Direct or Immersive? Comparing Smartphone-based Museum Guide Systems for Blind Visitors

Xiyue Wang
Miraikan – The National Museum of Emerging Science and Innovation
Tokyo, Japan
wang.xiyue@lab.miraikan.jst.go.jp

Seita Kayukawa
IBM Research - Tokyo
Tokyo, Japan
Seita.Kayukawa@ibm.com

Hironobu Takagi
IBM Research - Tokyo
Tokyo, Japan
takagih@jp.ibm.com

Giorgia Masoero*
Università degli Studi di Milano
Milan, Italy
giorgia.masoero@studenti.unimi.it

Chieko Asakawa
Miraikan – The National Museum of Emerging Science and Innovation
Tokyo, Japan
chieko.asakawa@jst.go.jp

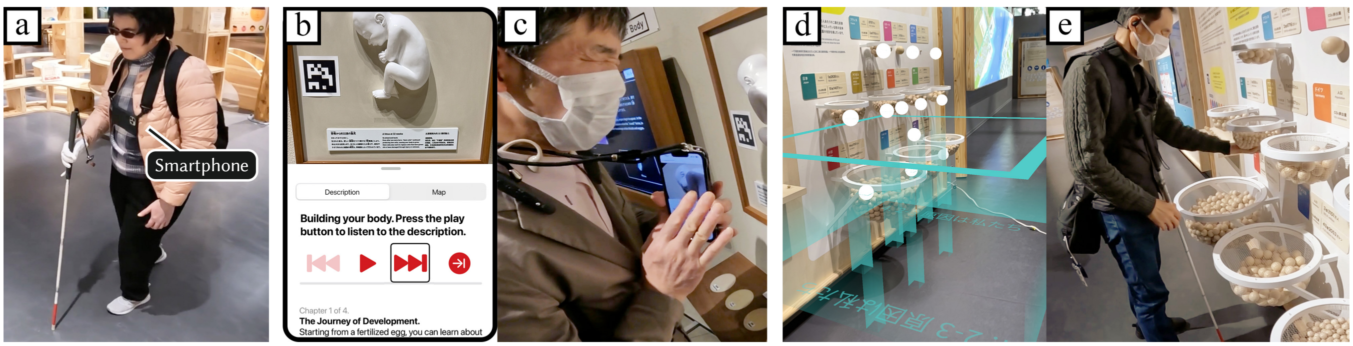

Figure 1: Museum guide systems using a smartphone. (a) User wears a smartphone for environmental sensing. (b) Direct system interface activated upon exhibit recognition, with audio description controls. (c) User browsing chapters on the phone. (d) Immersive system's virtual space with annotated locations. (e) User touching an exhibit, guided by the immersive system.

ABSTRACT

Guiding blind visitors to navigate and comprehend exhibits is crucial in museums. Two paradigms of smartphone-based guide systems have emerged: one provides direct interaction with turn-by-turn navigation and screen reader-controlled audio description, while the other offers immersive experiences with spatialized sound navigation and automatically playing audio content. However, it remains unclear which system better supports museum experiences. In a comparative study at a science museum with seven blind participants experiencing both systems, we found that immersive spatialized sound was more effective and preferred for navigation. For information provision, participants valued audio autoplay's minimal operation but expressed a need for on-demand direct control. The touch instructions provided by both systems were found inadequate for aiding interactions with tactile exhibits. Our findings suggest that a hybrid system, which adds direct interaction to the immersive experience and is adaptable to both environment and user requirements, could enhance the museum experience for blind visitors.

CCS CONCEPTS

• **Human-centered computing → Accessibility technologies**; • **Social and professional topics → People with disabilities**.

KEYWORDS

Visual impairment, Museum accessibility, Navigation, Information provision, Tactile exhibits

ACM Reference Format:
Xiyue Wang, Seita Kayukawa, Hironobu Takagi, Giorgia Masoero, and Chieko Asakawa. 2024. Direct or Immersive? Comparing Smartphone-based Museum Guide Systems for Blind Visitors. In *The 21st International Web for All Conference (W4A '24), May 13–14, 2024, Singapore, Singapore.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3677846.3677856

*This work was conducted while Giorgia Masoero was an intern at Miraikan – The National Museum of Emerging Science and Innovation.

W4A '24, May 13–14, 2024, Singapore, Singapore

1 INTRODUCTION

Museums are enhancing their accessibility to blind visitors through specialized tours [34, 57] and tactile, multisensory exhibits [15, 37]. Despite the progress, the complexity of museum environments still necessitates human assistance for navigation and information provision [32]. Blind visitors, however, aspire to an independent experience, exploring the space and engaging with exhibits at their own pace without the constant need for a human guide [43]. Technologies have been proposed to foster independence and autonomy in blind people's daily lives, including mobility assistance [44, 47] and audio descriptions of surroundings [52]. They have been adapted to museums [10, 43] to provide an independent visiting experience.

Smartphone-based guide systems that employ built-in cameras and sensors are gaining popularity due to their cost-effectiveness and independence from additional hardware like Bluetooth beacons [20, 23, 73]. Two distinct approaches have emerged in the design of these systems' user experience. The first approach focuses on direct interaction, leveraging established accessibility features like built-in screen readers such as iPhone's VoiceOver. This method provides language-based turn-by-turn navigation and accessible buttons to manage audio descriptions converted from text, ensuring clarity and practicality across diverse settings outside of museums [2, 10, 76]. On the other hand, another novel approach emerging in museums is to offer an immersive experience through audio-augmented reality. These systems guide users to a target location with spatialized audio cues and then activate the audio description or ambient sounds automatically based on the user's location and orientation in relation to the exhibits, significantly reducing the need for explicit system instructions or user input [41, 74].

While individual systems have shown promise, a comprehensive comparison between direct and immersive paradigms is lacking. Regarding navigation, spatialized sound was favored in a study for its effectiveness, with turn-by-turn instructions a close second [50]. Meanwhile, turn-by-turn navigation has been widely implemented for its straightforwardness in conveying precise information [2, 10, 76]. Museums, however, pose unique challenges with their closely arranged exhibits and irregular layouts, necessitating frequent navigation and orientation adjustments. As an integral part of the museum experience, navigation should also be enjoyable, with minimal cognitive demand. Regarding information provision, user-controlled audio descriptions have proven beneficial [1], while studies suggest that auto-activated audio-augmented reality enhances engagement and memorization [41, 74] although such findings mainly involve sighted individuals. Consequently, which type of museum guide system—direct or immersive—is best suited for the unique museum environment remains inconclusive.

To address this research gap, our study is among the first to compare different smartphone-based guide methods within a dynamic museum context. We evaluated two robust guide applications, collaboratively developed by academic and industrial teams, representing direct and immersive paradigms. The direct guide is an iPhone application that employs VoiceOver for turn-by-turn navigation and a user interface to control text-to-speech audio descriptions tied to specific exhibits, organized into chapters (Figure 1b and c). In contrast, the immersive guide leverages spatialized sound for navigation and automatically plays vivid and expressive human narration and ambient sounds upon arrival at exhibits (Figure 1d and e). We compared the paradigms and unique features through the following research questions:

- **RQ 1.** Direct vs. immersive: Which smartphone-based guide system is preferred for visiting a museum exhibition? In particular:
  - **RQ 1.1** Turn-by-turn vs. spatialized sound: Which navigation guide type facilitates effective travel within an exhibition?
  - **RQ 1.2** Button controls vs. audio autoplay: Which interface type provides efficient information provision?

Moreover, museums are increasingly offering tactile exhibits for touch-based exploration [15, 37]. Despite braille labels, such exploration often demands assistance. Interactive audio label innovations [28, 66] have promised autonomy but require additional preparations. Nonetheless, many current guide systems default to basic touch instructions without interactive feedback, which was adopted by both systems in our comparison. To evaluate their effectiveness and to identify potential improvements, we also investigated the following question:

- **RQ 2.** Accessible exhibits: How effective are touch instructions, and how can interactions be improved with tactile exhibits?

We implemented both systems in a science museum showcasing cutting-edge scientific themes. Each system was implemented in an exhibition, with tactile exhibits at most stops. We recruited seven totally blind participants to test both systems and collected data on their ratings and behaviors from video analysis. Our findings indicate a preference for spatialized sound in navigation, since it allowed for natural orientation adjustments while walking (RQ1.1). Although information delivery by either autoplay or manual button controls was comprehensible and enjoyable, participants preferred vivid narration to monotone VoiceOver and desired fewer button interactions to maintain tactile engagement. Meanwhile, autonomy in information access was crucial, indicating the need for control within immersive experiences (RQ1.2). There was a desire for a system that seamlessly blends immersive and direct elements—offering immersion by default and direct control upon user request while also adapting to user preferences and environmental factors (RQ1). The study also revealed that with only touch instructions, participants often missed tactile exhibits. Real-time overviews, accurate locational feedback, and detailed interactive guidance could enhance engagement with accessible exhibits (RQ2). Leveraging these insights, we propose a set of design considerations to advance guide systems that support the independence of blind museum visitors.

2 BACKGROUND AND RELATED WORKS

2.1 Museum Accessibility for Blind Visitors

Museums play a vital role in society, serving not only as exhibitors of historical and artistic artifacts but also as inclusive spaces that promote social engagement and empowerment [14, 36, 58, 59]. People with visual impairments are keen to engage with museum exhibits [9, 15, 16, 33] yet face considerable barriers. Accessibility needs span physical, sensory, and intellectual dimensions [68] and encompass wayfinding, tactile and multisensory exhibits, information in accessible formats, as well as staff assistance [9, 32, 37, 54].

Currently, efforts are underway in museums to improve accessibility for visually impaired visitors. These include the provision of guided tours and educational workshops [11, 15, 34, 57], the creation of accessible exhibits that engage other senses beyond sight [15, 37, 40, 67, 72], and the development of comprehensive audio descriptions [18, 26, 29]. Nonetheless, challenges remain, particularly for totally blind visitors. Guided tours are not always readily available, typically necessitating advance reservations [13, 15]. Tactile exhibitions often need additional explanations for full intellectual comprehension, provided either by sighted guides or through braille, although space for braille is often limited [56, 65] and not all blind people can read braille [60]. Furthermore, while audio guides provide valuable information, they frequently lack integration with the necessary wayfinding capabilities, hindering blind visitors from locating exhibits independently [10, 43]. The reliance on limited human resources for assistance with navigation and exhibit comprehension continues to be a significant barrier to the autonomous museum visits blind visitors desire [9, 32, 43].

Technological innovations for independent navigation, accessible information, and tactile engagement have been proposed. They are essential for creating user-friendly, independent, and enjoyable museum experiences for blind visitors.

2.2 Mobility Assistance Technologies

Mobile applications and robotics have been developed to guide visually impaired users to their destinations and help them avoid obstacles [44, 47, 52]. These innovations have been adapted for museums to facilitate wayfinding among exhibits [10, 39, 43, 55]. Indoor localization methods, such as Bluetooth beacons [2, 10, 45, 55] and RFID/IR sensors [4, 27], support wayfinding despite their use requiring the installation of external devices within the museum setting. In contrast, computer vision-based systems [51, 73], IMU sensors [5, 20], and visual-inertial odometry [23, 24] offer localization through wearable sensors, enabling users to employ their own or provided smartphones for an enhanced museum experience.

On the user experience front, turn-by-turn audio instructions have been generally employed, capitalizing on the directness and descriptiveness of human language [2, 10, 55, 76]. Spatialized audio is another method that has grown in popularity, efficiently conveying multidimensional data by indicating the location of destinations or objects [24, 42, 49, 53, 76]. Additionally, navigation robots [31] can serve as guides in museums, reducing cognitive load compared to audio instructions [43]. However, such robots come at a higher cost compared to audio-guided navigation apps available on smartphones. Comparative research by Loomis et al. [50] evaluated the navigation effectiveness of different modes and found that spatialized audio was the most effective and preferable, although turn-by-turn remained a close second. Klatzky et al. [46] found no significant difference between these modes under the condition of no cognitive load, but navigating with spatialized sound performed better when cognitive load was introduced. Nevertheless, turn-by-turn navigation remains prevalent due to its simplicity in providing precise route information, and spatialized audio cues may be too faint to hear clearly if the source is too distant [2, 76].

However, museums offer environments that differ significantly from the lab spaces or office floors typically used for indoor navigation studies, containing booths and divided rooms where exhibits are closely arranged [22, 70]. In such spaces, visitors often navigate short distances, must finely adjust their orientation to face exhibits, and frequently pause to navigate through crowds. Cognitive load should also be minimal to avoid detracting from the user's enjoyment of the exhibitions [43]. Moreover, it remains unknown whether these smartphone-based navigation instructions perform differently in museums.

2.3 Information Provision and Accessible Exhibitions

Audio guides, now common in museums, enhance exhibition understanding and are typically available through rented devices or visitors' smartphones [25, 61]. However, audio guides designed for sighted users may lack the features needed for blind people. To effectively serve them, audio guides should integrate navigation functions [10, 43], be easy to learn and use [1], and support an engaging information acquisition experience [38, 62].

There are two primary methods for delivering audio guides: one involves a control interface, and the other triggers and stops automatically. Navilens[1] offers a user interface that lets visitors choose what to listen to from a list of artifacts. Ahmetovic et al. [1] introduced MusA, an interface-based system providing interactive artwork descriptions via touch on a smartphone screen. In contrast, other research has explored environmental audio augmentation without explicit interfaces. Bederson [12] developed one of the first electronic museum guide prototypes that supplied audio content automatically based on the visitor's location. Yang et al. [74] enhanced sensory experiences in art by delivering spatialized audio cues, such as birdsong or pouring milk, through headphones that varied based on the user's location and direction relative to the artwork. Kaghat et al. [41] further personalized content delivery with an adaptive system, where the type of sound played is responsive to user interest as indicated by head gestures. These automatic systems were found to be user-friendly and enjoyable among sighted visitors, yet their effectiveness for blind visitors, who do not rely on visual context, remains to be investigated.

Increasing the accessibility of exhibits is also crucial for blind visitors [6, 15, 37, 69]. Integrating tactile objects with audio labels activated by touch has proven effective for learning [28, 35, 64]. Technologies such as push-buttons [35, 48], touch screens [30, 70], and computer vision [63, 66, 71] have been employed to create audio labels and guides on tactile objects. However, the widespread installation of such systems in museums is limited by costs associated with setup, annotation, and maintenance. A cost-effective alternative is to offer non-interactive touch instructions, which are adopted by the systems we compare. Our study examines their effectiveness and discusses their potential further needs.

While both direct (turn-by-turn navigation and screen reader controls for information) and immersive (spatialized sound navigation and audio content autoplay) paradigms have been argued as useful, their effectiveness was never compared within a real museum context. Our study fills this research gap by comparing them and investigating the most suitable methodology for the museum.

3 COMPARING DIRECT AND IMMERSIVE MUSEUM GUIDES

We conducted a study with seven blind participants to evaluate direct versus immersive guide systems inside museum exhibitions.

[1] https://www.navilens.com/

Both systems were installed on an iPhone 12 Pro housed in a hanging case, which the user wore around the neck with the camera facing forward (Figure 1a). Bluetooth open-ear headphones were used to allow simultaneous system audio and environmental sound. To evaluate system performance, we tracked participants' navigation time and errors, information recall rates, and mistakes in interacting with tactile exhibits, and collected their feedback.

3.1 Participants

Seven participants (female = 3, male = 4), aged 20 to 67 years (mean = 50.29, SD = 14.64), were recruited through an e-newsletter for people with visual impairments and compensated $45 plus travel costs. Eligibility criteria included total or legal blindness, proficiency with iPhone VoiceOver, and no prior visits to the exhibition venues. All participants were totally blind and white cane users. As seen in Table 1, they were frequent museum visitors: five (P1–P3, P6, P7) visited 2–3 times annually, while the other two (P4, P5) visited 4–6 times. Five participants (P1–P5) had previously used museum audio guides, and all were accustomed to the iPhone's VoiceOver.

3.2 System Implementation and Apparatus

The comparative study was conducted within two exhibition spaces featuring a variety of tactile exhibits. Each type of museum guide was implemented in one exhibition, in collaboration with the guide's developers and designers. The guides were not compared within the same exhibition to reduce time and cost in implementation and to ensure participants experienced fresh routes and content with each guide. Despite differing contents and paths, we ensured consistency across exhibitions in the number of stops, lengths of routes, and quantities of tactile exhibits. We also aligned the audio content duration for both exhibitions and ensured that the contents were understandable for middle schoolers and above.

3.2.1 The Direct Guide and a Biology Exhibition. The direct museum guide is an iOS-based application that provides direct guides inside a single exhibition. It provides turn-by-turn audio instructions for navigation and audio descriptions with user-controllable sequence and flow at the exhibit locations. Leveraging iOS ARKit's visual-inertial odometry [8], it tracks user movement and identifies exhibits via pre-defined image markers [7].

Based on turn-by-turn navigation [2, 76] and adjusted for short distances between exhibits, it offers three types of guidance accompanied by sonification: (1) "Go straight," followed by a soft bell sound, "Dinding," with a consistent pulse delay (the length of silence between each beep) to signify the action of moving forward. (2) "Turn Left" or "Turn Right" are followed by a distinct tap sound, "Pon," which decreases in pulse delay as the user turns correctly, culminating in a "Ding" sound to confirm correct orientation. (3) "Stop. You are near the exhibit" signals proximity within one meter of the target, with potential additional cues for fine orientation adjustment. Upon recognition of the exhibit's visual marker, the app emits a chime and automatically announces the exhibit's name.

A chapter-based screen reader interface facilitates learning about the exhibit. All descriptions are displayed on the screen and read out through VoiceOver. The initial chapter provides an overview, directing users to tactile exhibits with simple instructions, such as "Find the tactile exhibit on a shelf around waist height" or "Locate the model on the wall around eye level." The following chapters break down the exhibit's information into manageable segments, allowing users to freely navigate between "previous chapter," "play/pause," "next chapter," and "next exhibit" at any time. Completion of a chapter prompts the user to proceed to the next chapter or exhibit. Accompanying the app, a Bluetooth neckband speaker (Sharp AN-SS3) is provided, ensuring participants can hear the audio clearly.

The direct guide was implemented in a biology exhibition named "Cells in Progress," spotlighting the Nobel Prize-winning iPS cell-related knowledge and research. The stops in the guide are shown in Figure 2a. Participants learned the interface at the tutorial stop (ST) and then proceeded from ST to S0 (start location) to become familiar with navigation. Then they traveled to S1–4 in order, with an average travel distance of 3.95 m (STD = 1.59 m). The exhibit names, distances from the previous stop, and the accessibility features of each stop are summarized in Table 4.

3.2.2 The Immersive Guide and an Earth Science Exhibition. The other iOS-based app delivers an immersive museum experience through audio-augmented reality. It navigates users to their destinations with spatialized audio cues and automatically plays crafted and vivid audio content upon arrival. It senses the environment and user position through a novel markerless computer vision technique VPS (Visual Positioning System) [19], a method adapted for indoor navigation [75], and AR applications [17].

During navigation, the app emits a spatialized audio cue, a soft bell sound "Dinding" with a consistent pulse delay, from the direction of the exhibit with an increased volume as the user approaches it. Upon arrival, a distinctive chime rings, and the auto content plays. The audio seamlessly integrates vivid, expressive human voice narration with ambient sounds pertinent to the exhibit, such as ocean echoes for marine exhibits. Additionally, it provides tactile guidance (e.g., "Reach out and touch the exhibit. There are balls in a mesh bag. You can lift the bag to gauge its weight."), with thoughtfully timed pulses allowing users to interact with the exhibit. After the audio description of the exhibit, navigation to the next one begins automatically. Due to the neck-hung speaker's limitations in delivering spatialized sound, we employed open-ear earphones (Anker Soundcore AeroFit Pro with a detachable band) to ensure that users fully experience the immersive audio while remaining aware of their surroundings.

The app is implemented in an earth science exhibition named "Planetary Crisis," which explores the multifaceted aspects of global environmental issues. Participants received an orientation at Stop T before beginning their exploration from Start (S0). They progressed through four stops (S1–S4), spaced an average distance of 3.80 meters apart (STD = 0.99 m). A detailed layout is depicted in Figure 2b, and information for each stop is summarized in Table 4.

3.3 Procedure

The study was conducted in a single session that lasted for 1.5 hours. We first conducted a pre-study interview, and then the participants experienced both systems and were asked to recall exhibit information. The order of the systems was counterbalanced, where P1, P3, P5, and P7 started with the direct guide and the rest started with the immersive guide. Finally, we conducted a post-study interview to gather user ratings and comments.

Table 1: Participant demographic information.

ID	Age	Blind since	Gender	Museum visits	Museum audio guide usage	VoiceOver familiarity (1. Not at all, 7. Very much)
P1	20	2	Male	2–3 times/year	3 times	7
P2	54	10	Male	2–3 times/year	5 times	6
P3	54	26	Female	2–3 times/year	3 times	7
P4	50	35	Female	4–6 times/year	3 times	6
P5	49	3	Female	4–6 times/year	3 times	5
P6	58	7	Male	2–3 times/year	0 times	6
P7	67	50	Male	2–3 times/year	0 times	5

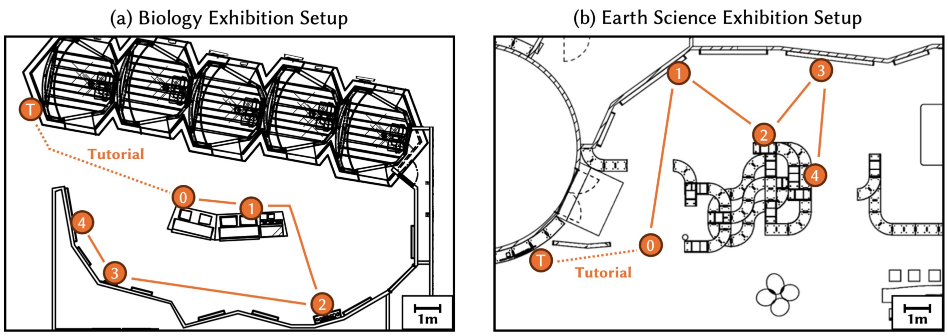

Figure 2: The study environment, including two exhibitions with numbered stops; "T" marks the tutorial stop. Lines illustrate the navigation routes.

3.3.1 Pre-study Interview. Before entering the exhibition for the study, we conducted a pre-study interview lasting roughly 10 minutes. We gathered information about participants' demographics, visual condition, museum experiences, and familiarity with museum audio guides and iPhone's VoiceOver. We also briefly explained the study procedure and informed them that after visiting each exhibition, we would ask them simple questions that required them to comprehend what they heard rather than just listening passively.

3.3.2 Main Study and Recall Test. Before the study with each system, we guided the participants to the tutorial stop and instructed them to put on the smartphone and earphones, making sure they were comfortable with them. We next advised the participants to avoid blocking the camera's view. They were allowed to use a white cane, and we might ask them to stop if there were crowds or obstacles in their path. We also informed them that the tutorial was followed by the main study, which started automatically. Before beginning with the direct guide, participants were given the option to adjust the VoiceOver speed. They were also told that while a skipping function was available, they could skip only one exhibit for the purposes of the study. The guide usage for both systems took approximately 15 minutes at normal speed.

After participants completed their use of the guide, we conducted a recall test in a quiet location. We asked them to describe each exhibit they visited in chronological order using one sentence. Their answer was marked correct if it matched the exhibit name or accessibility feature listed in Table 4 in the correct order. Each recall test took approximately 5 minutes. The participant was then escorted to the next system's tutorial stop, visited the exhibition using the system, and subsequently undertook the recall test.

3.3.3 Post-study Interview. We ended the study with an approximately 20-minute interview, asking participants about their experiences as related to the RQs. Participants gave scores on a scale from 1 (strongly disagree) to 7 (strongly agree), offered comments on their preferences, and made suggestions for improvements.

To investigate RQ 1.1, which compared turn-by-turn and spatialized sound navigation within an exhibit, we asked about the easiness and enjoyment of each type of navigation. Easiness encompassed learning and using the navigation, while enjoyment included feeling enjoyment with a minimal cognitive load. Autonomy was not assessed, as neither navigation type allowed participants to choose their destinations independently. To investigate RQ 1.2, which compared screen reader controls and audio autoplay, we evaluated the easiness, enjoyment, and autonomy of each type of information provision, where easiness included the ease of learning and using the method. We then asked the participants to compare the systems overall and choose their preferred system, corresponding to RQ1. Finally, investigating RQ.2, we asked about the effectiveness of touch instructions provided by both systems and solicited suggestions for further improvement.

3.4 Video Analysis

In order to gain an objective understanding of the participant's performance in navigating, acquiring information, and interacting with tactile exhibits, two research team members reviewed the recorded

Table 2: Navigation events, time, and time differences, in a format of turn-by-turn/spatialized sound.

ID	Assistance required	External factors	Time for 1 m (diff)
P1	1/0	0/2	3.29/3.66 (+0.37)
P2	1/0	2/0	6.01/3.48 (-2.52)
P3	2/0	1/0	5.92/3.82 (-2.10)
P4	4/0	0/1	7.38/4.75 (-2.63)
P5	1/0	1/0	6.36/4.71 (-1.65)
P6	0/1	1/1	6.76/4.31 (-2.45)
P7	0/0	0/1	6.62/4.71 (-1.91)

videos. They identified common events that disrupted the smooth experience, drawing on observations and previous research [2]. Discrepancies in video coding were resolved through additional video reviews and discussions until they reached a consensus.

Navigation performance was evaluated by counting two types of events:

- **Assistance Required:** Additional assistance from the experimenters was provided when participants needed help using a feature or making correct turns.
- **External Factors:** Occurrence of external interference, such as another person obstructing the path.

Additionally, we measured the **average time to travel one meter**, excluding segments with system errors or external disruptions. This metric served as an indicator of the participant's travel speed, movement smoothness, and veering behavior.

Information provision performance was measured by the occurrence of assistance:

- **Assistance Required:** Participants required help understanding functions or content, prompting further explanation by the experimenters.

Interaction with tactile exhibits was quantified by two events:

- **Not Found:** Upon arriving at a tactile exhibit, participants either failed to locate the intended exhibit within reach or found something different than what was instructed.
- **Not Followed:** During tactile exploration, participants failed to follow the audio instructions, touching incorrect locations and necessitating further guidance from experimenters.

4 RESULTS

4.1 Navigation Performance and Effectiveness

Table 2 summarizes user performance obtained from video analysis. During navigation, five participants (P1–P5) required additional assistance with turn-by-turn instructions. Three (P1, P2, and P5) continued moving forward despite turning cues, leading to confusing instructions that necessitated guidance for pausing and orientation correction. Two others (P3 and P4) needed reminders to initiate navigation by pressing the "next exhibit" button. In contrast, only one user with spatialized sound sought clarification on whether to keep the spatial sound directly ahead at all times. Encounters with crowds occurred for four participants (P2, P3, P5, P6) using turn-by-turn instructions and another four (P1, P4, P6, P7) using spatialized sound, with each system eliciting different responses. With spatialized sound, users were instructed to halt until the way was clear. With turn-by-turn instructions, participants sometimes experienced confusion during turn instructions, often requiring assistance to align their orientation to silence the turning sound, which could be irritating over time. Spatialized sound also resulted in a faster average travel time per meter (Mean = 4.21 s) compared to turn-by-turn (Mean = 6.05 s), with participants taking 1.65 to 2.63 seconds less, except for P1. We observed that veering was common to both systems. With spatialized sound, participants naturally adjusted their orientation while walking, resulting in a smoother, curved trajectory. With turn-by-turn instructions, those who took significantly longer displayed a zigzag walking pattern, often stopping to reorient themselves after veering off course.

The user ratings of system effectiveness, presented in Figure 3, indicate that spatialized sound was generally preferred over turn-by-turn instructions, despite both being positively received. Participants rated spatialized sound as easy to understand (median = 6) and to use (median = 7), while turn-by-turn instructions were rated lower in ease of understanding and use (median = 5). Confirming our previous observations, three participants (P1, P2, P5) noted that turn-by-turn instructions were more time-consuming due to the need to stop and reorient, a process they found counterintuitive.

> A1: *"The turn-by-turn guide took me longer because I had to slow down and turn until the turning sound stopped. With spatialized sound, I could continue walking toward the sound, adjusting my direction on the move. It was intuitive."* P5

> A2: *"(With spatialized sound) I could easily adjust my direction when I was a little off the track. It made walking easier, since it resembled my everyday experiences like seeking out a subway turnstile or a car by following their sounds."* P1

Furthermore, four participants (P1, P2, P4, P5) found turning instructions particularly counterintuitive and difficult to understand.

> A3: *"I understood the tutorial but still felt I needed more practice to master the turning cues. When the 'Pon' sound for turning came up, I couldn't intuitively grasp its meaning and how much I should turn."* P4

Nevertheless, P7 valued the control provided by turning instructions, enabling precise orientation adjustments. P3 observed that spatialized sound made it simpler to avoid people, as it was less disruptive when stopping, providing a clearer acoustic understanding of the environment.

Regarding enjoyment, participants strongly agreed that spatialized sound was enjoyable (median = 7) and agreed that it posed minimal cognitive load (median = 6). In comparison, turn-by-turn instructions were rated as somewhat enjoyable (median = 5) with a moderately minimal cognitive load (median = 5). One participant (P4) enjoyed the novelty of navigating with spatial sound, while two (P2 and P5) observed that processing turn-by-turn cues increased cognitive load, diminishing enjoyment.

> A4: *"With spatialized sound, I could move instantly, so it didn't feel it was burdensome. With turn-by-turn, I had to think about how to turn to the correct direction—I tried moving a bit left, then right, which I found annoying."* P2

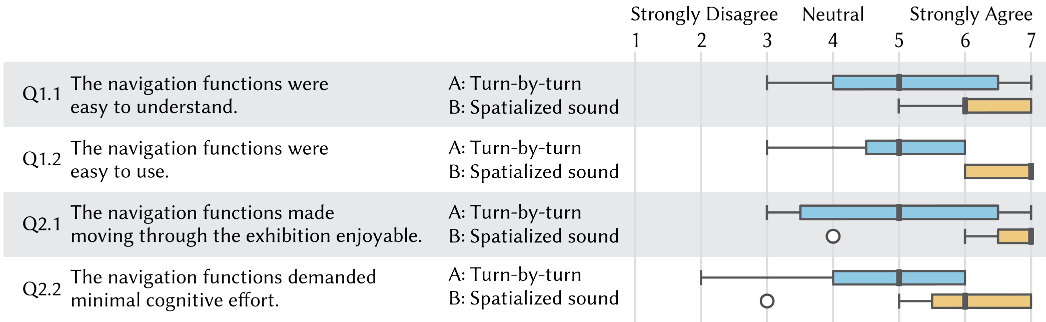

Figure 3: Questionnaire results of navigation easiness (Q1.1, Q1.2) and enjoyment (Q2.1, Q2.2) for A: turn-by-turn and B: spatialized sound navigation on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Table 3: Information provision speed, recall rate, and events occurring during interaction with tactile exhibits, in the format of screen reader controls/autoplay.

ID	Audio speed	Recall rate	Not found	Not followed
P1	95%/50%	100%/100%	0/0	1/1
P2	65%/50%	75%/100%	0/1	1/1
P3	80%/50%	50%/100%	1/2	0/2
P4	65%/50%	75%/75%	1/1	1/2
P5	55%/50%	100%/100%	1/0	1/1
P6	55%/50%	75%/100%	1/1	0/1
P7	65%/50%	100%/100%	1/0	0/1

Furthermore, one participant (P3) mentioned that tapping the "next exhibit" button for each navigation cue with turn-by-turn instructions was cognitively demanding.

To further enhance navigation, participants suggested several improvements. First, they desired more detailed location information when approaching exhibits to prevent bumping them or stopping too far away. P7 wanted cues like "Just move forward a little more," while P1 and P3 sought information on their exact stop position within an exhibit, such as on the far left or center, to facilitate finding tactile exhibits. P4 recommended tactile markers on the floor to help users stop at the correct spot. Second, they hoped the navigation systems would be responsive to their surroundings. P1 hoped it could detect and navigate around people, while P5 pointed out that the immersive system's background music during navigation could obscure important environmental cues. Third, they recommended receiving an overview of the exhibition before exploring individual exhibits. P6 proposed an initial walk-through accompanied by others to familiarize themselves with the layout, and P2 proposed this practice using navigation robots. P2 further expressed a desire to select specific exhibits themselves for autonomous exploration inside the exhibition.

4.2 Information Provision Performance and Effectiveness

Table 3 presents the participants' recall performance. Using screen reader controls, three out of seven participants recalled all exhibits correctly. Three had 75% accuracy, and one had a 50% recall rate. In contrast, with autoplay, six out of seven participants accurately remembered all exhibits, and one had a 75% recall rate. Notably, three participants remembered more information with autoplay than with screen reader controls. We also observed that while autoplay maintained a consistent normal speed, all participants using screen reader controls opted for a faster reading speed. No participant required assistance for functions or content.

User ratings of information provision effectiveness are presented in Figure 4. Participants unanimously agreed that screen reader controls were easy to understand (median = 7) and autoplay was also straightforward (median = 7). The two systems were rated as equally easy to use (median = 7 for both) and enjoyable in learning the contents (median = 7 for both). Despite having the same median scores, participant comments highlighted nuanced differences. Four participants (P2, P3, P4, P6) found that repeatedly pressing the button for the next chapter could be time-consuming and cumbersome.

> A5: *"Pressing the button every time was tiring. Holding a white cane, I had to operate the phone with the other hand. It was inconvenient since I also wanted to touch the exhibits."* P3

P3 further suggested implementing voice commands for system operation. Five participants (P1, P2, P4–P6) observed that the iPhone VoiceOver's monotone description was inferior to autoplay's natural voice narrations with background sounds, potentially impacting content understandability and enjoyment. P5, however, appreciated the screen reader control's structured approach to providing information.

> A6: *"The natural-sounding narration (in the immersive system) was easy to understand and enjoyable. The monotone VoiceOver made long paragraphs boring and hard to remember."* P2

> A7: *"The controls let me learn step by step. It started with an overview of the layout, which helped me identify items by touch. Then I could request more details. Its chapter numbers easily allowed me to revisit sections I wanted to hear again."* P5

In terms of autonomy, there was a slight preference for screen reader controls. Participants strongly agreed that they could understand the exhibits of their own interest and at their own pace using screen reader controls (median = 7), and they agreed it was possible using autoplay (median = 6). Three participants (P2, P4, P5), however, felt rushed by autoplay, expressing a preference for paused

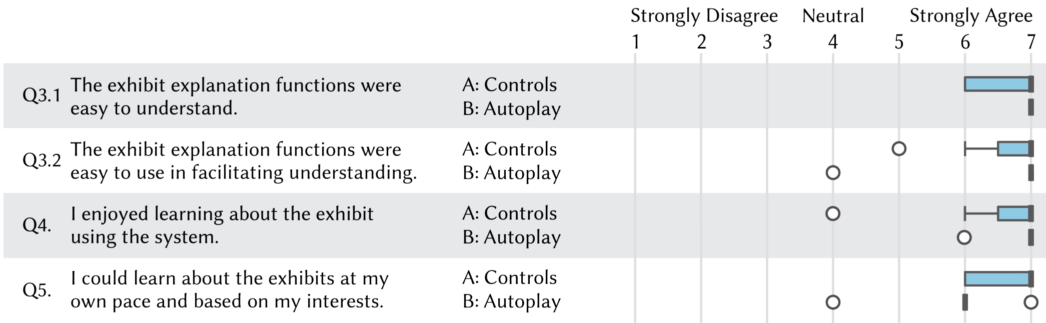

Figure 4: Questionnaire results of information provision easiness (Q3.1, Q3.2), enjoyment (Q4), and autonomy (Q5) for A: (screen reader) controls, and B: (audio) autoplay on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

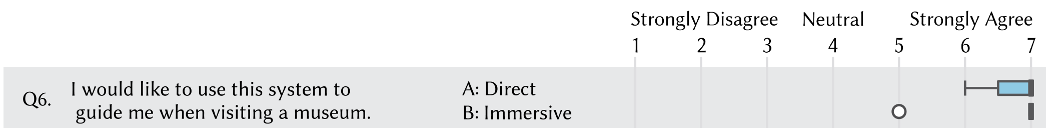

Figure 5: Questionnaire results of user favorability (Q6) for A: direct paradigm and B: immersive paradigm on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

narration, particularly when they found something not referred to in the narration or were engaging with braille displays.

> A8: *"The immersive system was understandable but didn't accommodate my pace. When I was touching something, I felt like I was scrambling to keep up with the narration. Sometimes, I finally understood the place it referred to when the narration ended. Thus I appreciate the direct system that can pause and replay."* P2

Six participants (P1–P6) expressed a desire to replay sections while using autoplay, and one participant (P7) felt the need to skip content. Moreover, looking to the future for both systems, P5 wished for the ability to ask follow-up questions based on the information received, such as "Could you repeat the name of the cells?"

4.3 Preference

As shown in Figure 5, participants unanimously favored using both systems for museum visits (median = 7 for both). When asked to choose between the two, four participants (P1, P2, P4, P6) favored the immersive system, while the remaining three (P3, P5, P7) preferred the direct system. Three participants (P1–P3) noted that it was their first time exploring a museum using an application designed for independence, and they appreciated the novelty of both systems. Four participants (P2, P3, P5, P7) favored the direct system for learning about exhibits, since it allowed them to control their learning pace by pausing, skipping, and replaying content. On the other hand, the immersive system was their preferred navigation choice due to its intuitiveness. Three participants (P1, P3, P7) highlighted the advantage of the immersive system in allowing them to concentrate while following along. Among them, two (P1, P3) suggested the potential to integrate the best features of both systems to provide control as well as ease of concentration.

> A9: *"With the immersive system, I was pleased that it required no manual operation. I could follow the spatial sound and focus on the exhibits. However, with the direct system, I had the flexibility to modify settings and replay content as needed. A system that combines both, allowing me to switch between automatic and manual modes, would be ideal."* P1

> A10: *"I value the ability to learn at my own pace, but I also want to maintain focus. Voice commands might be suitable for this."* P3

4.4 Interaction with Tactile Exhibitions

With both systems providing instructions for users to touch the exhibits, Table 3 includes user interaction errors observed in the video. Five participants (P3–P7) failed to locate the touchable exhibit at least once with the direct system, while four (P2–P4, P6) had the same difficulty using the immersive system. When attempting to touch in sync with the audio explanation, four participants (P1, P2, P4, P5) using the direct system were unable to touch precisely as instructed—either touching the wrong area or failing to find the object mentioned. A similar issue occurred with all participants while using the immersive system.

The rating (Figure 6) showed that participants agreed on the systems' capability to assist them in understanding touchable exhibits independently (median = 6). Nevertheless, their comments indicated a mismatch between the provided information and their tactile interaction. Four participants (P2, P4, P6, P7) expressed their uncertainty about whether they touched the correct locations or missed any element, while the others (P1, P3, P5) reflected on how they realized they had touched incorrectly only after the explanation had ended. Three participants (P4, P5, P7) expressed a desire for an introductory overview of the layout of tactile exhibits, and P5 noted that the direct system was more effective because it included

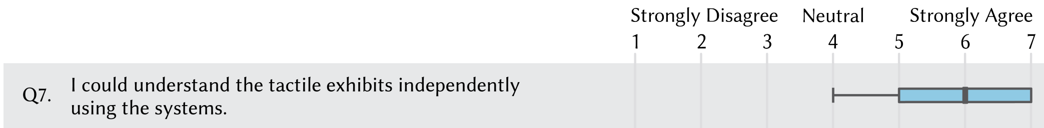

Figure 6: Questionnaire result of the effectiveness of guiding tactile exhibits (Q7) for both systems on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

such an overview as an initial chapter, whereas the immersive system's sequential left-to-right guidance could lead to errors if the starting position were misaligned.

> A11: *"I want to know if I'm touching the right thing. It's challenging with exhibits like the CO2 one (Figure 1e), where items are placed at varying heights. An overall image including sizes and the total number of objects would help me avoid missing things."* P4

Furthermore, three participants (P5–P7) noted that braille was their only means of verification. To improve interaction, they suggested incorporating hand-tracking feedback (P3, P4, P7), tangible cues such as tactile arrows (P6), or sounds emitted by the exhibits (P3) to confirm interaction with the correct elements.

5 DISCUSSION

5.1 Comparing Direct and Immersive Museum Guide Systems

To address RQ1, our study compared the effectiveness of direct and immersive guide systems for blind museum visitors. User ratings indicated both systems facilitated easy and enjoyable navigation within exhibitions and effective information acquisition about exhibits. Moreover, they supported autonomous learning tailored to individual interests and pace. Despite overall positive user ratings, detailed video analyses and user comments unveiled nuanced differences in experiences, highlighting preferences and identifying areas for improvement.

5.1.1 Navigation. From user behavior, we found that spatialized sound was easier and smoother to use than turn-by-turn guidance. While five participants needed additional assistance with the turn-by-turn guide for tasks like turning or initiating navigation, only one required confirmation while navigating with spatialized audio cues. Notably, turn-by-turn instructions led to increased veering and time consumption.

User evaluations also reflected a preference for spatialized sound, highlighting its ease of use, enjoyable experience, and reduced cognitive load. Participants particularly criticized turn-by-turn instructions' turning process as being counterintuitive, difficult to understand, and more cognitively demanding than adjusting orientation during walking using spatialized sound (A1–A4). Despite prior studies [2, 10, 76] endorsing the turn-by-turn guide for its easiness, our findings suggest that spatialized audio was more suited to the museum setting, due to the need for fine and frequent orientation adjustments in museums and the simplicity of the sound (A2, A4). Unlike typical buildings, museum spaces often require varied turning angles, more challenging than 90-degree turns [3]. Adjusting to these can be time-consuming and cognitively demanding without physical aids like tactile paving. Spatialized sound facilitated smoother gradual orientation changes, which reduced effort and correspondingly enhanced enjoyment. Furthermore, the single, consistent audio cue for walking facilitated straightforward movement. Conversely, multiple sounds, especially the turning cue sonification with changing pause delays, was harder to learn and interpret (A3), echoing previous findings that compared text-to-speech and sonification [21]. The less intrusive nature of spatialized sound also proved beneficial during pauses in movement, allowing better environmental awareness.

However, this study was confined to simple, short paths without significant obstacles. For more complex routes, turn-by-turn instructions might be more practical, particularly for precise navigation or when spatial audio comprehension is difficult, since P7 appreciated its precision. Future implementations could combine turn-by-turn instructions with spatialized audio for a balanced, context-sensitive solution that facilitates intuitive walking toward targets and allows for more precise directional adjustments when necessary. This hybrid approach could be suitable for diverse navigation scenarios within museum environments.

5.1.2 Information Provision. User ratings indicated that both systems were user-friendly and engaging, yet recall rate showed a better outcome with audio autoplay. This difference could stem from two factors. First, the lower cognitive load required during navigation with spatial audio may have allowed better memorization of exhibit contents. Second, autoplay's vivid and natural narration and ambient sounds enhanced engagement and comprehension, particularly with long texts, compared to VoiceOver's monotone delivery (A6), echoing findings that enriched audio descriptions improve user experiences [38, 62].

Another prominent finding was that participants perceived screen reader controls as offering more autonomy, appreciating their allowance for self-paced learning. Despite users criticizing the repetitive effort required to operate the phone to proceed (A5, A9), the availability of functions like replay and skip was valued. During the study, even though no skipping occurred, participants frequently replayed content with screen reader controls, noting that full-exhibit autoplay was hurried (A8). One participant highlighted screen reader controls' chapters, which provided structured content delivery from general overviews to detailed explanations, and the flexibility to pause (A7). These findings suggest that control options are essential and that they should be contextually adaptive. Voice commands could serve as a useful adjunct to the autoplay (A10), offering user controls when necessary.

5.1.3 Overall. In summary, while immersive experiences were appreciated for their ease and enjoyability, blind visitors also desired direct controls for autonomy. They hoped for a balance between them (A9, A10), which could be adaptive to user needs and the environment.

5.2 Enhancing Interaction with Tactile Exhibits

As museums enrich their tactile exhibits, it is crucial that guide systems assist blind visitors in independently navigating and comprehending these exhibits through touch. Presently, both systems offer simple touch instructions without interactive feedback. However, user performance and feedback suggest that such guidance may not be adequate to effectively support locating and engaging with the exhibits (RQ2).

First, difficulties in finding tactile exhibits arose if participants were not positioned precisely as assumed by the instructions. For instance, a user following the direction to find a tactile exhibit "1 meter ahead" could miss the object if their stopping point was skewed. This could be improved by narrowing the navigation target zone or better tracking where users stop and offering relative location feedback. Adding tactile markers on the floor (e.g., blister paving) is another simple fix, but it might require users to find them.

Second, when exhibits featured multiple tactile objects or parts, users often lost track of referenced locations. This issue affected over half of the participants using the direct system and all of those using the immersive system. Instructions like "on the far left" could be ineffective as blind users struggle to gauge the extent and the boundary, often leading to either insufficient or excessive movement. Braille is the current standard for tactile confirmation, but it is not accessible to all. Participants often expressed uncertainty about their touch location and lacked a comprehensive overview (A11). Tactile indicators or auditory cues associated with different parts of an exhibit might enhance the touch interaction. However, they require additional exhibit setups and thorough testing for effectiveness. Hand-tracking technology on 3D objects can offer adaptive information [63, 71] and has already been employed in devices like iPads [64]. Incorporating such technology in museum guides is needed to enable independent interaction with tactile exhibits.

5.3 Design Considerations for Smartphone-based Museum Guide Systems

To enhance museum guide systems, we recommend the following design considerations:

(1) Utilize immersive spatialized sound for navigating short distances and non-standard angles. When confusion arises or spatialized audio is inaccessible, switch to turn-by-turn instructions for precise guidance.
(2) Automatically provide engaging information using vivid natural voice and ambient sound upon arriving at an exhibit, allowing user control as needed. Organize content into structured chapters and employ voice commands or buttons for sequential navigation.
(3) Present an exhibit layout overview and guide users based on their relative location for seamless tactile exploration. Provide detailed guidance based on their hand positions for accurate tactile interactions.
(4) Enable automatic and smooth transition to the next exhibit once the current interaction concludes or when the user moves away. Alert users when they are close to the target.

We also suggest enhancing the guide's adaptability to the environment and user needs:

(5) Detect the crowds, navigate around them, or pause until the route is clear with the option to attenuate audio during waiting times.
(6) Respond to user inquiries throughout the navigation and exhibit engagement.

6 CONCLUSION AND FUTURE WORK

Our study investigated smartphone-based museum guides for blind visitors by comparing two emerging paradigms: direct (turn-by-turn navigation and screen reader controls for information) and immersive (spatialized sound navigation and audio content autoplay). We evaluated their effectiveness in navigation, information access, and tactile exhibit interaction with seven blind participants in a science museum, gathering preferences and suggestions. Although both systems were favorably rated, user behavior and feedback suggested an integration of them: enriching immersive experiences with direct control options and enhancing tactile exhibit engagement.

As the first study comparing different guide system paradigms in a museum setting, our investigation was limited to qualitative user experience evaluations with a small participant group, without a detailed verification of technical specifics. Despite its limitations, the comparison provides actionable design considerations for future museum guide systems. Future work should validate these advanced features with a broader user base in real-world museum contexts.

ACKNOWLEDGMENTS

We thank all the participants who took part in our user study. We also thank the EveryWare Lab at the University of Milan, GATARI Inc., and NOMLAB at the Nomura Group, for their contributions to the design and implementation of the direct and immersive systems in the museum.

REFERENCES

[1] Dragan Ahmetovic, Cristian Bernareggi, Kristian Keller, and Sergio Mascetti. 2021. MusA: artwork accessibility through augmented reality for people with low vision. In *Proceedings of the 18th International Web for All Conference* (Ljubljana, Slovenia) *(W4A '21)*. Association for Computing Machinery, New York, NY, USA, Article 1, 9 pages. https://doi.org/10.1145/3430263.3452441
[2] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: a navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 90–99. https://doi.org/10.1145/2935334.2935361
[3] Dragan Ahmetovic, Uran Oh, Sergio Mascetti, and Chieko Asakawa. 2018. Turn Right: Analysis of Rotation Errors in Turn-by-Turn Navigation for Individuals with Visual Impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 333–339. https://doi.org/10.1145/3234695.3236363
[4] Saleh Alghamdi, Ron van Schyndel, and Ahmed Alahmadi. 2013. Indoor navigational aid using active RFID and QR-code for sighted and blind people. In *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*. 18–22. https://doi.org/10.1109/ISSNIP.2013.6529756
[5] E Apostolopoulos, Navid Fallah, Eelke Folmer, and Kostas E Bekris. 2010. Feasibility of interactive localization and navigation of people with visual impairments. *Proceedings of the 11th IEEE Intelligent Autonomous Systems (IAS-10), Ottawa, ON, Canada* (2010), 22–32.
[6] Vassilios S Argyropoulos and Charikleia Kanari. 2015. Re-imagining the museum through "touch": Reflections of individuals with visual disability on their experience of museum-visiting in Greece. *Alter* 9, 2 (2015), 130–143.

[7] ARKit. 2024. Detecting Images in an AR Experience. https://developer.apple.com/documentation/arkit/arkit_in_ios/content_anchors/detecting_images_in_an_ar_experience

[8] ARKit. 2024. Understanding World Tracking. https://developer.apple.com/documentation/arkit/arkit_in_ios/configuration_objects/understanding_world_tracking

[9] Saki Asakawa, João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2018. The Present and Future of Museum Accessibility for People with Visual Impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 382–384. https://doi.org/10.1145/3234695.3240997

[10] Saki Asakawa, João Guerreiro, Daisuke Sato, Hironobu Takagi, Dragan Ahmetovic, Desi Gonzalez, Kris M. Kitani, and Chieko Asakawa. 2019. An Independent and Interactive Museum Experience for Blind People. In *Proceedings of the 16th International Web for All Conference* (San Francisco, CA, USA) *(W4A '19)*. Association for Computing Machinery, New York, NY, USA, Article 30, 9 pages. https://doi.org/10.1145/3315002.3317557

[11] Alternative Athens. 2024. Sensing the Acropolis: Acropolis Visit for the Blind. https://www.alternativeathens.com/tours/sensing-acropolis-a-tour-visually-impaired/.

[12] Benjamin B. Bederson. 1995. Audio augmented reality: a prototype automated tour guide. In *Conference Companion on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95)*. Association for Computing Machinery, New York, NY, USA, 210–211. https://doi.org/10.1145/223355.223526

[13] Ruth Bieber and John Rae. 2013. From the Mind's Eye: Museum and Art Gallery Appreciation for the Blind–Canadian Perspectives. *Disability Studies Quarterly* 33, 3 (2013).

[14] Graham Black. 2012. *The engaging museum: Developing museums for visitor involvement*. Routledge, London.

[15] Matthew Butler, Erica J Tandori, Vince Dziekan, Kirsten Ellis, Jenna Hall, Leona M Holloway, Ruth G Nagassa, and Kim Marriott. 2023. A Gallery In My Hand: A Multi-Exhibition Investigation of Accessible and Inclusive Gallery Experiences for Blind and Low Vision Visitors. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (, New York, NY, USA,) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 9, 15 pages. https://doi.org/10.1145/3597638.3608391

[16] Serap Buyurgan. 2009. The Expectations of the Visually Impaired University Students from Museums. *Educational Sciences: Theory and Practice* 9, 3 (2009), 1191–1204.

[17] Kelvin Cheng, Kensuke Koda, and Soh Masuko. 2022. Reimagining the Stadium Spectator Experience using Augmented Reality and Visual Positioning System. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 786–787. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00167

[18] Alison F. Eardley, Louise Fryer, Rachel Hutchinson, Matthew Cock, Peter Ride, and Joselia Neves. 2017. *Enriched Audio Description: Working Towards an Inclusive Museum Experience*. Springer International Publishing, Cham, 195–207. https://doi.org/10.1007/978-3-319-55224-8_13

[19] Masashi Eshima, Takaaki Kato, and Andreas Kuhn. 2024. Make it possible for all devices to understand the real world "3D Environment Sensing". https://www.sony.com/en/SonyInfo/research/technologies/3D_environment_sensing/

[20] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2012. The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 425–432. https://doi.org/10.1145/2207676.2207735

[21] Alexander Fiannaca, Ilias Apostolopoulous, and Eelke Folmer. 2014. Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility* (Rochester, New York, USA) *(ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/2661334.2661453

[22] Natália Filová, Lea Rollová, and Zuzana Čerešňová. 2022. Route options in inclusive museums: Case studies from Central Europe. *Architecture Papers of the Faculty of Architecture and Design STU* 27, 1 (2022), 12–24.

[23] German Flores and Roberto Manduchi. 2018. Easy Return: An App for Indoor Backtracking Assistance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (, Montreal QC, Canada,) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173591

[24] Giovanni Fusco and James M. Coughlan. 2018. Indoor Localization Using Computer Vision and Visual-Inertial Odometry. In *Computers Helping People with Special Needs*, Klaus Miesenberger and Georgios Kouroupetroglou (Eds.). Springer International Publishing, Cham, 86–93.

[25] Sandra Gebbensleben, Jana Dittmann, and Claus Vielhauer. 2006. Multimodal audio guide for museums and exhibitions. In *Multimedia on Mobile Devices II*, Vol. 6074. SPIE, 249–260.

[26] Sandra Gebbensleben, Jana Dittmann, and Claus Vielhauer. 2006. Multimodal audio guide for museums and exhibitions - art. no. 60740S. *Proceedings of SPIE - The International Society for Optical Engineering* 6074 (02 2006). https://doi.org/10.1117/12.641404

[27] Giuseppe Ghiani, Barbara Leporini, and Fabio Paternò. 2008. Supporting orientation for blind people using museum guides. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (Florence, Italy) *(CHI EA '08)*. Association for Computing Machinery, New York, NY, USA, 3417–3422. https://doi.org/10.1145/1358628.1358867

[28] Stéphanie Giraud, Anke M Brock, Marc J-M Macé, and Christophe Jouffrais. 2017. Map learning with a 3D printed interactive small-scale model: Improvement of space and text memorization in visually impaired students. *Frontiers in psychology* 8 (2017), 930.

[29] Desi Gonzalez. 2017. A path with choice: What we learned from designing an inclusive audio guide. In *Museums and the Web*.

[30] Timo Götzelmann. 2016. LucentMaps: 3D Printed Audiovisual Tactile Maps for Blind and Visually Impaired People. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) *(ASSETS '16)*. Association for Computing Machinery, New York, NY, USA, 81–90. https://doi.org/10.1145/2982142.2982163

[31] João Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M. Kitani, and Chieko Asakawa. 2019. CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 68–82. https://doi.org/10.1145/3308561.3353771

[32] Kozue Handa, Hitoshi Dairoku, and Yoshiko Toriyama. 2010. Investigation of priority needs in terms of museum service accessibility for visually impaired visitors. *British journal of visual impairment* 28, 3 (2010), 221–234.

[33] Simon Hayhoe. 2013. Expanding our vision of museum education and perception: An analysis of three case studies of independent blind arts learners. *Harvard Educational Review* 83, 1 (2013), 67–86.

[34] Phoebe Hillemann. 2024. Art Museum Tours For Visitors Who Are Blind — americanart.si.edu. https://americanart.si.edu/blog/eye-level/2016/09/317/art-museum-tours-visitors-who-are-blind.

[35] Leona Holloway, Kim Marriott, and Matthew Butler. 2018. Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (, Montreal QC, Canada,) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173772

[36] Eilean Hooper-Greenhill, Richard Sandell, Theano Moussouri, Helen O'Riain, et al. 2000. *Museums and social inclusion: The GLLAM report*. University of Leicester, Leicester, UK.

[37] Yuru Huang, Jingling Zhang, Xiaofu Jin, and Mingming Fan. 2023. Understanding Curators' Practices and Challenge of Making Exhibitions More Accessible for People with Visual Impairments. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (, New York, NY, USA,) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 10, 18 pages. https://doi.org/10.1145/3597638.3608384

[38] R Hutchinson and AF Eardley. 2023. 'I felt I was right there with them': the impact of sound-enriched audio description on experiencing and remembering artworks, for blind and sighted museum audiences. *Museum Management and Curatorship* (2023), 1–18.

[39] Dhruv Jain. 2014. Pilot evaluation of a path-guided indoor navigation system for visually impaired in a public museum. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility* (Rochester, New York, USA) *(ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 273–274. https://doi.org/10.1145/2661334.2661405

[40] Gunnar Jansson, Massimo Bergamasco, and Antonio Frisoli. 2009. A new option for the visually impaired to experience 3D art at museums: Manual exploration of virtual copies. *Visual Impairment Research* 5 (07 2009), 1–12. https://doi.org/10.1076/vimr.5.1.1.15973

[41] Fatima Zahra Kaghat, Ahmed Azough, Mohammed Fakhour, and Mohammed Meknassi. 2020. A new audio augmented reality interaction and adaptation model for museum visits. *Computers & Electrical Engineering* 84 (2020), 106606.

[42] Brian FG Katz, Slim Kammoun, Gaëtan Parseihian, Olivier Gutierrez, Adrien Brilhault, Malika Auvray, Philippe Truillet, Michel Denis, Simon Thorpe, and Christophe Jouffrais. 2012. NAVIG: Augmented reality guidance system for the visually impaired: Combining object localization, GNSS, and spatial audio. *Virtual Reality* 16 (2012), 253–269.

[43] Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2023. Enhancing Blind Visitor's Autonomy in a Science Museum Using an Autonomous Navigation Robot. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (, Hamburg, Germany,) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 541, 14 pages. https://doi.org/10.1145/3544548.3581220

[44] Sulaiman Khan, Shah Nazir, and Habib Ullah Khan. 2021. Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review. *IEEE Access* 9 (2021), 26712–26734. https://doi.org/10.1109/ACCESS.2021.3052415

[45] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (Pisa, Italy) *(SAC '16)*. Association for Computing Machinery, New York, NY, USA, 604–611. https://doi.org/10.1145/2851613.2851716

[46] Roberta L Klatzky, James R Marston, Nicholas A Giudice, Reginald G Golledge, and Jack M Loomis. 2006. Cognitive load of navigating without vision when guided by virtual sound versus spatial language. *Journal of experimental psychology: Applied* 12, 4 (2006), 223.

[47] Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. 2022. Tools and technologies for blind and visually impaired navigation support: a review. *IETE Technical Review* 39, 1 (2022), 3–18.

[48] Barbara Leporini, Valentina Rossetti, Francesco Furfari, Susanna Pelagatti, and Andrea Quarta. 2020. Design guidelines for an interactive 3D model as a supporting tool for exploring a cultural site by visually impaired and sighted people. *ACM Transactions on Accessible Computing (TACCESS)* 13, 3 (2020), 1–39.

[49] Jacobus C. Lock, Iain D. Gilchrist, Iain D. Gilchrist, Grzegorz Cielniak, and Nicola Bellotto. 2020. Experimental Analysis of a Spatialised Audio Interface for People with Visual Impairments. *ACM Trans. Access. Comput.* 13, 4, Article 17 (oct 2020), 21 pages. https://doi.org/10.1145/3412325

[50] Jack M Loomis, Reginald G Golledge, and Roberta L Klatzky. 1998. Navigation system for the blind: Auditory display modes and guidance. *Presence* 7, 2 (1998), 193–203.

[51] Roberto Manduchi and James M. Coughlan. 2014. The last meter: blind visual guidance to a target. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3113–3122. https://doi.org/10.1145/2556288.2557328

[52] Kanak Manjari, Madhushi Verma, and Gaurav Singal. 2020. A survey on Assistive Technology for visually impaired. *Internet of Things* 11 (2020), 100188. https://doi.org/10.1016/j.iot.2020.100188

[53] James R. Marston, Jack M. Loomis, Roberta L. Klatzky, Reginald G. Golledge, and Ethan L. Smith. 2006. Evaluation of spatial displays for navigation without sight. *ACM Trans. Appl. Percept.* 3, 2 (apr 2006), 110–124. https://doi.org/10.1145/1141897.1141900

[54] Rebecca McMillen. 2012. The Inclusive Art Museum: Determining Disability Access. *The International Journal of the Inclusive Museum* 4, 1 (2012), 101–116. https://www.proquest.com/scholarly-journals/inclusive-art-museum-determining-disability/docview/2793133339/se-2 Copyright - Copyright © 2012, Common Ground Research Networks, All Rights Reserved; Last updated - 2024-01-26.

[55] Apostolos Meliones and Demetrios Sampson. 2018. Blind MuseumTourer: A System for Self-Guided Tours in Museums and Blind Indoor Navigation. *Technologies* 6, 1 (2018). https://doi.org/10.3390/technologies6010004

[56] Jerzy Montusiewicz, Marcin Barszcz, and Sylwester Korga. 2022. Preparation of 3D Models of Cultural Heritage Objects to Be Recognised by Touch by the Blind—Case Studies. *Applied Sciences* 12, 23 (2022). https://doi.org/10.3390/app122311910

[57] The Metropolitan Museum of Art. 2024. For Visitors Who Are Blind or Partially Sighted — metmuseum.org. https://www.metmuseum.org/events/programs/access/visitors-who-are-blind-or-partially-sighted.

[58] ICOM The International Council of Museums. 2020. International Museum Day 2020 - Museums for Equality: Diversity and Inclusion. https://imd.icom.museum/past-editions/2020-museums-for-equality-diversity-and-inclusion/

[59] ICOM The International Council of Museums. 2022. International Museum Day 2022: The Power of Museums. https://imd.icom.museum/international-museum-day-2022-the-power-of-museums/

[60] National Federation of the Blind. 2009. The braille literacy crisis in America: Facing the truth, reversing the trend, empowering the blind. *Braille Monitor* 52, 5 (2009). https://nfb.org/images/nfb/documents/pdf/braille_literacy_report_web.pdf

[61] Mohd Kamal Othman, Helen Petrie, and Christopher Power. 2013. Measuring the usability of a smartphone delivered museum guide. *Procedia-Social and Behavioral Sciences* 97 (2013), 629–637.

[62] Kyle Rector, Keith Salmon, Dan Thornton, Neel Joshi, and Meredith Ringel Morris. 2017. Eyes-Free Art: Exploring Proxemic Audio Interfaces For Blind and Low Vision Art Engagement. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 93 (sep 2017), 21 pages. https://doi.org/10.1145/3130958

[63] Andreas Reichinger, Helena Garcia Carrizosa, Joanna Wood, Svenja Schröder, Christian Löw, Laura Rosalia Luidolt, Maria Schimkowitsch, Anton Fuhrmann, Stefan Maierhofer, and Werner Purgathofer. 2018. Pictures in Your Mind: Using Interactive Gesture-Controlled Reliefs to Explore Art. *ACM Trans. Access. Comput.* 11, 1, Article 2 (mar 2018), 39 pages. https://doi.org/10.1145/3155286

[64] Lei Shi, Holly Lawson, Zhuohao Zhang, and Shiri Azenkot. 2019. Designing Interactive 3D Printed Models with Teachers of the Visually Impaired. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300427

[65] Lei Shi, Idan Zelzer, Catherine Feng, and Shiri Azenkot. 2016. Tickers and Talker: An Accessible Labeling Toolkit for 3D Printed Models. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4896–4907. https://doi.org/10.1145/2858036.2858507

[66] Lei Shi, Yuhang Zhao, and Shiri Azenkot. 2017. Markit and Talkit: A Low-Barrier Toolkit to Augment 3D Printed Models with Audio Annotations. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 493–506. https://doi.org/10.1145/3126594.3126650

[67] Ludovico Solima et al. 2016. Do not Touch! How 3D Printing can open the way to an accessible museum. In *Management in a Digital World: Decisions, Production, Communication*. CUEIM Comunicazione.

[68] Roberto Vaz, Diamantino Freitas, and António Coelho. 2021. Perspectives of Visually Impaired Visitors on Museums: Towards an Integrative and Multisensory Framework to Enhance the Museum Experience. In *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion* (Online, Portugal) *(DSAI '20)*. Association for Computing Machinery, New York, NY, USA, 17–21. https://doi.org/10.1145/3439231.3439272

[69] Diana Walters. 2009. Approaches in museums towards disability in the United Kingdom and the United States. *Museum management and curatorship* 24, 1 (2009), 29–46.

[70] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2022. BentoMuseum: 3D and Layered Interactive Museum Map for Blind Visitors. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 35, 14 pages. https://doi.org/10.1145/3517428.3544811

[71] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2023. TouchPilot: Designing a Guidance System that Assists Blind People in Learning Complex 3D Structures. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (, New York, NY, USA,) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 5, 18 pages. https://doi.org/10.1145/3597638.3608426

[72] Paul F. Wilson, Janet Stott, Jason M. Warnett, Alex Attridge, M. Paul Smith, and Mark A. Williams. 2017. Evaluation of Touchable 3D-Printed Replicas in Museums. *Curator: The Museum Journal* 60, 4 (2017), 445–465. https://doi.org/10.1111/cura.12244 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cura.12244

[73] Yutaro Yamanaka, Seita Kayukawa, Hironobu Takagi, Yuichi Nagaoka, Yoshimune Hiratsuka, and Satoshi Kurihara. 2022. One-Shot Wayfinding Method for Blind People via OCR and Arrow Analysis with a 360-Degree Smartphone Camera. In *Mobile and Ubiquitous Systems: Computing, Networking and Services*, Takahiro Hara and Hirozumi Yamaguchi (Eds.). Springer International Publishing, Cham, 150–168.

[74] Jing Yang and Gábor Sörös. 2019. Audio-augmented museum experiences using wearable visual-inertial odometry. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) *(MUM '19)*. Association for Computing Machinery, New York, NY, USA, Article 21, 6 pages. https://doi.org/10.1145/3365610.3365616

[75] Han Zhang, Mengsi Guo, Wenqing Zhao, Jin Huang, Ziyao Meng, Ping Lu, Liu Sen, and Bin Sheng. 2022. Visual Indoor Navigation Using Mobile Augmented Reality. In *Advances in Computer Graphics*, Nadia Magnenat-Thalmann, Jian Zhang, Jinman Kim, George Papagiannakis, Bin Sheng, Daniel Thalmann, and Marina Gavrilova (Eds.). Springer Nature Switzerland, Cham, 145–156.

[76] Yuhang Zhao, Elizabeth Kupferstein, Hathaitorn Rojnirun, Leah Findlater, and Shiri Azenkot. 2020. The Effectiveness of Visual and Audio Wayfinding Guidance on Smartglasses for People with Low Vision. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (, Honolulu, HI, USA,) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376516

A DETAILED INFORMATION OF EXHIBITIONS IN THE STUDY

Table 4: Overview of exhibitions, including stops in the guide systems, distances from previous stop, exhibit names, and accessibility features.

Biology Exhibition "Cells in Progress"			
	Distance	Exhibit	Accessibility features
S1	3.0	Comparing the stem cells	Tactile graphs for comparing somatic, ES, and iPS cells' microscope images
S2	4.0	Building your body	Life-sized fetal development models from conception to 56 days on a shelf; a 220-day fetus on the wall
S3	6.5	The basic structure of cells	A wall-mounted cell structure model accessible at eye level
S4	2.3	Prolonging life with cell research	None

Earth Science Exhibition "Planetary Crisis"			
	Distance	Exhibit	Accessibility features
S1	5.5	Progress of climate change	A tactile graph showing global average temperatures from 1850 to 2050
S2	3.4	Sea-level surface in the future	Two touchable planks showing projected sea levels for 2100 under different greenhouse gas emissions scenarios
S3	3.0	CO2 emission global comparison	Three rows of baskets with wooden balls representing CO2 emissions of various countries and regions
S4	3.3	Stacked wooden blocks	Guiding participants to sit and feel the wooden blocks comprising the exhibition space