Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

Article

Image Accessibility for Screen Reader Users: A Systematic Review And A Road Map

Uran Oh [1]*, Hwayeon Joh [1], and YunJung Lee [1]

[1] Department of Computer Science and Engineering, Ewha Womans University, Seoul, South Korea; e.jhy@ewhain.net; yunjung_lee20@ewhain.net

* Correspondence: uran.oh@ewha.ac.kr

Abstract: A number of studies have been conducted to improve the accessibility of images using touchscreen devices for screen reader users. In this study, we conducted a systematic review of 33 papers to get a holistic understanding of existing approaches and to suggest a research road map given identified gaps. As a result, we identified types of images, visual information, input device and feedback modalities that were studied for improving image accessibility using touchscreen devices. Findings also revealed that little has studied how to automate the generation of image-related information, and that screen reader users play important roles during the evaluation but the design process. Then we introduce two of our recent studies on the accessibility of artwork and comics, AccessArt and AccessComics respectively. Based on the identified key challenges, we suggest a research agenda for improving image accessibility for screen reader users.

Keywords: Image accessibility; touchscreen; nonvisual feedback; blind; visual impairment; systematic review

1. Introduction

A number of studies have been conducted to improve the accessibility of images for people who are blind or have low vision (BLV) using custom-made tactile versions of images [1–7]. Cavazos *et al.* [4], for instance, proposed a 2.5D tactile representation of the artwork where blind users can feel the artwork by touch while listening to audio feedback. Holloway *et al.* [6] also investigated tactile graphics and 3D models to deliver map information such as the number of entrances, location and direction of certain landmarks. While these approaches with extra tactile feedback are found to be effective as it conveys spatial understanding by touch [8–10], it requires additional equipment (*e.g.*, 3D printer, custom devices), whose access is limited. Moreover, it often needs to be built for individual images, which is not ideal for supporting a large number of various images in terms of the cost and time.

Meanwhile, others relied on digital devices that are commercially available (*e.g.*, PC, tablets, smartphones) for conveying image descriptions (also known as alternative text or alt text) on the web in particular [11–13]. For instance, Zhong *et al.* [12] generated alt text for images on the web that are identified as important using crowdsourcing. On the other hand, Stangl *et al.* [11] used natural language processing and computer vision techniques to automatically extract visual descriptions (alt text) on online shopping websites for clothes. Unlike tactile approaches, this software-based approaches are more scalable especially with the help of crowds or advanced machine learning techniques. However, listening to a set of verbal descriptions of an image may not be sufficient for understanding its spatial layout of content or objects within each image.

To leverage these issue of two different approaches above, researchers have worked on touchscreen-based image accessibility that enables users to explore different regions on images by touch to help them have a better spatial understanding. In this paper, to gain a more holistic perspective of this approach by examining the current states and

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

identifying the challenges to be solved, we conducted a systematic literature review of 33 papers, following PRISMA guidelines [14]. To be specific, we were interested in identifying the followings: supported image types, provided information, how the information was collected and delivered, participation of BLV people.

As a result, we found that researches on touchscreen-based image accessibility have been mostly focused on maps (*e.g.*, directions, distance), graphs (*e.g.*, graph type, values) and geometric shapes (*e.g.*, shape, size, length) using audio and haptic feedback. Moreover, it revealed that the majority of them manually generated image-related information or assumed that the information is given. We also confirmed that while most user studies are conducted with participants who are blind or have low vision for user evaluation, little has involved the target users during the design process.

To demonstrate how other types of images can be made accessible using touchscreen devices, we then introduce two of our previous systems; AccessArt [15–17] for artwork and AccessComics [18] for digital comics.

Based on the challenges and limitations identified by conducting systematic review and from our own experience of improving image accessibility for screen reader users, we suggest a road map for future studies in this field of research. The contributions of this research are as follow:

- A systematic review of touchscreen-based image accessibility for screen reader users.
- A summary of the systematic review in terms of image type, information type, methods for collecting and delivering information, the involvement screen reader users.
- The identifications of key challenges and limitations.
- Recommendations for future research directions.

2. Related Work

Our work is inspired by prior work on image accessibility and touchscreen accessibility for BLV people.

2.1. Image Accessibility

Screen readers cannot describe an image unless its metadata is present such as alt text. To improve the accessibility of images, various solutions have been proposed to provide accurate descriptions for individual images on the web or on mobile devices [11–13,19–21]. Winters *et al.*, for instance, [13] proposed an auditory display for social media which can automatically detect the overall mood of an image and gender and emotion of any faces using Microsoft's computer vision and optical character recognition (OCR) APIs. Similarly, Stangl *et al.* [11] developed computer vision (CV) and national language processing (NLP) modules to extract information about clothing images on online shopping mall. To be specific, The CV module automatically generates a description of the entire outfit shown in a product image while the NLP module is responsible for extracting price, material, and description from the web page. Goncu and Marriott [21], on the other hand, demonstrated the idea of creating accessible images by the general public using a web-based tool. In addition, Morris *et al.* [19] proposed a mobile interface that provides screen reader users with rich information of visual contents prepared using real-time crowdsourcing and friend-sourcing rather than using machine learning techniques. It allowed users to listen to the alt text of a photograph and ask questions using voice input while touching specific regions with their fingers.

While most of the studies for improving the accessibility of images that can be accessed with digital devices tend to focus on how to collect the metadata that can be read out to screen reader users, others investigated how to deliver image-related information with tactile feedback [1,3,4,22,23]. Götzelmann *et al.* [22], for instance, presented a 3D-printed map to convey geographic information by touch. Yet, some worked on using computational methods to automatically generate tactile representations [2,24]. For example, Rodrigues *et al.* [24] proposes an algorithm for creating tactile representations of objects of an artwork varying shape and depth. While promising, we focused on improving image accessibility on a touchscreen which is widely adopted to personal devices such as smartphones and tablets since it does not require additional hardware.

2.2. Touchscreen Accessibility

While we chose to focus on touchscreen-based image accessibility, touchscreen devices are innately inaccessible as it requires accurate hand-eye coordination [25]. Thus, various studies have been conducted to improve touchscreen accessibility by providing tactile feedback using additional hardware devices [26–28]. *TouchCam*, for example, designed and implemented a camera-based wearable device that can be worn on a finger, which is used to access own's personal touchscreen devices by interacting with their skin surface to provide extra tactile and proprioceptive feedback. Physical overlays that can be placed on the top of a touchscreen were also investigated [5,29]. For instance, *TouchPlates* [5] allows people with visual impairments to interact with touchscreen devices by placing tactile overlays on the top of the touch display. Meanwhile, software-based approaches have been proposed as well such as supporting touchscreen gestures that can be performed anywhere on the screen [25,30–33]. *BrailleTouch* [31] and No-Look Notes [33], for example, proposed software solutions for supporting eyes-free text entry for blind users by using multi-touch gestures. Similarly, smartphones on the market also offer screen reader modules with location-insensitive gestures: iOS's VoiceOver[1] and Android's Talkback[2]. These screen readers reads out the contents on the screen if focused, and users can navigate different items by directional swipes (*i.e.*, left-to-right and right-to-left swipe gestures) or by exploration-by-touch.

Again, we are interested in how touchscreen devices can be used to improve image accessibility mainly because it is readily available to a large of end-users even BLV people as they have their own personal devices with touchscreens. In addition, as touchscreen devices offer screen reader functionality, it is accessible.

3. Method

To identify the road map of future research directions on image accessibility for people with visual impairments using touchscreen devices, we conducted a systematic review following PRISMA guidelines [34]. The process is shown in Figure 1.

---

1 https://www.apple.com/accessibility/vision/

2 https://support.google.com/accessibility/android/answer/6283677?hl=en

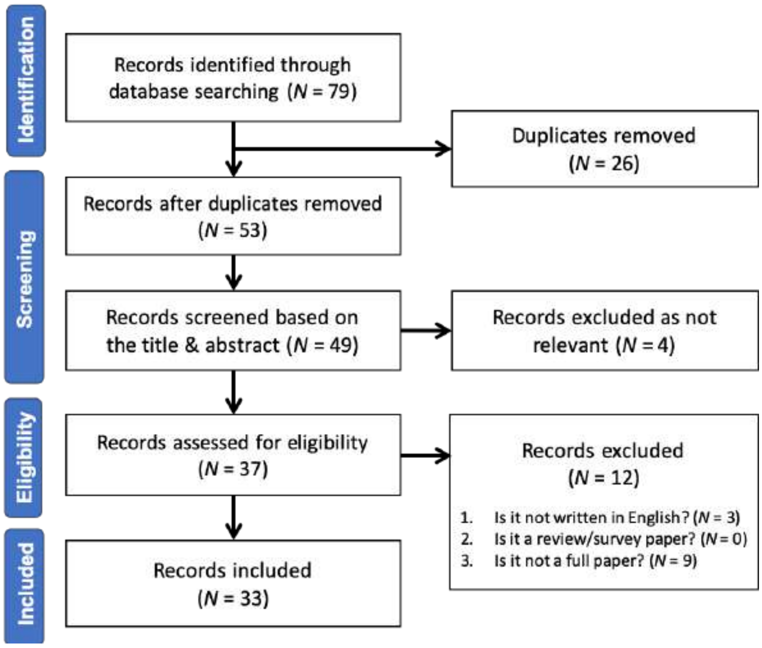


Figure 1. A PRISMA flow diagram that shows the process of identifying eligible papers on touchscreen-based image accessibility for people who are blind or have visual impairments.

3.1. Research Questions

We had five specific research questions for this systematic review:

- *RQ1. What types of images have been studied for image accessibility?*
- *RQ2. What types of image-related information have been supported for BLV?*
- *RQ3. How has image-related information collected?*
- *RQ4. How has image-related information delivered?*
- *RQ5. How have BLV involved in the design and evaluation process?*

3.2. Identification

To identify research papers related to touchscreen-based image accessibility for BLV, we checked if at least one of the following search keywords from each category — target user (*User*), target object (*Object*), supported feature (*Feature*), supported device (*Device*) — is included either in the title, the abstract or authors' keywords:

- *User: "blind", "visual impairment", "visually impaired", "low vision", "vision loss"*
- *Object: "image", "picture", "photo", "figure", "drawing", "painting", "graphic", "map", "diagram"*
- *Feature: "description", "feedback"*
- *Device: "touchscreens", "touch screens"*

As a result, a total of 79 papers were identified from three databases: Scopus (N = 50), ACM digital library (N = 25) and IEEE Xplore (N = 4). Then we removed 26 duplicates.

3.3. Screening

Then we examined the titles and abstracts of the rest 53 unique papers and excluded four papers that are not relevant, which were all conference reviews.

3.4. Eligibility

Of 49 remaining papers, we excluded 12 papers that met the following exclusion criteria:

1. Not written in English (N = 3).
2. A survey or review paper (N = 0).

3. Not a full paper such as posters, workshop and case study papers (*N* = 9).

Then papers were included if and only if the goal of the paper is improving the accessibility of any type of images for people with visual impairments.

4. Results

As the result of the systematic review, 33 papers were considered as eligible for the analysis. We summarize the papers mainly in terms of the research questions specified in subsection 3.1.

4.1. Overview

As shown in Figure 2, the first paper was published in 2008; note that the first generation of iPhone was released in 2007. While the topic of this area is not that active, at least three papers have been published each year constantly since 2013. As for the country of authors' affiliations, United States had the dominant number of papers, which was 17. It followed by Germany (*N* = 4), Canada (*N* = 3), Australia (*N* = 2) and Japan (*N* = 2). Others countries that had a single paper were Austria, Brazil, China, France, Italy, Lebanon, South Korea and Spain.

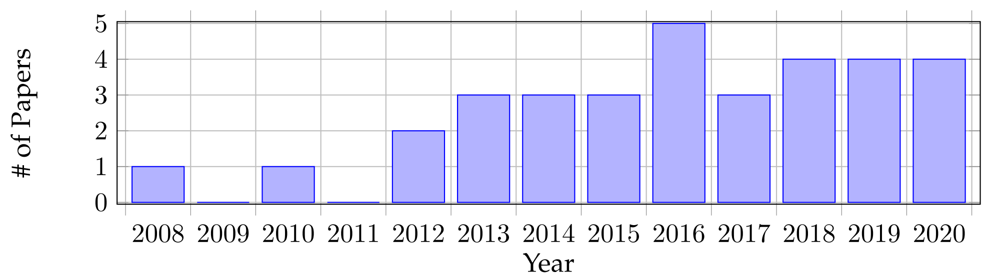


Figure 2. The number papers related to touchscreen-based image accessibility.

4.2. Supported Image Types

The types of images that have been studied in prior works for providing better accessibility for screen reader users (*RQ1*) are summarized in Table 1. While three papers were designed to support any type of images in general [35–37], most of the papers focused on specific types of images. For instance, images of maps were investigated the most (*N* = 10), followed by graphs (*N* = 6). Interestingly, while the accessibility of photographs for BLV were largely investigated in terms of web accessibility [20,38], only three out of 33 papers aimed to support photographs particularly using touchscreen devices. In addition, as touchscreen devices themselves have accessibility issues for people with visual impairments requiring accurate hand-eye coordination [25], four papers focused on improving the accessibility of the touchscreen-based interface itself such as soft buttons [39–41] and gestures [42].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

Image Type	Count	Papers
Map	10	[22,23,37,43–49]
Graph	6	[37,50–54]
Geometric Shapes	4	[50,52,55,56]
Images in General	3	[35–37]
Lines	3	[57–59]
Photographs	3	[37,60,61]
Touchscreen UI	3	[39–41]
Floor plan	3	[62–64]
Diagram	2	[37,65]
Artwork	1	[66]
Touchscreen Gestures	1	[42]

Table 1. Types of supported images for BLV.

4.3. Information Type

We have identified types of information provided to improve image accessibility (*RQ2*), which is shown in Table 2. The name of the object, mainly the object that is touched, was provided the most ($N = 10$), followed by spatial layout of various objects in each image ($N = 8$). As expected, types of provided information differ depending on the image type. For instance, direction/orientation, distance and other geographic information were provided for map images. On the other hand, shape, size, boundary of objects, length were mostly offered for images related to geometric figures. Scene descriptions ($N = 5$), textures ($N = 4$), graph values ($N = 2$) and texts written in images ($N = 2$). Note that only two papers supported color information for photographs [35] and artwork [66]. Other information types include types of graph [53] and weather [23]. Five papers did not specified the information they provide.

Info Type	Count	Papers
Object Names	10	[22,23,43,46,49,55,62–64,66]
Spatial Layout	8	[46,47,49,54,55,60–62]
Direction/Orientation	8	[43–45,47,51,59,61,65]
Distance	5	[43,44,46,47,49]
Geographic information	5	[22,46–49]
Scene	5	[36,37,53,63,66]
Shape	4	[50–52,55]
Texture	4	[35,40,56,66]
UI Element	3	[40,41,61]
Area/Size	2	[56,59]
Object Boundary	2	[35,58]
Stroke	2	[42,54]
Length	2	[51,65]
Values in Graphs	2	[51,53]
Texts in Images	2	[53,61]
Color	2	[35,66]
Not specified	5	[39,50,55,57,64]

Table 2. Types of information provided to improve image accessibility.

4.4. Information Preparation

In addition to types of information supported to improve image accessibility (*RQ3*). As shown in Table 3, we have found that most studies have not specified how the visual information is collected or created. This suggests that the aim of many of these studies is "delivering" visual information that is inaccessible to BLV as is while assuming that the information is given rather than "retrieving" the information. Meanwhile, close to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

one-third of the studies seem to manually created the data they need to provide (*N* = 9), or use metadata such as alternative text (alt text) or textual descriptions that are paired with the images (*N* = 4). Others relied on automatic approaches to extract information of interests from images such as optical character recognition (OCR) and computer vision (*N* = 5). Meanwhile, two papers proposed a system where the image descriptions are provided by crowdworkers.

Data Collection	Count	Papers
Manually Created	9	[39,40,45,53,54,56,58,59,66]
Metadata (*e.g.*, alt text)	4	[41,43,64,66]
Image Processing	2	[61,62]
Machine Learning (*e.g.*, computer vision)	2	[35,36]
Crowdsourcing	2	[60,61]
OCR	1	[61]
Not specified	18	[22,23,37,41,42,44–52,55,57,63,65]

Table 3. Data preparation/collection methods for providing image-related information.

4.5. Interaction Types

We were also interested in how image-related information is delivered using touch-screen devices (*RQ4*). We have identified the interaction type in terms of input types and output modalities as follows:

Input types. As shown in Table 4, the major input type is touch as expected; most of the studies allowed users to explore images by touch with their bare hands (*N* = 28 out of 33). Moreover, touchscreen gestures were also used as input (*N* = 5). On the other hand, physical input devices (*N* = 4 for *keyboard*, *N* = 3 for *stylus*, and *N* = 2 for *mouse*) were used in addition to touchscreen devices. While it is known that aiming a camera towards a target direction is difficult for BLV [67], a camera was also used as a type of input where users are allowed to share image feeds from cameras with others so that they get information about their surrounding physical objects such as touch panels on a microwave [60,61].

Input	Count	Papers
Touch	28	[23,35–37,39,42–46,48–59,61–66]
Gesture	5	[37,42,46,47,49]
Keyboard	4	[23,35–37]
Stylus	3	[37,40,61]
Mouse	2	[35,37]
Voice command	2	[46,49]
Camera	2	[60,61]
Physical UI	1	[41]

Table 4. Types of input used for improving images on touchscreen devices.

Output Modalities. As for output modalities, various audio and haptic feedback techniques were used (see Table 5); audio (*N* = 25), haptic (*N* = 19) and multimodal feedback with both audio and haptic (*N* = 12). The mostly widely used output was speech feedback that verbally describes images to BLV users using an audio channel as a screen reader reads out what is on the screen using text-to-speech (*e.g.*, Apple's VoiceOver). On the other hand, non-speech audio feedback (*e.g.*, sonification) was also used. For instance, different pitch of a sound [23,35,36,42] or rhythm [55] were used to convey image-related information. Meanwhile, vibration was as popular as non-speech audio feedback while some used tactile feedback to convey information. For example, Gotzelmann *et al.* [22] used 3D-printed tactile map. Zhang *et al.* [41] also made user interface elements (*e.g.*, buttons, sliders) with a 3D printer to improve the accessibility of touchscreen-based interfaces in general by replacing virtual elements on a touchscreen with physical ones. Moreover, Hausberger *et al.* [56] proposed an interesting approach using kinesthetic feedback along with frictions. Their system dynamically changes the position and the orientation of a touchscreen device in a 3D space for BLV to explore shapes and textures of images on a touchscreen device.

Output	Count	Papers
Speech	19	[23,36,41,42,45–49,51,53,54,58,60–63,65,66]
Non-speech Audio	14	[22,23,23,35–37,42,43,48,51,52,55,55,65]
Vibration	14	[39,40,43–45,47,50–55,57,59]
Tactile Feedback	8	[22,23,39,41,44,47,56,63]
Force Feedback	3	[39,40,56]

Table 5. Output modalities used for improving images on touchscreen devices.

4.6. Involvement of BLV

Finally, we checked if BLV, the target users, were involved in the system development and evaluation processes. We first examined if user evaluation was conducted regardless of whether target users are involved or not. As a result, we found that all studies but two had tested their system with human subjects. Of the remaining 31 papers, three papers did had user studies but there were no BLV participants. The rest of the 28 papers had evaluated their system with participants from the target user group. In addition to user evaluation, some papers had BLV participated in their formative survey and design process (N = 2 and N = 7, respectively).

Involvement	Count	Papers
Evaluation	28	[23,36,37,39–53,57–66]
Design	7	[23,50,61–65]
Survey	2	[23,65]
No BLV	3	[54–56]
No user study	2	[50,63]

Table 6. BLV's involvement in system design and evaluation. Note that first three rows are the ones that involved BLV.

5. AccessArt and AccessComics

Based on our systematic review results, we have confirmed that various types of images were studied to improve their accessibility for BLV people. However, most of the studies have focused on providing knowledge or information based on facts (*e.g.*, maps, graphs) to users rather than offering accessible experience that BLV users can enjoy. Thus, we focused on two types of images in particular that are rarely studied for screen reader users: artwork and comics. Here we briefly introduce two systems we have developed for each image type: *AccessArt* [15–17] and *AccessComics* [18] (see Table 7 as well).

	AccessArt	AccessComics
Image Type	2D paintings	Digital comics
Information Type	Object names, spatial layout, shape, size, color	Panel number, background, character appearance and actions
Data Preparation	Manual, crowdsourced	Manual (visual descriptions) open dataset (panel, balloon, script)
Interaction	touch/gesture input, speech output	touch/gesture input, speech output
User Study with BLV	semi-structured interview, design probe study	online survey, semi-structured interview, design probe study

Table 7. A summary of AccessArt and AccessComics following identifying factors used in our systematic review.

5.1. AccessArt for Artwork Accessibility

BLV people are interested in visiting museums and wish to know about artwork [68–70]. However, a number of accessibility issues exist when visiting and navigating inside a museum [71]. While audio guide services are in operation for some exhibition sites [72–75], it can still be difficult for BLV people to understand the spatial arrangement of objects within each painting. Tactile versions of artwork, on the other hand, allows BLV people to learn the spatial layout of objects in the scene by touch [1–4]. However, it is not feasible to make these replica for every exhibited artwork. Thus, we began to design and implement touchscreen-based artwork exploration tool called *AccessArt*.

AccessArt Ver1. The very first version of AccessArt is shown in Figure 3 where it had four paintings varying genre: landscape, portrait, abstract and still life [16]. As for the object-level labels, we segmented each object along with descriptions. Then we developed a web application which allowed BLV users to (1) select one of the four paintings they wish to explore, and (2) scan objects within each painting by touch with its corresponding verbal description including object-level information such as the name, color and position of the object, For example, if a user touches the moon on "The Starry Night", then the system reads out the following: *"Moon, shining. Its color is yellow and it's located at the top right corner"*. Users can either use swipe gestures to go through a list of objects or freely explore objects in a painting by touch to better understand objects' location within an image. In addition, users could also specify objects and attributes they wish to explore using filtering options. Eight participants with visual impairments were recruited for a semi-structured interview study using our prototype, and received positive feedback.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

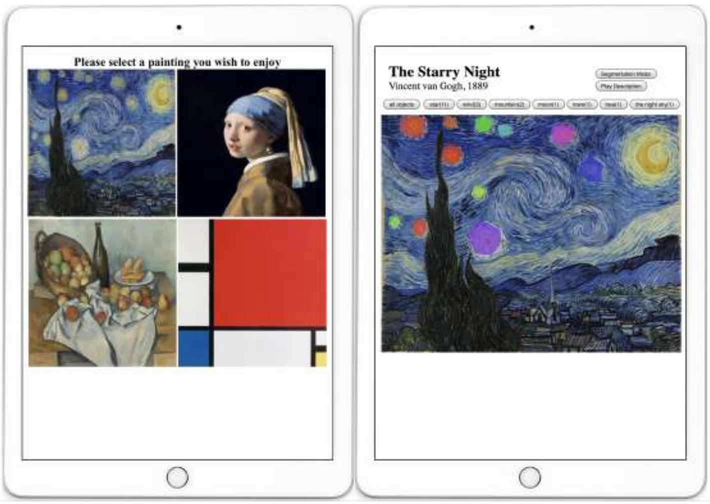


Figure 3. User interface prototype with two interaction modes: *Overview* (left) and *Exploration* (right). As for *Exploration Mode* example, *star* is selected as object of interest, highlighted in various colors.

AccessArt Ver2. The major problem with the first version of AccessArt was the object segmentation process, which was not scalable as it was all manually done by a couple of researchers. Thus, we investigated the feasibility of relying on crowdworkers who are not expected have expertise in art [15]. We used Amazon Mechanical Turk[3] for collecting object-label metadata for eight different paintings from anonymous crowd. Then we assessed the effectiveness of the descriptions generated by crowd with nine participants with visual impairments where they were asked to go through four steps of *Felman Model of Criticism* [76]: *description*, *analysis*, *interpretation*, and *judgment*). Findings showed that object-level descriptions provided by anonymous crowds were sufficient for supporting BLV's artwork appreciation.

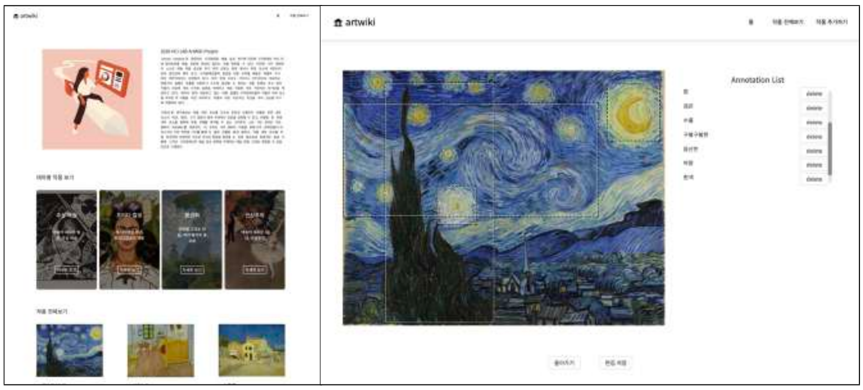


Figure 4. Screenshot examples of the main page (left) and edit page (right) of AccessArt Ver3.

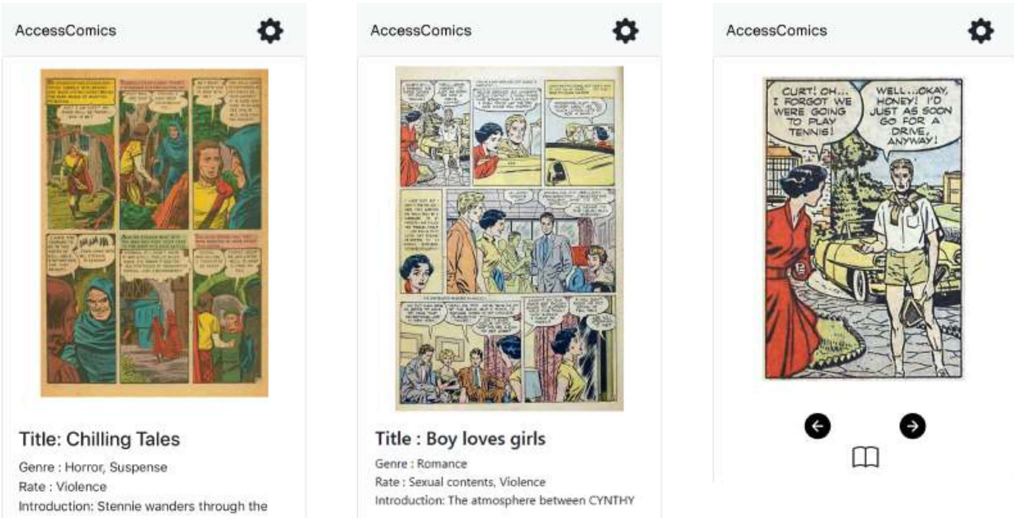
AccessArt Ver3. As a final step, we have implemented an online platform [4] as shown in Figure 4. It is designed to allow anonymous users can freely volunteer to provide object segmentation and description inspired by Wikipedia [5]. While no user evaluation is conducted with the final version yet, we expect this platform to serve as an accessible online art gallery for BLV people where the metadata are collected and maintained by crowd to support a greater number of artwork which can be accessed anywhere using one's personal device.

[3] https://www.mturk.com/
[4] https://artwiki-hci2020.vercel.app/
[5] www.wikipedia.org

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

5.2. AccessComics for Comics Accessibility

Compared to artwork accessibility, fewer number of studies have been conducted to improve the accessibility of comics. For instance, Ponsard *et al.* [77] proposed a system for people who have low vision or have motor impairments, which can automatically retrieve necessary information (*e.g.*, panel detection and ordering) from images of digital comics and reads out the content on a desktop computer controlled with a TV remote. *ALCOVE* [78] is another web-based digital comic book reader for people who have low vision. The authors conducted a user study with 11 people who have low vision and most of them preferred their system over pdf version of digital comics. Inspired by this study, our system, AccessComics [18], is designed to provide BLV users with overview (as shown in Figure 5(a)-(b)), various reading units (*i.e.*, page, strip, panel), magnifier, text-to-speech, autoplay as well. Moreover, we mapped different voices with different characters to offer high sense of immersion in addition to improved accessibility similar to how Wang *et al.* [79] used voice synthesis technique that can express emotion given scripts with the voice that suits each character. Here we briefly describe how the system is implemented.



(a) Introduction page ex1 **(b)** Introduction page ex2 **(c)** Reading page Ex

Figure 5. Screenshot examples of introduction page of two comics: (a) Chilling Tales, (b) Boy Loves Girls. (c) is an example of reading page, which plays the following information: the panel number within a page, the background such as places and spatial arrangements, visual descriptions of characters such as their appearance and action, and scripts.

Data Preparations. As for the information we wish to provide about comics, we used *eBDtheque* [80], which is a dataset of comics consisting of pairs of an image file (.jpg) and a metadata file (.svg). The metadata has segmented information of *panel*, *character*, and *balloon* of a comic page. In addition, we manually added visual descriptions such as background and appearance and actions of characters.

Interaction. Similar to *AccessArt*, *AccessComics* allow users to select a comic book they wish to read and navigate to different elements in each panel, panels themselves as well as pages and listen to displayed content. For example, as for Figure 5(c), the following would be played using the audio channel starting with the panel number, followed by background and character-related visual descriptions:

1. *[Narration Voice] "Panel One. The place is the front yard of a house and in front of a car. Cynthy is on the left, and the Curt is on the right. Cynthy turns her head and speaks to the man. She has black hair wearing a long red dress. CURT is holding a tennis racket with his left hand. His right hand is outstretched. His face is neutral. Curt has yellow hair wearing a white shirt, beige short pants and white shoes."*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

2. *[Narration Voice] "CYNTHY says," [Woman's Voice1] "CURT! Oh... I forgot we were going to play tennis!"*
3. *[Narration Voice] "Curt says," [Man's Voice1] "Well... okay, honey! I'd just as soon go for a drive, anyway!"*

6. Discussion

Here we discuss the current state of researches on touchscreen-based image accessibility and missing gaps to be investigated in the future based on the findings from systematic review and our own experience of designing two systems for artwork and comics accessibility.

6.1. From Static Images to Dynamic Images

Various types of images displayed on touchscreen devices were studied in terms of accessibility over a decade since the year of 2008. However, all but except one [42] have supported still images without motion. However, dynamically changing images such as animations and videos (*e.g.*, movies, TV programs, games, video conferences) has rarely explored in terms of accessibility for touchscreen devices for BLV users. Considering the rapid growth of YouTube [81] and its use for gaining knowledge [82], videos are another type of images (a series of images) that have various accessibility issues. While the area has been explored as well regardless of the medium [83–85], it would be interesting to examine how it can be supported for touchscreen devices.

6.2. Limited Rooms for Subjective Interpretations

As found in prior work that BLV wish to get different types of information depending on the context [20], different types of information was provided for different image types. For example, geographic information such as building locations, direction and distance were offered for map and graph images and shape, size and line-length information for geometric objects. However, little has studied about other types of images although specific location or spatial relationships of objects within each images such as photographs, touchscreen user interface, although spatial arrangements is considered important [19,29]. Moreover, the majority of the studies have been prioritized images that contains useful information (*e.g.*, facts, knowledge) over images that can be interpreted subjectively, different from one person to another such as artwork using touchscreen devices. Although many studies have taken the approach of providing tactile feedback to enable BLV people to explore, enjoy and appreciate artistic images [1–4], we recommend touchscreen-based approaches as well to make larger number of images accessible to more number of BLV people.

6.3. Automatic Retrieval of Metadata of Images for Scalability

The dominant number of studies that we have identified in our systematic review assumed that image-related information is given or manually created the information. However, a number of images on the web do not have alt text although it is recommended by Web Content Accessibility Guidelines (WCAG) [6]. Moreover, it is not feasible for a couple of researchers to generate metadata for individual images. Thus, machine learning-based automatic approaches have been studied [11,12,86]. Yet, since the accuracy is not as accurate when compared to descriptions produced by human, we recommend the crowdsourcing approach [17,19] for collecting accurate annotations. It can be served as a human-AI collaboration [12]. Eventually, the data can be used to training machine learning models for implementing fully automated image description generation system.

[6] https://www.w3.org/TR/WCAG21/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

6.4. *Limited Input and Output Modalities of Touchscreen Devices*

Unlike other assistive systems that require BLV people to physically visit certain locations (*e.g.* [75,87]) or that require special hardware devices with tactile cues (*e.g.*, [2,3]), touchscreen devices benefit from being portable where a variety of images can be accessed using a personal device with less physical and time constraints. However, the input and output modalities that a touchscreen devices can offer is limited to audio and vibration feedback. To provide more intuitive and rich feedback, studies on how to ease the designing a 2.5D, or 3D models and reduce the cost and time for producing tactile representations of images can be investigated more in depth. One way to do so is open-sourcing the process as in Instructables [7]

6.5. *Limited Involvement of BLV People During Design Process*

The findings of our systematic review revealed that most studies had user evaluation of proposed systems with BLV participants after the design and implementation. However, it is important to have target users take part in at an early state during design process when designing a new technology [88]. A formative study with surveys or semi-structured interviews is recommended to understand the current needs and challenges of BLV people before making design decisions. Iterative participatory design process is also great way to reflect BLV participants' opinions into the design especially for users with special needs [89–91].

7. Conclusions

To have complete understanding of existing approaches and identify challenges to be solved as the next step, we conducted a systematic review of 33 papers on touchscreen-based image accessibility for screen reader users. The results revealed that image types other than maps, graphs and geometric shapes such as artwork and comics are rarely studied. Also, we found that only about one-third of the papers provide multi-modal feedback of audio and haptic. Moreover, our findings show that how to collect image descriptions was out of the interests for most studies, suggesting that automatic retrievals of image-related information is one of the bottlenecks for making images accessible in a large scale. Finally, while the majority of studies did not involve people who are blind or have low vision during the system design process, future studies should consider inviting target users early in advance and reflect their comments for making design decisions.

Author Contributions: Conceptualization, U.O.; methodology, U.O.; validation, U.O; formal analysis, H.J. and Y.L.; investigation, U.O., H.J. and Y.L.; data curation, U.O., H.J. and Y.L.; writing—original draft preparation, U.O.; writing—review and editing, U.O.; visualization, U.O.; supervision, U.O.; project administration, U.O.; funding acquisition, U.O. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Science Technology and Humanity Converging Research Program of National Research Foundation of Korea (2018M3C1B6061353).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gyoshev, S.; Karastoyanov, D.; Stoimenov, N.; Cantoni, V.; Lombardi, L.; Setti, A. Exploiting a graphical Braille display for art masterpieces. International Conference on Computers Helping People with Special Needs. Springer, 2018, pp. 237–245.
2. Cantoni, V.; Lombardi, L.; Setti, A.; Gyoshev, S.; Karastoyanov, D.; Stoimenov, N. Art Masterpieces Accessibility for Blind and Visually Impaired People. International Conference on Computers Helping People with Special Needs. Springer, 2018, pp. 267–274.

[7] https://www.instructables.com/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

3. Iranzo Bartolome, J.; Cavazos Quero, L.; Kim, S.; Um, M.Y.; Cho, J. Exploring Art with a Voice Controlled Multimodal Guide for Blind People. Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction. ACM, 2019, pp. 383–390.
4. Cavazos Quero, L.; Iranzo Bartolomé, J.; Lee, S.; Han, E.; Kim, S.; Cho, J. An Interactive Multimodal Guide to Improve Art Accessibility for Blind People. Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 2018, pp. 346–348.
5. Kane, S.K.; Morris, M.R.; Wobbrock, J.O. Touchplates: low-cost tactile overlays for visually impaired touch screen users. Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, 2013, pp. 1–8.
6. Holloway, L.; Marriott, K.; Butler, M. Accessible maps for the blind: Comparing 3D printed models with tactile graphics. Proceedings of the 2018 chi conference on human factors in computing systems, 2018, pp. 1–13.
7. Panotopoulou, A.; Zhang, X.; Qiu, T.; Yang, X.D.; Whiting, E. Tactile line drawings for improved shape understanding in blind and visually impaired users. *ACM Transactions on Graphics (TOG)* **2020**, *39*, 89–1.
8. Paterson, M. 'Seeing with the hands': Blindness, touch and the Enlightenment spatial imaginary. *British Journal of Visual Impairment* **2006**, *24*, 52–59.
9. Heller, M.A. *The psychology of touch*; Psychology Press, 2013.
10. Watanabe, T.; Yamaguchi, T.; Nakagawa, M. Development of software for automatic creation of embossed graphs. International Conference on Computers for Handicapped Persons. Springer, 2012, pp. 174–181.
11. Stangl, A.J.; Kothari, E.; Jain, S.D.; Yeh, T.; Grauman, K.; Gurari, D. BrowseWithMe: An Online Clothes Shopping Assistant for People with Visual Impairments. Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 2018, pp. 107–118.
12. Zhong, Y.; Matsubara, M.; Morishima, A. Identification of Important Images for Understanding Web Pages. 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018, pp. 3568–3574.
13. Winters, R.M.; Joshi, N.; Cutrell, E.; Morris, M.R. Strategies for auditory display of Social Media. *Ergonomics in Design* **2019**, *27*, 11–15.
14. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; others. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* **2010**, *8*, 336–341.
15. Nahyun Kwon, Y.L.; Oh, U. Supporting a Crowd-powered Accessible Online Art Gallery for People with Visual Impairments: A Feasibility Study. *Universal Access and Information Society* **2021**, p. Accepted.
16. Kwon, N.; Koh, Y.; Oh, U. Supporting object-level exploration of artworks by touch for people with visual impairments. 2019, pp. 600–602. cited By 1; Conference of 21st International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2019 ; Conference Date: 28 October 2019 Through 30 October 2019; Conference Code:153194, doi:10.1145/3308561.3354620.
17. Soohyun Yeo, S.C.; Oh, U. Crowdsourcing Based Online Art Gallery for Visual Impairment People. *HCI Korea Annual Conference* **2021**, pp. 531–535.
18. Yunjung Lee, Hwayeon Joh, S.Y.; Oh, U. AccessComics : An Accessible Digital Comic Book Reader for People with Visual Impairments. Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, 2021, p. Accepted.
19. Morris, M.R.; Johnson, J.; Bennett, C.L.; Cutrell, E. Rich representations of visual content for screen reader users. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 2018, p. 59.
20. Stangl, A.; Morris, M.R.; Gurari, D. " Person, Shoes, Tree. Is the Person Naked?" What People with Vision Impairments Want in Image Descriptions. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–13.
21. Goncu, C.; Marriott, K. Creating ebooks with accessible graphics content. Proceedings of the 2015 ACM symposium on document engineering, 2015, pp. 89–92.
22. Götzelmann, T. CapMaps: Capacitive sensing 3D printed audio-tactile maps **2016**. *9759*, 146–152. cited By 5; Conference of 15th International Conference on Computers Helping People with Special Needs, ICCHP 2016 ; Conference Date: 13 July 2016 Through 15 July 2016; Conference Code:178059, doi:10.1007/978-3-319-41267-2_20.
23. Lazar, J.; Chakraborty, S.; Carroll, D.; Weir, R.; Sizemore, B.; Henderson, H. Development and Evaluation of Two Prototypes for Providing Weather Map Data to Blind Users through Sonification. *J. Usability Studies* **2013**, *8*, 93–110.
24. Rodrigues, J.B.; Ferreira, A.V.M.; Maia, I.M.O.; Junior, G.B.; de Almeida, J.D.S.; de Paiva, A.C. Image Processing of Artworks for Construction of 3D Models Accessible to the Visually Impaired. International Conference on Applied Human Factors and Ergonomics. Springer, 2018, pp. 243–253.
25. McGookin, D.; Brewster, S.; Jiang, W. Investigating touchscreen accessibility for people with visual impairments. Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges, 2008, pp. 298–307.
26. Vanderheiden, G.C. Use of audio-haptic interface techniques to allow nonvisual access to touchscreen appliances. Proceedings of the human factors and ergonomics society annual meeting. SAGE Publications Sage CA: Los Angeles, CA, 1996, Vol. 40, pp. 1266–1266.
27. Landau, S.; Wells, L. Merging tactile sensory input and audio data by means of the Talking Tactile Tablet. Proceedings of EuroHaptics, 2003, Vol. 3, pp. 414–418.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

28. Stearns, L.; Oh, U.; Findlater, L.; Froehlich, J.E. TouchCam: Realtime Recognition of Location-Specific On-Body Gestures to Support Users with Visual Impairments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2018**, *1*, 1–23.
29. Kane, S.K.; Morris, M.R.; Perkins, A.Z.; Wigdor, D.; Ladner, R.E.; Wobbrock, J.O. Access overlays: improving non-visual access to large touch screens for blind users. Proceedings of the 24th annual ACM symposium on User interface software and technology, 2011, pp. 273–282.
30. Kane, S.K.; Bigham, J.P.; Wobbrock, J.O. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, 2008, pp. 73–80.
31. Frey, B.; Southern, C.; Romero, M. Brailletouch: mobile texting for the visually impaired. International Conference on Universal Access in Human-Computer Interaction. Springer, 2011, pp. 19–25.
32. Guerreiro, T.; Lagoa, P.; Nicolau, H.; Goncalves, D.; Jorge, J.A. From Tapping to Touching: Making Touch Screens Accessible to Blind Users (vol 15, pg 48, 2008). *IEEE MULTIMEDIA* **2009**, *16*, 13–13.
33. Bonner, M.N.; Brudvik, J.T.; Abowd, G.D.; Edwards, W.K. No-look notes: Accessible eyes-free multi-touch text entry. International Conference on Pervasive Computing. Springer, 2010, pp. 409–426.
34. Al-Gobari, M.; El Khatib, C.; Pillon, F.; Gueyffier, F. Beta-blockers for the prevention of sudden cardiac death in heart failure patients: a meta-analysis of randomized controlled trials. *BMC cardiovascular disorders* **2013**, *13*, 52.
35. Toffa, O.; Mignotte, M. A Hierarchical Visual Feature-Based Approach for Image Sonification **2021**. *23*, 706–715. cited By 0, doi:10.1109/TMM.2020.2987710.
36. Banf, M.; Blanz, V. Sonification of images for the visually impaired using a multi-level approach; , 2013; pp. 162–169. cited By 9; Conference of 4th Augmented Human International Conference, AH 2013 ; Conference Date: 7 March 2013 Through 8 March 2013; Conference Code:96522, doi:10.1145/2459236.2459264.
37. Neto, J.; Freire, A.; Souto, S.; Abílio, R. Usability evaluation of a web system for spatially oriented audio descriptions of images addressed to visually impaired people **2014**. *8514 LNCS*, 154–165. cited By 0; Conference of 8th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2014 - Held as Part of 16th International Conference on Human-Computer Interaction, HCI International 2014 ; Conference Date: 22 June 2014 Through 27 June 2014; Conference Code:105561, doi:10.1007/978-3-319-07440-5_15.
38. Zhao, Y.; Wu, S.; Reynolds, L.; Azenkot, S. The Effect of Computer-Generated Descriptions on Photo-Sharing Experiences of People with Visual Impairments. *Proc. ACM Hum.-Comput. Interact.* **2017**, *1*. doi:10.1145/3134756.
39. Onishi, J.; Sakajiri, M.; Miura, T.; Ono, T. Fundamental study on tactile cognition through haptic feedback touchscreen; , 2013; pp. 4207–4212. cited By 2; Conference of 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013 ; Conference Date: 13 October 2013 Through 16 October 2013; Conference Code:102429, doi:10.1109/SMC.2013.717.
40. Kyung, K.U.; Lee, J.Y.; Park, J. Haptic stylus and empirical studies on braille, button, and texture display **2008**. *2008*. cited By 19, doi:10.1155/2008/369651.
41. Zhang, X.; Tran, T.; Sun, Y.; Culhane, I.; Jain, S.; Fogarty, J.; Mankoff, J. Interactiles: 3D printed tactile interfaces to enhance mobile touchscreen accessibility. 2018, pp. 131–142. cited By 10; Conference of 20th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2018 ; Conference Date: 22 October 2018 Through 24 October 2018; Conference Code:141110, doi:10.1145/3234695.3236349.
42. Oh, U.; Branham, S.; Findlater, L.; Kane, S. Audio-based feedback techniques for teaching touchscreen gestures **2015**. *7*. cited By 12, doi:10.1145/2764917.
43. Rodriguez-Sanchez, M.; Moreno-Alvarez, M.; Martin, E.; Borromeo, S.; Hernandez-Tamames, J. Accessible smartphones for blind users: A case study for a wayfinding system **2014**. *41*, 7210–7222. cited By 39, doi:10.1016/j.eswa.2014.05.031.
44. Watanabe, T.; Kaga, H.; Yagi, T. Evaluation of virtual tactile dots on touchscreens in map reading: Perception of distance and direction **2017**. *21*, 79–86. cited By 0, doi:10.20965/jaciii.2017.p0079.
45. Grussenmeyer, W.; Garcia, J.; Jiang, F. Feasibility of using haptic directions through maps with a tablet and smart watch for people who are blind and visually impaired. 2016, pp. 83–89. cited By 2; Conference of 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2016 ; Conference Date: 6 September 2016 Through 9 September 2016; Conference Code:123722, doi:10.1145/2935334.2935367.
46. Götzelmann, T. LucentMaps: 3D printed audiovisual tactile maps for blind and visually impaired people. 2016, pp. 81–90. cited By 28; Conference of 18th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2016 ; Conference Date: 24 October 2016 Through 26 October 2016; Conference Code:124437, doi:10.1145/2982142.2982163.
47. Yatani, K.; Banovic, N.; Truong, K. SpaceSense: Representing geographical information to visually impaired people using spatial tactile feedback; , 2012; pp. 415–424. cited By 67; Conference of 30th ACM Conference on Human Factors in Computing Systems, CHI 2012 ; Conference Date: 5 May 2012 Through 10 May 2012; Conference Code:89842, doi:10.1145/2207676.2207734.
48. Su, J.; Rosenzweig, A.; Goel, A.; De Lara, E.; Truong, K. Timbremap: Enabling the visually-impaired to use maps on touch-enabled devices; , 2010; pp. 17–26. cited By 60; Conference of 12th International Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI2010 ; Conference Date: 7 September 2010 Through 10 September 2010; Conference Code:82356, doi:10.1145/1851600.1851606.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

49. Götzelmann, T. Visually augmented audio-tactile graphics for visually impaired people. *ACM Transactions on Accessible Computing (TACCESS)* **2018**, *11*, 1–31.
50. Bateman, A.; Zhao, O.; Bajcsy, A.; Jennings, M.; Toth, B.; Cohen, A.; Horton, E.; Khattar, A.; Kuo, R.; Lee, F.; Lim, M.; Migasiuk, L.; Renganathan, R.; Zhang, A.; Oliveira, M. A user-centered design and analysis of an electrostatic haptic touchscreen system for students with visual impairments **2018**. *109*, 102–111. cited By 14, doi:10.1016/j.ijhcs.2017.09.004.
51. Gorlewicz, J.; Tennison, J.; Uesbeck, P.; Richard, M.; Palani, H.; Stefik, A.; Smith, D.; Giudice, N. Design Guidelines and Recommendations for Multimodal, Touchscreen-based Graphics **2020**. *13*. cited By 0, doi:10.1145/3403933.
52. Tekli, J.; Issa, Y.; Chbeir, R. Evaluating touch-screen vibration modality for blind users to access simple shapes and graphics **2018**. *110*, 115–133. cited By 11, doi:10.1016/j.ijhcs.2017.10.009.
53. Hahn, M.; Mueller, C.; Gorlewicz, J. The Comprehension of STEM Graphics via a Multisensory Tablet Electronic Device by Students with Visual Impairments **2019**. *113*, 404–418. cited By 1, doi:10.1177/0145482X19876463.
54. Klatzky, R.; Giudice, N.; Bennett, C.; Loomis, J. Touch-screen technology for the dynamic display of 2d spatial information without vision: Promise and progress **2014**. *27*, 359–378. cited By 45, doi:10.1163/22134808-00002447.
55. Silva, P.; Pappas, T.; Atkins, J.; West, J.; Hartmann, W. Acoustic-tactile rendering of visual information; , 2012; Vol. 8291. cited By 0; Conference of Human Vision and Electronic Imaging XVII ; Conference Date: 23 January 2012 Through 26 January 2012; Conference Code:89284, doi:10.1117/12.916166.
56. Hausberger, T.; Terzer, M.; Enneking, F.; Jonas, Z.; Kim, Y. SurfTics-Kinesthetic and tactile feedback on a touchscreen device. 2017, pp. 472–477. cited By 7; Conference of 7th IEEE World Haptics Conference, WHC 2017 ; Conference Date: 6 June 2017 Through 9 June 2017; Conference Code:129557, doi:10.1109/WHC.2017.7989947.
57. Palani, H.; Fink, P.; Giudice, N. Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices **2020**. *36*, 1393–1414. cited By 1, doi:10.1080/10447318.2020.1752464.
58. Tennison, J.; Gorlewicz, J. Non-visual Perception of Lines on a Multimodal Touchscreen Tablet **2019**. *16*. cited By 3, doi:10.1145/3301415.
59. Palani, H.; Tennison, J.; Giudice, G.; Giudice, N. Touchscreen-based haptic information access for assisting blind and visually-impaired users: Perceptual parameters and design guidelines **2019**. *794*, 837–847. cited By 0; Conference of AHFE International Conferences on Usability and User Experience and Human Factors and Assistive Technology, 2018 ; Conference Date: 21 July 2018 Through 25 July 2018; Conference Code:215439, doi:10.1007/978-3-319-94947-5_82.
60. Zhong, Y.; Lasecki, W.; Brady, E.; Bigham, J. Regionspeak: Quick comprehensive spatial descriptionsof complex images for blind users. 2015, Vol. 2015-April, pp. 2353–2362. cited By 31; Conference of 33rd Annual CHI Conference on Human Factors in Computing Systems, CHI 2015 ; Conference Date: 18 April 2015 Through 23 April 2015; Conference Code:116824, doi:10.1145/2702123.2702437.
61. Guo, A.; Kong, J.; Rivera, M.; Xu, F.; Bigham, J. StateLens: A reverse engineering solution for making existing dynamic touchscreens accessible. 2019, pp. 371–385. cited By 6; Conference of 32nd Annual ACM Symposium on User Interface Software and Technology, UIST 2019 ; Conference Date: 20 October 2019 Through 23 October 2019; Conference Code:153032, doi:10.1145/3332165.3347873.
62. Goncu, C.; Madugalla, A.; Marinai, S.; Marriott, K. Accessible on-line floor plans. 2015, pp. 388–398. cited By 11; Conference of 24th International Conference on World Wide Web, WWW 2015 ; Conference Date: 18 May 2015 Through 22 May 2015; Conference Code:119450, doi:10.1145/2736277.2741660.
63. Madugalla, A.; Marriott, K.; Marinai, S.; Capobianco, S.; Goncu, C. Creating Accessible Online Floor Plans for Visually Impaired Readers **2020**. *13*. cited By 0, doi:10.1145/3410446.
64. Taylor, B.; Dey, A.; Siewiorek, D.; Smailagic, A. Customizable 3D printed tactile maps as interactive overlays. 2016, pp. 71–79. cited By 32; Conference of 18th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2016 ; Conference Date: 24 October 2016 Through 26 October 2016; Conference Code:124437, doi:10.1145/2982142.2982167.
65. Grussenmeyer, W.; Folmer, E. AudioDraw: User preferences in non-visual diagram drawing for touchscreens. 2016. cited By 8; Conference of 13th Web for All Conference, W4A 2016 ; Conference Date: 11 April 2016 Through 13 April 2016; Conference Code:122862, doi:10.1145/2899475.2899483.
66. Reinholt, K.; Guinness, D.; Kane, S. Eyedescribe: Combining eye gaze and speech to automatically create accessible touch screen artwork. 2019, pp. 101–112. cited By 0; Conference of 14th ACM International Conference on Interactive Surfaces and Spaces, ISS 2019 ; Conference Date: 10 November 2019 Through 13 November 2019; Conference Code:155017, doi:10.1145/10.1145/3343055.3359722.
67. Jayant, C.; Ji, H.; White, S.; Bigham, J.P. Supporting blind photography. The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility, 2011, pp. 203–210.
68. Hayhoe, S. Expanding our vision of museum education and perception: An analysis of three case studies of independent blind arts learners. *Harvard Educational Review* **2013**, *83*, 67–86.
69. Handa, K.; Dairoku, H.; Toriyama, Y. Investigation of priority needs in terms of museum service accessibility for visually impaired visitors. *British journal of visual impairment* **2010**, *28*, 221–234.
70. Hayhoe, S. Why Do We Think That People who are Visually Impaired Don't Want to Know About the Visual Arts II? Creative Sandpit@ the (virtual) American Museum, 2020.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 March 2021 doi:10.20944/preprints202103.0218.v1

71. Asakawa, S.; Guerreiro, J.; Ahmetovic, D.; Kitani, K.M.; Asakawa, C. The Present and Future of Museum Accessibility for People with Visual Impairments. Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 2018, pp. 382–384.
72. The Audio Guide by Metropolitan Museum of Art. https://www.metmuseum.org/visit/audio-guide. Accessed: 2019-04-20.
73. Accessibility at Museum of Modern Art. http://www.moma.org/learn/disabilities/sight. Accessed: 2019-04-20.
74. Art Beyond Sight Tours at Seattle Art Museum. http://www.seattleartmuseum.org/visit/accessibility#tou. Accessed: 2019-04-20.
75. Out Loud, The Andy Warhol Museum's inclusive audio guide. https://itunes.apple.com/us/app/the-warhol-out-loud/id1103407119.
76. Feldman, E.B. Practical Art Criticism, 1994.
77. Ponsard, C.; Fries, V. An accessible viewer for digital comic books. International Conference on Computers for Handicapped Persons. Springer, 2008, pp. 569–577.
78. Rayar, F.; Oriola, B.; Jouffrais, C. ALCOVE: an accessible comic reader for people with low vision. Proceedings of the 25th International Conference on Intelligent User Interfaces, 2020, pp. 410–418.
79. Wang, Y.; Wang, W.; Liang, W.; Yu, L.F. Comic-guided speech synthesis. *ACM Transactions on Graphics (TOG)* **2019**, *38*, 1–14.
80. Guérin, C.; Rigaud, C.; Mercier, A.; Ammar-Boudjelal, F.; Bertet, K.; Bouju, A.; Burie, J.C.; Louis, G.; Ogier, J.M.; Revel, A. eBDtheque: a representative database of comics. 2013 12th International Conference on Document Analysis and Recognition. IEEE, 2013, pp. 1145–1149.
81. Figueiredo, F.; Benevenuto, F.; Almeida, J.M. The tube over time: characterizing popularity growth of youtube videos. Proceedings of the fourth ACM international conference on Web search and data mining, 2011, pp. 745–754.
82. Hartley, J. Uses of YouTube: Digital literacy and the growth of knowledge. *YouTube: Online video and participatory culture* **2009**, pp. 126–143.
83. Yuksel, B.F.; Fazli, P.; Mathur, U.; Bisht, V.; Kim, S.J.; Lee, J.J.; Jin, S.J.; Siu, Y.T.; Miele, J.A.; Yoon, I. Human-in-the-Loop Machine Learning to Increase Video Accessibility for Visually Impaired and Blind Users. Proceedings of the 2020 ACM Designing Interactive Systems Conference, 2020, pp. 47–60.
84. Encelle, B.; Ollagnier-Beldame, M.; Pouchot, S.; Prié, Y. Annotation-based video enrichment for blind people: A pilot study on the use of earcons and speech synthesis. The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility, 2011, pp. 123–130.
85. Maneesaeng, N.; Punyabukkana, P.; Suchato, A. Accessible video-call application on android for the blind. *Lecture Notes on Software Engineering* **2016**, *4*, 95.
86. Wu, S.; Wieland, J.; Farivar, O.; Schiller, J. Automatic alt-text: Computer-generated image descriptions for blind users on a social network service. Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, 2017, pp. 1180–1192.
87. Rector, K.; Salmon, K.; Thornton, D.; Joshi, N.; Morris, M.R. 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* **2017**, *1*, 93.
88. Fischer, B.; Peine, A.; Östlund, B. The importance of user involvement: a systematic review of involving older users in technology design. *The Gerontologist* **2020**, *60*, e513–e523.
89. Schuler, D.; Namioka, A. *Participatory design: Principles and practices*; CRC Press, 1993.
90. Sanders, E.B. From user-centered to participatory design approaches. *Design and the social sciences: Making connections* **2002**, *1*, 1.
91. Sahib, N.G.; Stockman, T.; Tombros, A.; Metatla, O. Participatory design with blind users: a scenario-based approach. IFIP Conference on Human-Computer Interaction. Springer, 2013, pp. 685–701.