# MusA: Artwork Accessibility through Augmented Reality for People with Low Vision


Dragan Ahmetovic
dragan.ahmetovic@unimi.it
Università degli Studi di Milano

Cristian Bernareggi
cristian.bernareggi@unimi.it
Università degli Studi di Milano

Kristian Keller
kristian.keller@studenti.unimi.it
Università degli Studi di Milano

Sergio Mascetti
sergio.mascetti@unimi.it
Università degli Studi di Milano



## ABSTRACT

People with disabilities rarely visit museums due to mobility issues in reaching and navigating museum buildings and difficulties in accessing artworks. This work addresses the latter problem, considering people with low vision. To this end, we present MusA, an inclusive mobile app aimed at providing interactive artwork descriptions to museum visitors in AR. This contribution describes the mobile app design process, consisting in two iterations with end-users. The evaluation, conducted with participants with low vision, shows that AR is an effective support for accessing visual artworks. MusA is more engaging than traditional audio guides and it is also highly usable. It provides useful functionalities and it raises interest of low vision people in visiting museums.



## CCS CONCEPTS

• **Human-centered computing** → *Accessibility systems and tools*; • **Applied computing** → *Arts and humanities.*

## KEYWORDS

Art accessibility; Audio-tactile interfaces; Visual impairment.

**ACM Reference Format:**

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 (W4A '21), April 19–20, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3430263.3452441


## 1 INTRODUCTION

All over the world, museums have legal obligations to provide accessibility for visitors with disabilities. Article 30 of the United Nations Convention on the Rights of People with Disabilities [31] states that the signatory countries recognize the right of people with disabilities to take part on an equal basis with others in cultural life, and shall take all appropriate measures to ensure that people with disabilities enjoy access to cultural places such as museums. Since 2006, 177 countries have signed this convention, including all EU countries. In the United States, which did not take part in this agreement, equal access to museums is upheld by the Americans with Disabilities Act (titles II, III) [39].

To enforce these rights, many different solutions have been adopted to guarantee equal access to museums for people with visual impairments or blindness (VIB). The two main problems concern the visitors' mobility in art venues and visual artwork accessibility. This contribution focuses on the second problem, which has so far been addressed with tactile and audiotactile reproductions, audioguides and specialized tours. These solutions, however, present limitations, in particular for people with Low Vision (LV). Tactile reproductions are specifically designed for blind people, but not for people with LV. They are often placed in separate rooms, hence lacking in inclusiveness. They also take time and resources to be created and they can provide only some morphological details about the artwork. Audioguides, instead, are often very verbose and lack in interaction [6]. Moreover, they are delivered on devices that are not easy to use (*e.g.*, access to a description is achieved by typing a number). Specialized tours are highly appreciated to access a specific part of the museum. Nonetheless, they are not inclusive and the artworks to visit are predefined.

To address these problems, we present *MusA*[1], a mobile app supporting sighted and LV visitors in accessing 2D visual artworks (*e.g.*, paintings) through interactive artwork descriptions in augmented reality (AR). We describe the iterative process that was adopted for the design of *MusA*, which involved two iterations and studies with end-users. During the development of *MusA* we faced a number of challenges that required to address the following research questions:

- **RQ1** Is AR effective to support people with LV in artwork accessibility? In particular, in the given context:
  - **RQ1.1** Can people with LV easily frame a visual artwork?
  - **RQ1.2** Is visual feedback in AR useful to people with LV?
  - **RQ1.3** Can people with LV interact with visual information shown in AR on a touchscreen device?
- **RQ2** Is *MusA* considered usable and useful by people with LV? In particular:
  - **RQ2.1** How does *MusA* compare to an audio guide baseline in terms of usability?
  - **RQ2.2** Are *MusA* functionalities useful and easy to use?

[1] *MusA* ("Muse" in Italian) stands for "Museo Accessibile" ("accessible museum").



*W4A '21, April 19–20, 2021, Ljubljana, Slovenia*



ACM ISBN 978-1-4503-8212-0/21/04…$15.00
https://doi.org/10.1145/3430263.3452441

Empirical evaluation with participants with LV shows that *MusA* is effective in supporting people with LV during artwork accessibility (**RQ1**). Indeed, participants were able to easily frame the target artworks with their mobile device camera (**RQ1.1**), to access the visual feedback provided in AR (**RQ1.2**) and also to interact with the visual information shown in AR (**RQ1.3**). Achieving these results required to improve *MusA* between the first and second interaction. In the second iteration we also compared *MusA* with an audio guide (**RQ2.1**), showing that *MusA* was considered usable at a similar level but it resulted more engaging, and its novel functionalities were considered useful and easy to use (**RQ2.2**). Most importantly *MusA* was reported to increase the intention of people with LV to visit museums.

# 2 BACKGROUND

We survey the related work on the use of AR in assistive technologies for people with VIB, and we review the solutions for improving their access to museums. Additionally, we present the *DescriVedendo* mathodology for creating accessible descriptions, which are used by our system.

## 2.1 Augmented Reality for People with VIB

The applicability of AR for assistive technologies aimed at people with VIB has been investigated for different tasks, such as object recognition [3], magnification [35] and for improving color contrast and font readability [21]. In particular, AR assistive technologies have been frequently proposed to support way finding applications for people with VIB, using audio [26] or visual guidance messages [40].

Comparing audio instructions and visual guidance in AR, people with LV were shown to make fewer mistakes and to experience a lower cognitive load with visual feedback [41]. Obstacle avoidance systems in AR have also been proposed for blind people [32], people with tunnel vision [36, 37] or retinitis pigmentosa [5], showing that guidance based on AR can reduce obstacle collisions by 50% for people with VIB.

## 2.2 Museum Accessibility for People with VIB

People with disabilities, in particular people with VIB, are interested in accessing museums and enjoying artworks [7, 12, 18, 19]. Nonetheless, they meet two main barriers: mobility issues and inaccessible artworks [7, 29].

Considering mobility issues, a number of studies have investigated navigation assistance for people with VIB in museum environments. Museum visitors with VIB can be guided using speech or haptic instructions provided by a mobile or a wearable device that automatically localizes the person in the museum through Bluetooth low energy beacons [8, 22, 27], dead-reckoning [30], RFID systems [15, 16] or IR systems, coupled with object recognition [14].

An alternative approach is to have museum staff members guide people with disabilities through a specific route. Both solutions are accepted by people with VIB, the former ones are preferred because they support autonomous mobility through the museum and the visitors are free to choose what to visit and how long they stay in front of an artwork. [7].

For what concerns artwork accessibility, many solutions have been proposed to present artworks in accessible forms, in particular: tactile reproductions, audio guides and specialized workshops or visits. Specialized visits [17, 19, 20, 38] are appreciated by persons with visual impairments. However, they are not inclusive, and VIB people report difficulties in booking and attending these activities while other friends are visiting the museum [7].

Tactile reproductions are considered effective to convey morphological characteristics of an artwork [6, 24]. Nonetheless, they are not adequate to provide further information (*e.g.*, historical details). Additionally, tactile reproductions are often supported by Braille label descriptions, which can be read by only a minority of people with VIB, most of which blind.

Audio guides can also be designed to provide morphological characteristics and background details about an artwork [33, 34]. Prior works highlight that informative audio descriptions should be recorded by a neutral voice, and should include an introduction, historical details, a visual description and the technique used to create the artwork [7, 13, 25]. Moreover, it should be possible to navigate through the description, skipping from section to section [7]. However, even when well structured, audio guides are scarcely interactive.

Tactile reproductions combined with audio descriptions prove to be more effective to describe artworks [4]. However, both tactile and audiotactile reproductions often require to be installed in a separate room. This is an obstacle to ease of access and inclusion [6, 7]. The cost to produce tactile representations can be a limit for the diffusion of this solutions, too.

## 2.3 *DescriVedendo*

*DescriVedendo*[2] is a methodology for creating morphological artwork descriptions accessible to people with VIB. It was conceived by "Associazione Nazionale Subvedenti" (ANS)[3], and it was experimented in various important museums in Italy, uncovering that the provided descriptions are effective not only for visitors with VIB, but also sighted visitors. In addition to guided tours provided in selected museums, *DescriVedendo* descriptions can also be accessed online from *DescriVedendo* website[4]. Users can access online descriptions when visiting museums or from home. However, descriptions accessed in such way only provide basic web page navigation and no interaction capabilities.

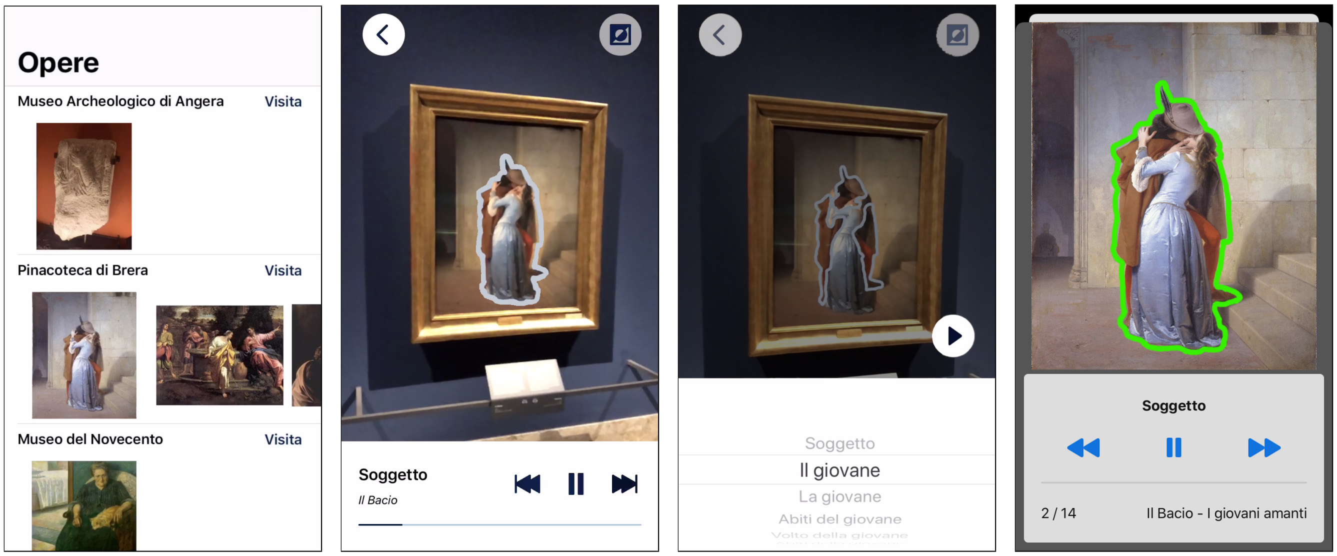
# 3 FIRST ITERATION

*MusA* was developed with a user-centric approach, based on two iterations, the first one described in this section.

## 3.1 System Analysis, Design & Implementation

*MusA* stems from a collaboration with ANS as an effort to present *DescriVedendo* descriptions in an interactive way, accessible to people with LV. Thus, the requirements analysis of system features has been conducted in collaboration with the *DescriVedendo* team, involving researchers with VIB and experts in the fields of visual disabilities, artwork descriptions and museum accessibility.

[2]The word is a crasis of the Italian words "descrivere" (describe) and "vedendo" (seeing).
[3]ANS is an Italian NPO supporting people with VIB
[4]https://www.descrivedendo.it/




(a) Museum selection (b) Chapters navigation (first iteration) (c) Chapter selection (first iteration) (d) Virtual mode (second iteration)

**Figure 1: Different screens of the *MusA* app.**

The analysis focused on addressing the problem of how museum visitors can access *DescriVedendo* descriptions during a visit, resulting in the following requirements for the *MusA* system:

**r1)** It should be accessible from visitors' own devices.
**r2)** It should be accessible to people with LV.
**r3)** It should help the user quickly find descriptions;
**r4)** It should help the user navigate the description;
**r5)** It should augment descriptions with visual information;
**r6)** It should work in multiple museums, for all artworks with *DescriVedendo* descriptions.

Addressing the above requirements, we designed and implemented *MusA* as a prototype mobile app (requirement **r1**) for *iOS*. As shown in Figure 1(a), the app is designed to work in multiple museums where *DescriVedendo* descriptions are available (requirement **r6**). Starting from a *DescriVedendo* description and the corresponding picture of an artwork, museum staff can load artworks easily using a web interface.

To quickly access the *DescriVedendo* descriptions (requirement **r3**), the app recognizes the target artworks when they are framed by the device camera and automatically selects the corresponding description, which is then read with a text-to-speech software. QR codes and tags for audio guides, could not be used because the museum director where we conducted the evaluation expressed concerns about the esthetical effects of QR codes (2AC). Furthermore, prior research reports that QR codes could be hard to find for people with VIB [10].

With this solution one question arises: can visitors with LV easily frame a 2D visual artwork with mobile device camera (**RQ1.1**)? Note that the addressed problem is not how to navigate towards the artwork (which has been addressed in prior literature [8]). Rather, the problem is to frame the artwork with the mobile device camera, when the user is already close to it.

Addressing requirement **r4**, *MusA* partitions the *DescriVedendo* descriptions into *chapters*. The idea is that each chapter describes one area of the artwork. The user can move to the previous/next chapter and pause/resume the reading (see Figure 1(b)). This logical organization of the descriptions also enables another form of navigation through the chapters: the user can tap on the artwork, which is displayed on the screen in AR, to select visual elements and play the associated chapter. A similar interaction approach has been previously proposed for web based access to artwork images [2].

To enable this type of navigation, one key interaction design challenge needs to be addressed: how to interact with the elements presented in AR? (**RQ1.3**). In particular, since the elements described in two or more chapters can overlap, there are areas of the artwork that are associated with more than one chapter. For example, considering the painting in Figure 1(b), there is one chapter describing the main painting figures (*i.e.*, the two lovers), and another chapter describing the male subject only. To address this problem, we designed the following solution (see Figure 1(c)). Upon tapping on the screen over the painting, the app shows a picker listing all the chapters describing elements that overlap with the tapped area. The user can then select a chapter from the picker.

For augmenting verbal description with visual information (requirement **r5**), each chapter is associated to an image and, when a chapter plays, the corresponding image is shown, overlaid on top of the artwork in AR. While any image can be used, we currently use this functionality to highlight the element described in each chapter with a contour, as depicted in Figure 1(b). This contour is currently created manually by system administrators, but automated segmentation and highlighting could also be used in the future. Thus a research question is whether such visual feedback is useful also for people with LV (**RQ1.2**).

To address requirement **r2**, *MusA* was designed to be accessible to users that can rely on residual sight. Thus, *MusA* is compatible with system magnifier and enlarged fonts. Additionally, the interface design presents only a few key elements at a time, thus avoiding visual clutter and supporting an easier focus on the key items with users' residual sight. Also, we implemented a functionality for applying visual filters to the artworks in AR, in order to improve their visualization for users with LV. As an example, we provided the negative image filter.

## 3.2 Experimental Evaluation

We conducted an observational study of the first iteration of *MusA* with four individuals with LV. Since *MusA* was designed to be accessible by users with LV, participants were required to be able to interact with their devices with residual vision, without a screen reader. The goal of this study was to identify the limitations of the preliminary design of the *MusA* app in order to further refine it and to provide a preliminary assessment for the outlined research questions related to AR (**RQ1**).

The experiment was conducted at the *Pinacoteca di Brera*[5], one of the most important Italian museums. The evaluation was conducted on an iPhone X device, provided by the experimenters. The study was divided into two tasks for each participant. During each task the participant was guided to a target location, in front of a painting, and was asked to use *MusA* to access its *DescriVedendo* description. The two paintings used for the tasks were "Il bacio" ("The kiss") by Francesco Hayez[6] and "Cena in Emmaus" ("Supper at Emmaus") by Caravaggio[7].

While exploring the painting, the participants were instructed to interact with the main functionalities of *Musa*: framing the picture, navigating the chapters, interacting with the image, and visualizing the image with the negative filter. A *think-aloud* protocol was adopted: participants were asked to verbally report their thoughts and actions during the test, while a member of the research team transcribed all participants' comments and another member took note of their actions and possible non-verbal reactions (*e.g.*, facial expressions).

At the end of each task the supervisor administered a series of open ended questions to understand whether the system functionalities were useful, easy to use and which issues were encountered (**RQ2.2**). In particular the questions focused on the following functionalities: camera pointing, description navigation and touch exploration of the artwork picture.

## 3.3 Results

From the analysis of study transcriptions it emerges that all participants were able to easily frame the paintings (**RQ1.1**) although two participants did not immediately understand that the painting was recognized and kept searching for a while before realizing it. This suggests that a clearer feedback needs to be provided when the visual artwork is recognized. Despite this problem, participants reported that framing the painting is easy and fast.

Two unexpected problems emerged regarding the picture framing. First, we observed that, after some time, the participants would point the device towards the ground, which resulted in the painting not being framed any more. We discovered, thanks to participants' comments, that this is due to the fact that keeping the smartphone pointed towards the painting for a long time is physically demanding. For example, two participants reported:[8]

> *"I would prefer not to keep my arm like this [pointing toward the painting] because it is tiring"*

> *"after a while it is tiring"*

The second problem is that all participants, when framing the "Cena in Emmaus" painting, tried to rotate the device in landscape, since the painting has this orientation. However, the initial *MusA* design did not consider this user need.

Considering **RQ1.2**, some participant reported that, for "Cena in Emmaus", the graphical information provided in AR was helpful to better understand the description:

> *"the red contour is helpful"*

> *"I like the idea of the red contour"*

Conversely, for "Il bacio", one participants reported that the AR graphic was not always clearly visible because it did not contrast enough with respect to the painting:

> *"I cannot distinguish the contour. Now I cannot detect the female figure because the contour is white and the figure is white as well."*

Also, the participants did not find the image filter very useful but suggested to implement a zoom functionality.

The participants experienced problems with image interaction (**RQ1.3**). In particular they did not understand that each point in the image could map to multiple chapters, and that they had to select the desired chapter from the picker. So, the problem was mainly with the effect of the touch interaction, rather than with the interaction itself. However, one participant also reported that it was hard to select a point on the image, because the image itself was not clearly visible.

> *"I have to say that it is hard to select [a point of] the painting because it is very dark"*

We should note that the painting was indeed dark and the ambient light was intentionally low.

Finally, users with LV suggested to make the chapters navigation tab (play/pause, next/previous) more visible, by using larger and more distant buttons.

# 4 SECOND ITERATION

The objective of the second iteration was to improve the system, addressing the problems in the AR artwork access and interaction (**RQ1**), which were observed during the evaluation of the first iteration. Furthermore, we aimed to assess the overall usability and usefulness of the system (**RQ2**), in particular in relation to a baseline solution: a mobile application replicating the functionality of a traditional audio guide.

[5] https://pinacotecabrera.org/en/
[6] https://pinacotecabrera.org/en/collezione-online/opere/the-kiss/
[7] https://pinacotecabrera.org/en/collezione-online/opere/supper-at-emmaus/

[8] We report here the English translation of the comments provided in Italian.

**Table 1: Participants' demographic data**

| ID | Sex | Age | Visual Impairment | | Mobile device | | Art venue visit freq. |
|---|---|---|---|---|---|---|---|
| | | | Level | Onset | Expertise | ATs used | |
| $P1$ | F | 25 | r: blind, l: 1/20 | Birth | 5 | VoiceOver | 3 / y. |
| $P2$ | F | 66 | r: blind, l: 3/20 | Birth | 2 | VoiceOver, zoom | 3 / y. |
| $P3$ | M | 62 | 1/20 | 45 y. | 4 | VoiceOver, zoom, enhanced colors | 4 / y. |
| $P4$ | F | 25 | 1/20, 3% fov | Birth | 3 | VoiceOver | rarely |
| $P5$ | M | 34 | 1/20 | Birth | 5 | VoiceOver | 5 / y. |
| $P6$ | M | 32 | 1/20 | 10 y. | 5 | VoiceOver, zoom, enhanced colors and fonts | 3 / y. |
| $P7$ | M | 59 | 1/20, 3% fov | Birth | 5 | VoiceOver, zoom, enhanced colors | 2 / y. |

## 4.1 System Improvements

From the point of view of end-user interaction, in the second iteration *MusA* was improved along the following directions. To make the artwork framing functionality more effective (**RQ1.1**), we improved the feedback that is provided when the target artwork is detected. For this, we use a combination of audio and vibro-haptic cues. Additionally, we included the support for both portrait and landscape device orientation modes.

To address the problem of the participants getting tired during prolonged camera framing, we added a *virtual mode*. After the artwork is recognized, if the user points towards the ground, the app exits from the AR modality and shows the artwork on the screen (see Figure 1(d)). When in *virtual mode* the app behaves like in the *AR mode* but without the need to keep the artwork framed: the user can navigate among the chapters, tap on the picture and see visual overlays over the picture. In *virtual mode* the user can also zoom on the image (which is instead not possible in *AR mode*). When the user moves the device back to the vertical position, the app automatically returns to the *AR mode*.

Considering the problems with the low contrast of the visual feedback provided in AR (**RQ1.2**), we designed contours with higher contrast with respect to the image. Indeed, the white contour in the previous version of the app (see Figure 1(b)) is less contrasted than the new contour (see Figure 1(d)). Since the filter functionality was not considered useful, we removed this functionality from the new version of *MusA*.

Finally, we changed how users interact with the artwork in AR (**RQ1.3**). In the second iteration we designed a different solution to select chapters by tapping on the image displayed on the touchscreen. This solution avoids using the picker, which was confusing for the participants: once the user taps on the image, the app starts reproducing the chapter that describes the area containing the tapped point. In case there is more than one of such chapters, the app selects the first one, following the one that is currently being played.

## 4.2 Experimental Evaluation

After implementing the improvements in the second iteration, we conducted a user study, involving 7 participants with LV, which was aimed at assessing the effectiveness of the improved AR system (**RQ1**) and the overall app usability, usefulness, and ease of use (**RQ2**). The evaluation compares *MusA* with a baseline app, which provides the same functionalities as a traditional audio guide.

Due to 2020 COVID-19 lockdown, we were unable to conduct the study at a museum. Instead, we designed a remote study protocol which we conducted telephonically, with participants being at their home. We recruited 7 participants with LV through ANS. Among these, $P3$ also partecipated to the first study. Nonetheless, as we see in the following, the answers to the questionnaires were similar between $P3$ and the other participants. All participants used an iPhone, besides $P2$ and $P5$ who used an iPad. Participants demographic data is presented in Table 1.

Since we could not provide an actual audio guide to the participants, in addition to *MusA*, we developed an app that simulates the behavior of a typical audio guide. This app presents only a numerical keyboard for inserting codes corresponding to different artworks, along with two buttons for play and pause. Inserting a specific code would start the reproduction of the audio description of an artwork, which could be paused and resumed with the corresponding buttons.

As the study was conducted remptely, the participants needed to use their own mobile device for the test. Therefore we shared *MusA* and the audio guide app with the participants through Testflight[9]. In order to test the artwork selection through camera pointing with *MusA*, we also shared a picture of a painting ("Cena in Emmaus") with them, to be displayed on their computer and used in place of the actual artwork.

We first described to the participants the scope of our research and the goal of the evaluation, and we collected their demographic data. We then had the participants access the description of an artwork with one of the two systems: *MusA* or audio guide, and afterwards with the other. The two conditions were counterbalanced to limit effects of order.

At the end of the study, participants answered to a questionnaire about the performed task, which included System Usability Scale (SUS) [11] (see Table 3), and a series of Likert scale questions related the specific functionalities of the accessed system, (see Table 2). The participants also provided additional comments and suggestions for improvement.

## 4.3 Results

*4.3.1 Improvements to the AR Artwork Access.* The improvements to the *MusA* system affected the AR artwork access effectiveness for participants with LV (**RQ1**). In particular, regarding the ability to correctly frame the artwork with the device camera (**RQ1.1**),

[9]Testflight is iOS app testing framework: https://developer.apple.com/testflight/

**Table 2: Usefulness and Ease of use questions and results**

| Question | *P*1 | *P*2 | *P*3 | *P*4 | *P*5 | *P*6 | *P*7 |
|---|---|---|---|---|---|---|---|
| ***MusA*** | | | | | | | |
| **m1** ease of use of camera artwork selection | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| **m2** ease of use of description navigation | 5 | 4 | 5 | 5 | 4 | 5 | 5 |
| **m3** fatigue from listening to the description | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| **m4** usefulness of the overlay images | 5 | 5 | 5 | 2 | 3 | 5 | 5 |
| **m5** clarity of the overlay images | 2 | 4 | 4 | 2 | 1 | 5 | 5 |
| **m6** usefulness of chapter selection by image touch | 5 | 5 | 5 | 3 | 5 | 5 | 5 |
| **m7** clarity of chapter selection by image touch | 5 | 1 | 5 | 3 | 4 | 5 | 5 |
| **m8** usefulness of using a personal device | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| **m9** usefulness of being able to use *MusA* at home | 5 | 5 | 5 | 5 | 5 | 3 | 5 |
| **m10** overall usefulness of *MusA* | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| **m11** would visit art venues more often with *MusA* | 4 | 4 | 4 | 3 | 5 | 5 | 5 |
| **Audio guide** | | | | | | | |
| **a1** ease of use of the artwork selection by code | 5 | 5 | 5 | 5 | 5 | 2 | 2 |
| **a2** ease of use of description navigation | 5 | 5 | 4 | 5 | 3 | 4 | 5 |
| **a3** overall usefulness of the audio guide app | 5 | 5 | 2 | 5 | 5 | 4 | 2 |

the addition of vibro-haptic feedback upon the recognition of the artwork made it immediate for the participants to realize that the artwork has been detected. Indeed, all participants reported that artwork framing using camera was easy (**m1**).

Most participants (*P*2, *P*3, *P*6, *P*7) found the visual feedback that is provided through image overlays (**RQ1.2**) clear (**m5**) after the contrast improvements. However, for the participants with highest degree of visual impairment, who use solely VoiceOver to access their mobile device (*P*1, *P*4, *P*5), the contrast was still not sufficient. Indeed, *P*4 and *P*5, due to the gravity of their visual impairment, were the only participants that did not consider image overlays very useful (**m4**), while *P*1 still considered it useful but would prefer further clarity adjustments.

The zoom functionality was tested by a few participants and it was appreciated. In particular *P*7 requested to add this functionality also to the AR mode. Finally, considering the improved interaction with the AR chapter selection (**RQ1.3**), participants did not report any interaction issues, and the functionality was considered useful (**m6**) and clear (**m7**) by most participants. Only *P*2 perceived it as unclear due to the small touch area.

*4.3.2 System Usability Compared to Audio Guide Baseline.* SUS results, comparing the usability of *MusA* to the audio guide baseline (**RQ2.1**), are shown in Figure 2 and reported in details in Table 3. Overall, SUS reported very high scores for the audio guide baseline (average score of 95). This was expected because the audio guide app provides minimal functionalities and only few interactions, which are well known and easy to perform. However, with respect to this upper baseline, *MusA* performed well, reaching an average score of 92. Both scores are considered "Excellent" based on the analysis of prior collections of SUS surveys [9].

In particular, we notice that *MusA* received higher scores for the SUS question 1, suggesting that participants would use it more frequently, and for question 9, highlighting that the participants felt more confident with *MusA* (see Figure 2). These scores are also confirmed by the answers to question **m11** (see Table 2), indicating that *MusA* would stimulate the users to visit museums more frequently.

Detailed answers to the additional questions are reported in Table 2. Overall, *MusA* was found to be useful (see **m10**), and participants appreciated the ability to use their own mobile device (**m8**) and use *MusA* at home (**m9**). Audio guide was also considered useful (**a3**), but somewhat less. Indeed, *P*3, who gave a lower score for audio guide further explained:

> *"(Compared to audio guide) the freedom of exploration is much higher in MusA."*

*P*7 further explained that, without someone to provide the number of an artwork to insert in the audio guide, its usefulness is lower:

> *"It is OK if there is someone to give me the number. Otherwise how can I access the painting? I'd say* 2*."*

*4.3.3 Usefulness and Ease of Use of MusA Functionalities.* We assess the usefulness and the ease of use of *MusA* functionalities (**RQ1.2**), comparing them, when appropriate, with the corresponding audio guide functionalities. In both conditions, selecting the artwork (**m1** and **a1**) was considered easy. Ease of use of description navigation functionalities (**m2** and **a2**) scored mildly better for *MusA* than for the audio guide app.

During the study we also noticed that the participants did not experience fatigue as in the preliminary study (**m3**). This might be due to the introduction of the virtual mode, as there was no need to keep the phone constantly pointing towards the artwork, but we were not able to assess this intuition due to the fact that the study was conducted telephonically.

## 5 DISCUSSION

We discuss the key findings and limitations of our work.

**Table 3: SUS questions and results, per participant, comparing *MusA* with the audio guide**

| Question | *P*1 | *P*2 | *P*3 | *P*4 | *P*5 | *P*6 | *P*7 | *P*1 | *P*2 | *P*3 | *P*4 | *P*5 | *P*6 | *P*7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***MusA*** | | | | | | | ***Audio guide*** | | | | | | |
| **s1** I think that I would like to use this system frequently | 5 | 5 | 5 | 3 | 5 | 4 | 5 | 5 | 5 | 5 | 3 | 5 | 2 | 4 |
| **s2** I found the system unnecessarily complex | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| **s3** I thought the system was easy to use | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 |
| **s4** I would need support of a technical person to be able to use this system | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **s5** I found the various functions in this system were well integrated | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 3 | 5 |
| **s6** I thought there was too much inconsistency in this system | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 |
| **s7** I immagine that most people would learn to use this system very quickly | 4 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 5 |
| **s8** I found the system very cumbersome to use | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| **s9** I felt very confident using the system | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 4 | 5 | 4 | 5 |
| **s10** I needed to learn a lot of things before I could get going with this system | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 |

## 5.1 Limitations of the User Study

While we initially planned to conduct the evaluation in a controlled environment (in the museum), due to the 2020 COVID-19 lockdown the second iteration user study was conducted remotely. This posed some additional challenges. First of all, since we administered the test by telephone, we had not been able to observe the users. Indeed, during the preliminary evaluation the direct observation provided several hints, including the fact that users were getting tired pointing towards the painting. Instead, it was more difficult to perceive the difficulties faced by the users over the phone. A possible approach to mitigate this issue is the use of videoconferencing software to observe the participants during the study. However, this setup may be harder to implement for the participants and may also raise privacy concerns. An alternative approach could rely on remote usage data collection [1], with a follow-up analysis on such data in order to identify difficult interactions.

Another challenge is the participants' selection: since we conducted the test with the users' own devices and the two apps (*MusA* and the audio guide) have been developed for the latest version of *iOS*, users who did not have access to a compatible device had to be excluded from the test. This happened for 7 out of the 14 candidate participants. Also, among the 7 participants there were differences in terms of the device used for running *MusA* (five participants used an iPhone, while two used an iPad). To address this issue, the researchers can ensure that the prototype software runs on diverse devices and platforms, for example by using cross-platform development frameworks [28].

In addition to the challenges of conducting the test remotely, there are limitations due to the fact that the experiment is conducted at home and not at the museum. First of all, we did not use a real audio guide. On the one hand, the simulated audio guide we used can be harder to use than the physical one, which have buttons in relief (this was also observed by *P*2). On the other hand, the fact that the simulated audio guide runs on the users' device can result in the user perceiving it as easier to use [23].

A second problem is that the different environment can impact on users' behaviour and needs. For example, since at home users are sitting at the desk, no fatigue was reported in holding the same position for a long time. Similarly, since the user knows where their computer is, the framing process is simpler. In our opinion, simulating an experimental setting similar to the real world scenario might not be possible. The study design should therefore focus on those aspects of the interaction that are not impacted by the different experimental setting.

## 5.2 Participants' comments and suggestions

During the study we collected a number of comments and suggestions. In particular *P*1 suggested to allow the user to personalize the text-to-speech voice speed. This is technically simple and we intend to implement this suggestion in the next iteration. Instead, *P*5 suggests to add the functionality to move back/forward in the audio description (*e.g.*, 10 seconds). This could be helpful, for example, if the user is distracted for a short time and wants to listen again a few seconds of description. We will take this functionality into account in the next versions, also balancing the number of buttons available in the interface with their actual usefulness. *P*7 suggested that, for longer descriptions, it would be useful to turn off the screen while listening.

*P*7 also suggested the possibility to insert personal comments and annotations, as well as to prepare descriptions of custom images and share them with friends.

> *"I would like to be able to record my thoughts or annotations while listening to the descriptions, for example to search more information on the author or the subject."*

Other participants asked for image filters, to tune luminosity, contrast and color (*P*5) or reported that the painting used for the test was too dark. For example, *P*1 commented:

> *"The painting is indeed quite dark"*

To support users with LV in accessing artwork visually, even when they are poorly illuminated or contrasted, we are currently working on tunable image filters, which will be included in the next iteration of the app.

## 5.3 Results' Generalizability

Our research provides some insights on the more general problem of usability of AR systems by people with LV. Indeed, our research takes into account three of the most important AR features: framing a target object with the camera, accessing visual feedback, and touchscreen interaction.

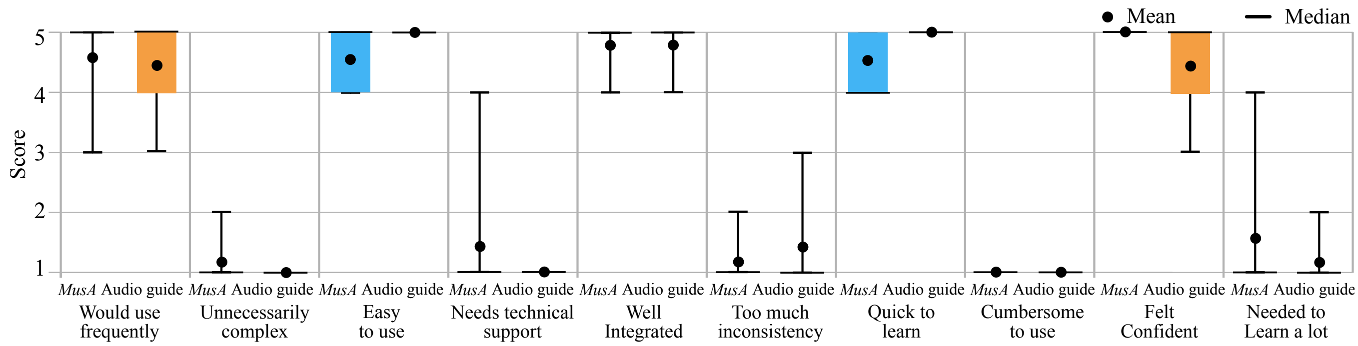



**Figure 2: Aggregated SUS results comparing *MusA* with the audio guide**

For the first feature, our contribution shows that, if the user is in the right position, framing a clearly visible target is simple, provided that a clear feedback is provided once the target is recognized. In the more general case in which the user is not in the right position, a navigation system is needed; since AR-based navigation systems have already been proposed in the literature, an integration of the two systems is possible, to provide, through AR, both navigation and description of elements in the physical world.

The results of our experiments also suggests that both device orientations should be supported and that a fatigue effect emerges if users are required to hold the position for a long time. While we observed these phenomena with participant with LV, we suspect that the same also holds for sighted users.

Considering the second feature (*i.e.*, visual feedback), our research suggests that users believe that this type of feedback is useful (see question **m4** in Table 2). However, designing clearly visible graphics is a challenge. Indeed, following the comments in the first iteration, in the second iteration we devoted an effort in designed clearly visible contours but 3 out of 7 participants still found them not clear enough (see question **m5** in Table 2). We believe that customization (according to users' needs) and adaptation (*e.g.*, to the background image) can help improving visibility. Despite these limitations, this result suggests that other AR applications are possible for people with LV. For example, we envision techniques to highlight objects in the real word (*e.g.*, a switch) with virtual graphical information.

Finally, considering the interaction with graphical information shown on the touchscreen, we initially expected this functionality to raise problems to the users due to the fact that visual information on the display can be hard to perceive. During the first iteration, we indeed observed problems with this functionality but we conjectured that they were related to the app design rather than the interaction itself. This conjecture was confirmed during the second iteration in which, using the same interaction but a different app design, no problem emerged with this functionality.

We observe that the interaction with the touchscreen is a functionality that is common in many AR applications but has been disregarded in the scientific literature for people with LV. One reason is that existing contributions mainly address the problem of way finding, so the focus is more on conveying information to the users, rather than interacting with them.

## 6 CONCLUSION AND FUTURE WORK

The problem of making museums accessible to people with disabilities can be addressed from different perspectives: increasing the building physical accessibility, providing accessible online information that is available before the visit takes place, navigating in the building, creating suitable descriptions and making them accessible to the end-users. This paper addresses the last problem and presents a solution, designed with a user-centric approach, targeted to people with LV. The proposed system relies on modern technologies, in particular mobile devices and AR, to fulfill the requirements identified during the analysis phase. The two rounds of evaluation provided valuable feedback to improve the app, and confirmed that *MusA* can indeed be an effective solution to support museum visits and to spur people with LV to visit museums.

This contribution also provide some insights on the usability of AR systems by people with LV. In particular our results confirm previous findings about the fact that visual information presented in AR can be useful for people with LV. In addition to this, our research suggests that two additional functionalities, typical of AR systems, can be accessible to people with LV: framing a target with the camera and interacting with physical objects through the touchscreen.

As a future work, from the point of view of technical development, we plan to make *MusA* available to iOS and Android users and to include all *DescriVedendo* descriptions that are currently available for artworks exposed in 8 museums. This will require a third iteration with the users, in particular to address the comments received during the second iteration. Once the app is published, we intend to remotely collect usage data, to better understand user behaviour, preferences and possibly identify accessibility problems.

There are several open research directions. First, we intend to address the problem of making *MusA* accessible to people with other forms of disabilities, including blind people. We expect the existing *DescriVedendo* descriptions to be suitable for blind users, but it might be challenging to provide accessible image interaction. Another interesting research direction is about the accessibility of three-dimensional artworks, (*e.g.*, statues and buildings). In this case the user should also be guided to move with respect to the artwork, and this can be challenging.

# REFERENCES


[1] Dragan Ahmetovic, Cristian Bernareggi, Mattia Ducci, Andrea Gerino, and Sergio Mascetti. 2021. Remote Usage Data Collection and Analysis for Mobile Accessibility Applications. In *International Conference on Pervasive Computing and Communications (PerCom) - Mobile and Pervasive Assistive Technologies Workshop (MPAT)*. IEEE.

[2] Dragan Ahmetovic, Nahyun Kwon, Uran Oh, Cristian Bernareggi, and Sergio Mascetti. 2021. Touch Screen Exploration of Visual Artwork for Blind People. In *The Web Conference (WWW)*. ACM.

[3] Dragan Ahmetovic, Daisuke Sato, Uran Oh, Tatsuya Ishihara, Kris Kitani, and Chieko Asakawa. 2020. ReCog: Supporting Blind People in Recognizing Personal Objects. In *ACM Human Factors in Computing Systems (CHI)*. ACM.

[4] Giorgos Anagnostakis, Michalis Antoniou, Elena Kardamitsi, Thodoris Sachinidis, Panayiotis Koutsabasis, Modestos Stavrakis, Spyros Vosinakis, and Dimitris Zissis. 2016. Accessible museum collections for the visually impaired: combining tactile exploration, audio descriptions and mobile gestures. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 1021–1025.

[5] Anastasios Nikolas Angelopoulos, Hossein Ameri, Debbie Mitra, and Mark Humayun. 2019. Enhanced Depth Navigation Through Augmented Reality Depth Mapping in Patients with Low Vision. *Scientific reports* 9, 1 (2019), 1–10.

[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] 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*. 382–384.

[8] 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 *International Cross-Disciplinary Conference on Web Accessibility (W4A)*.

[9] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[10] Sandra Birnstiel, Benedikt Steinmüller, Kerstin Bissinger, Simone Doll-Gerstendörfer, and Stephan Huber. 2019. Gartenfreund: Exploring the Botanical Garden with an Inclusive App. In *Proceedings of Mensch und Computer 2019*. 499–502.

[11] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

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

[13] Amanda Cachia. 2013. Talking blind: disability, access, and the discursive turn. *Disability Studies Quarterly* 33, 3 (2013).

[14] Paul Föckler, Thomas Zeidler, Benjamin Brombach, Erich Bruns, and Oliver Bimber. 2005. PhoneGuide: museum guidance supported by on-device object recognition on mobile phones. In *Proceedings of the 4th international conference on Mobile and ubiquitous multimedia*. 3–10.

[15] 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*. 3417–3422.

[16] Giuseppe Ghiani, Barbara Leporini, and Fabio Paternò. 2008. Vibrotactile feedback as an orientation aid for blind users of mobile guides. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. 431–434.

[17] Barry Ginley. 2013. Museums: A whole new world for visually impaired people. *Disability Studies Quarterly* 33, 3 (2013).

[18] 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.

[19] 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.

[20] Phoebe Hillemann. 2016. Art Museum Tours For Visitors Who Are Blind-Smithsonian American Art Museum.(2016).

[21] Jonathan Huang, Max Kinateder, Matt J Dunn, Wojciech Jarosz, Xing-Dong Yang, and Emily A Cooper. 2019. An augmented reality sign-reading assistant for users with reduced vision. *PloS one* 14, 1 (2019), e0210630.

[22] 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*. 273–274.

[23] Anna Jankowska, Agnieszka Szarkowska, Krzysztof Krejtz, Anita Fidyka, Jaroslaw Kowalski, and Marcin Wichrowski. 2017. Smartphone app as a museum guide. Testing the Open Art application with blind, deaf, and sighted users. *Rivista internazionale di tecnica della traduzione* 19 (2017), 113–130.

[24] Gunnar Jansson, Massimo Bergamasco, and Antonio Frisoli. 2003. A new option for the visually impaired to experience 3D art at museums: manual exploration of virtual copies. *Visual Impairment Research* 5, 1 (2003), 1–12.

[25] Catalina Jiménez Hurtado. 2015. Museum Accessibility through Translation: A Corpus Study of Pictorial Audio Description. *Audiovisual translation: Taking stock* (2015).

[26] 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. *Virtual Reality* 16, 4 (2012), 253–269.

[27] Steven Landau, William Wiener, Koorosh Naghshineh, and Ellen Giusti. 2005. Creating accessible science museums with user-activated environmental audio beacons (Ping!). *Assistive Technology* 17, 2 (2005), 133–143.

[28] Sergio Mascetti, Mattia Ducci, Niccoló Cantù, Paolo Pecis, and Dragan Ahmetovic. 2020. Developing Accessible Mobile Applications with Cross-Platform Development Frameworks. *arXiv preprint arXiv:2005.06875* (2020).

[29] Rebecca McMillen. 2012. The Inclusive Art Museum: Determining Disability Access. *International Journal of the Inclusive Museum* 4, 1 (2012).

[30] Apostolos Meliones and Demetrios Sampson. 2018. Blind MuseumTourer: A system for self-guided tours in museums and blind indoor navigation. *Technologies* 6, 1 (2018), 4.

[31] United Nations Convention on the Rights of People with Disabilities. Accessed: 2020-05-01. https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html

[32] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. 2019. WatchOut: Obstacle sonification for people with visual impairment or blindness. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 402–413.

[33] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–21.

[34] Carmen Santoro, Fabio Paterno, Giulia Ricci, and Barbara Leporini. 2007. A multimodal mobile museum guide for all. *Mobile Interaction with the Real World (MIRW 2007)* (2007), 21–25.

[35] Lee Stearns, Victor DeSouza, Jessica Yin, Leah Findlater, and Jon E Froehlich. 2017. Augmented reality magnification for low vision users with the microsoft hololens and a finger-worn camera. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. 361–362.

[36] E. J. Toledo, J. J. Martinez, E. J. Garrigos, and J. M. Ferrandez. 2005. FPGA implementation of an augmented reality application for visually impaired people. In *International Conference on Field Programmable Logic and Applications, 2005.* 723–724.

[37] Fernando Vargas-Martin, Eli Peli, et al. 2002. Augmented-view for restricted visual field: multiple device implementations. *Optometry and Vision Science* 79, 11 (2002), 715–723.

[38] 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.

[39] Americans with Disabilities Act. Accessed: 2020-05-01. www.ada.gov

[40] Yuhang Zhao, Michele Hu, Shafeka Hashash, and Shiri Azenkot. 2017. Understanding Low Vision People's Visual Perception on Commercial Augmented Reality Glasses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 4170–4181.

[41] 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*. 1–14.