2015, doi: 10.1155/2015/798748.
[37] G. Ghiani, B. Leporini, and F. Paternò, “Vibrotactile feedback to aid blind users of mobile guides,” *J. Vis. Lang. Comput.*, vol. 20, no. 5, pp. 305–317, Oct. 2009, doi: 10.1016/j.jvlc.2009.07.004.
[38] L. Lobo, D. Travieso, D. M. Jacobs, M. Rodger, and C. M. Craig, “Sensory substitution: Using a vibrotactile device to orient and walk to targets.,” *J. Exp. Psychol. Appl.*, vol. 24, no. 1, pp. 108–124, Mar. 2018, doi: 10.1037/xap0000154.
[39] M. Srikulwong and E. O’Neill, “A direct experimental comparison of back array and waist-belt tactile interfaces for indicating direction,” in *Workshop on Multimodal Location Based Techniques for Extreme Navigation at Pervasive*, 2010, pp. 5–8.
[40] S. U. König *et al.*, “Learning New Sensorimotor Contingencies: Effects of Long-Term Use of Sensory Augmentation on the Brain and Conscious Perception,” *PLOS ONE*, vol. 11, no. 12, p. e0166647, Dec. 2016, doi: 10.1371/journal.pone.0166647.
[41] E. **Faugloire** and L. Lejeune, “Evaluation of heading performance with vibrotactile guidance: The benefits of information–movement coupling compared with spatial language.,” *J. Exp. Psychol. Appl.*, vol. 20, no. 4, pp. 397–410, 2014, doi: 10.1037/xap0000032.
[42] J. B. F. V. Erp, H. A. H. C. V. Veen, C. Jansen, and T. Dobbins, “Waypoint navigation with a vibrotactile waist belt,” *ACM Trans. Appl. Percept.*, vol. 2, no. 2, pp. 106–117, Apr. 2005, doi: 10.1145/1060581.1060585.
[43] T. A. Stoffregen and B. **Mantel**, “Exploratory movement and affordances in design,” *Artif. Intell. Eng. Des. Anal. Manuf.*, vol. 29, no. 3, pp. 257–265, Aug. 2015, doi: 10.1017/S0890060415000190.
[44] B. **Mantel**, P. Hoppenot, and E. Colle, “Perceiving for Acting With Teleoperated Robots: Ecological Principles to Human–Robot Interaction Design,” *IEEE Trans. Syst. Man Cybern. - Part Syst. Hum.*, vol. 42, no. 6, pp. 1460–1475, Nov. 2012, doi: 10.1109/TSMCA.2012.2190400.
[45] E. **Faugloire**, L. Lejeune, M.-A. Rivière, and B. **Mantel**, “Spatiotemporal influences on the recognition of two-dimensional vibrotactile patterns on the abdomen,” *J. Exp. Psychol. Appl.*, In Press.
[46] M.-A. Rivière, 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. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 110–113.
[47] E. **Pissaloux**, R. Velazquez, M. Hersh, and G. Uzan, “Towards a Cognitive Model of Human Mobility: An Investigation of Tactile Perception for use in Mobility Devices,” *J. Navig.*, vol. 70, no. 01, pp. 1–17, Jan. 2017, doi: 10.1017/S0373463316000461.
[48] 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, doi: 10.1109/THMS.2017.2736888.
[49] E. **Pissaloux** and R. Velázquez, Eds., *Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies*. Cham: Springer, 2018.
[50] 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.
[51] M.-A. Rivière ***et al.***, “Vers une aide à la mobilité basée sur la connaissance de l’espace pour les personnes ayant une déficience visuelle,” in *HANDICAP*, Paris, 2018, p. 5.
[52] M.-A. Riviere ***et al.***, “NAV-VIR: an audio-tactile virtual environment to assist visually impaired people*,” in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Francisco, CA, USA, Mar. 2019, pp. 1038–1041.
[53] W. Elmannai and K. Elleithy, “Sensor-Based Assistive Devices for Visually-Impaired People: Current Status, Challenges, and Future Directions,” *Sensors*, vol. 17, no. 3, p. 565, Mar. 2017, doi: 10.3390/s17030565.
[54] B. Li, J. P. Munoz, X. Rong, J. Xiao, Y. Tian, and A. Arditi, “ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind,” in *European Conference on Computer Vision*, 2016, pp. 448–462, Accessed: Sep. 04, 2017.
[55] M. Gori, G. Cappagli, A. Tonelli, G. Baud-Bovy, and S. Finocchietti, “Devices for visually impaired people: High technological devices with

low user acceptance and no adaptability for children," *Neurosci. Biobehav. Rev.*, vol. 69, pp. 79–88, Oct. 2016.
[56] S. Maidenbaum, S. Abboud, and A. Amedi, "Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation," *Neurosci. Biobehav. Rev.*, vol. 41, pp. 3–15, Apr. 2014, doi: 10.1016/j.neubiorev.2013.11.007.
[57] R. Farcy and Y. Bellik, "Locomotion Assistance for the Blind," in *Universal Access and Assistive Technology*, S. Keates, P. Langdon, P. J. Clarkson, and P. Robinson, Eds. London: Springer London, 2002, pp. 277–284.
[58] M. A. Hersh (ed, R. Farcy, R. Leroux, and A. Jucha, "Electronic Travel Aids and Electronic Orientation Aids for blind people: technical, rehabilitation and everyday life points of view," 2006.
[59] L. F. Cuturi, E. Aggius-Vella, C. Campus, A. Parmiggiani, and M. Gori, "From science to technology: Orientation and mobility in blind children and adults," *Neurosci. Biobehav. Rev.*, vol. 71, pp. 240–251, Dec. 2016, doi: 10.1016/j.neubiorev.2016.08.019.
[60] Á. Kristjánsson *et al.*, "Designing sensory-substitution devices: Principles, pitfalls and potentials," *Restor. Neurol. Neurosci.*, vol. 34, no. 5, pp. 769–787, Sep. 2016, doi: 10.3233/RNN-160647.
[61] C. Lenay, O. Gapenne, S. Hanneton, C. Marque, and C. Genouëlle, "Sensory substitution: Limits and perspectives," *Touching Knowing*, pp. 275–292, 2003.
[62] D. Waller and L. Nadel, Eds., *Handbook of spatial cognition.* Washington: American Psychological Association, 2013.
[63] A. D. Ekstrom and E. A. Isham, "Human spatial navigation: representations across dimensions and scales," *Curr. Opin. Behav. Sci.*, vol. 17, pp. 84–89, Oct. 2017, doi: 10.1016/j.cobeha.2017.06.005.
[64] 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, doi: 10.1523/JNEUROSCI.4767-13.2014.
[65] H. Atilgan *et al.*, "Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding," *Neuron*, vol. 97, no. 3, pp. 640-655.e4, Feb. 2018, doi: 10.1016/j.neuron.2017.12.034.
[66] R. Hoffmann, S. Spagnol, Á. Kristjánsson, and R. Unnthorsson, "Evaluation of an Audio-haptic Sensory Substitution Device for Enhancing Spatial Awareness for the Visually Impaired:," *Optom. Vis. Sci.*, vol. 95, no. 9, pp. 757–765, Sep. 2018.
[67] 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.
[68] 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.
[69] T. Madl, K. Chen, D. Montaldi, and R. Trappl, "Computational cognitive models of spatial memory in navigation space: A review," *Neural Netw.*, vol. 65, pp. 18–43, May 2015, doi: 10.1016/j.neunet.2015.01.002.
[70] M.-B. Moser, D. C. Rowland, and E. I. Moser, "Place cells, grid cells, and memory," *Cold Spring Harb. Perspect. Biol.*, vol. 7, no. 2, p. a021808, Feb. 2015, doi: 10.1101/cshperspect.a021808.
[71] P. A. LaChance, T. P. Todd, and J. S. Taube, "A sense of space in postrhinal cortex," *Science*, vol. 365, no. 6449, Jul. 2019.
[72] V. Ayzenberg and S. F. Lourenco, "Skeletal descriptions of shape provide unique perceptual information for object recognition," *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Jun. 2019, doi: 10.1038/s41598-019-45268-y.
[73] M. F. Bonner and R. A. Epstein, "Computational mechanisms underlying cortical responses to the affordance properties of visual scenes," *PLOS Comput. Biol.*, vol. 14, no. 4, p. e1006111, Apr. 2018, doi: 10.1371/journal.pcbi.1006111.
[74] M. Auvray, S. Hanneton, C. Lenay, and K. O'REGAN, "There is something out there: distal attribution in sensory substitution, twenty years later," *J. Integr. Neurosci.*, vol. 4, no. 04, pp. 505–521, 2005.
[75] J. J. Gibson, *The senses considered as perceptual systems.* Oxford, England: Houghton Mifflin, 1966.
[76] M. Giroux, J. Barra, P.-A. Barraud, C. **Graff**, and M. Guerraz, "From Embodiment of a Point-Light Display in Virtual Reality to Perception of One's Own Movements," *Neuroscience*, vol. 416, pp. 30–40, Sep. 2019, doi: 10.1016/j.neuroscience.2019.07.043.
[77] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," *Behav. Brain Sci.*, vol. 24, no. 05, pp. 939–973, 2001.
[78] D. N. Lee, "Guiding Movement by Coupling Taus," *Ecol. Psychol.*, vol. 10, no. 3–4, pp. 221–250, Sep. 1998.
[79] O. Lahav, D. W. Schloerb, and M. A. Srinivasan, "Rehabilitation program integrating virtual environment to improve orientation and mobility skills for people who are blind," *Comput. Educ.*, vol. 80, pp. 1–14, Jan. 2015, doi: 10.1016/j.compedu.2014.08.003.
[80] F. R. Sarlegna and R. L. Sainburg, "The Roles of Vision and Proprioception in the Planning of Reaching Movements," in *Progress in Motor Control*, vol. 629, D. Sternad, Ed. Boston, MA: Springer US, 2009, pp. 317–335.
[81] D.-R. Chebat, S. Maidenbaum, and A. Amedi, "Navigation Using Sensory Substitution in Real and Virtual Mazes," *PloS One*, 2015.
[82] L. Thevin and A. M. Brock, "Augmented Reality for People with Visual Impairments: Designing and Creating Audio-Tactile Content from Existing Objects," in *Computers Helping People with Special Needs*, vol. 10897, K. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 193–200.
[83] A. Guezou-Philippe, S. **Huet**, D. **Pellerin**, and C. **Graff**, "Prototyping and Evaluating Sensory Substitution Devices by Spatial Immersion in Virtual Environments:," in *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Funchal, Madeira, Portugal, 2018, pp. 596–602, doi: 10.5220/0006637705960602.
[84] 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. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 110–113.
[85] S. **Ferrand**, F. **Alouges**, and M. Aussal, "An Augmented Reality Audio Device Helping Blind People Navigation," in *Computers Helping People with Special Needs*, vol. 10897, K. Miesenberger and G. Kouroupetroglou, Eds. Cham: Springer International Publishing, 2018, pp. 28–35.
[86] J. F. Soechting and M. Flanders, "Moving in Three-Dimensional Space: Frames of Reference, Vectors, and Coordinate Systems," *Annu. Rev. Neurosci.*, vol. 15, no. 1, pp. 167–191, Mar. 1992, doi: 10.1146/annurev.ne.15.030192.001123.
[87] C. **Graff**, G. Kaminski, M. Gresty, and T. Ohlmann, "Fish Perform Spatial Pattern Recognition and Abstraction by Exclusive Use of Active Electrolocation," *Curr. Biol.*, vol. 14, no. 9, pp. 818–823, May 2004, doi: 10.1016/j.cub.2004.04.039.
[88] Y. Doh, F. Delfour, E. Augier, H. Glotin, C. **Graff**, and O. Adam, "Bottlenose dolphin (Tursiops truncatus) sonar slacks off before touching a non-alimentary target," *Behav. Processes*, vol. 157, pp. 337–345, Dec. 2018, doi: 10.1016/j.beproc.2018.07.015.
[89] M. Aussal, "Numerical methods for spatial sound, from simulation to binaural synthesis," Theses, Ecole Polytechnique X, 2014.
[90] E. **Pissaloux** and R. Velazquez, Eds., *Mobility of Visually Impaired People*. Cham: Springer International Publishing, 2018.