

Review

# Image Accessibility for Screen Reader Users: A Systematic Review and a Road Map

Uran Oh \*, Hwayeon Joh and YunJung Lee

Department of Computer Science and Engineering, Ewha Womans University, Seoul 03760, Korea; e.jhy@ewhain.net (H.J.); yunjung\_lee20@ewhain.net (Y.L.)

\* 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 there is little study how the generation of image-related information can be automated. Moreover, we confirmed that the involvement of screen reader users is mostly limited to evaluations, while input from target users during the design process is particularly important for the development of assistive technologies. 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.



**Citation:** Oh, U.; Joh, H.; Lee, Y. Image Accessibility for Screen Reader Users: A Systematic Review and a Road Map. *Electronics* **2021**, *10*, 953. <https://doi.org/10.3390/electronics10080953>

Academic Editor: Juan M. Corchado

Received: 8 March 2021

Accepted: 5 April 2021

Published: 16 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

According to the World Health Organization, at least 2.2 billion people have a visual impairment, and the number is likely to increase with population growth and aging [1]. For them, understanding visual information is one of the main challenges.

To improve the accessibility of images for people who are blind or have low vision (BLV), a number of studies have been conducted to assess the effectiveness of custom-made tactile versions of images [2–8]. Cavazos et al. [5], 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. [7] also investigated tactile graphics and 3D models to deliver map information such as the number of entrances, location and direction of certain landmarks. This approach with extra tactile feedback is found to be effective as it can deepen one's spatial understanding of images by touch [9–11]. However, it requires additional equipment, which potential users have limited access to (e.g., 3D printer, custom devices). Moreover, tactile representations need to be designed and built for each image, and thus it is not ideal for supporting a number of different images in terms of time and cost.

Meanwhile, others have 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 [12–14]. For instance, Zhong et al. [13] generated alt text for images on the web that are identified as important using crowdsourcing. In addition, Stangl et al. [12] 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 approach is 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 the issues 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 identifying the challenges to be solved, we conducted a systematic literature review of 33 papers, following PRISMA guidelines [15]. To be specific, our goal is to identify the following: supported image types, provided information, collection and the delivery method of the information, and the involvement of screen reader users during the design and development process.

As a result, we found that research studies 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 was given. We also confirmed that while most user studies are conducted with participants who are blind or have low vision for user evaluation, a few studies involved target users during the design process.

In addition, to demonstrate how other types of images can be made accessible using touchscreen devices, we introduce two of our systems: AccessArt [16–18] for artwork and AccessComics [19] 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 following are the contributions of our work:

- 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, and the involvement of screen reader users.
- The identifications of key challenges and limitations of studying image accessibility of screen reader users using touchscreen devices.
- Recommendations for future research directions.

The rest of the content covers a summary of prior studies on image accessibility and touchscreen accessibility for BLV people (Section 2), followed by a description of how we conducted a systematic review (Section 3), and the results (Section 4), demonstrations of two systems for improving the accessibility of artwork and digital comics (Section 5), discussions on the current limitations and potentials of existing work and suggestions on future work (Section 6), and conclusions (Section 7).

## 2. Related Work

Our work is inspired by prior work on image accessibility and touchscreen accessibility for people who are blind or who have low vision.

### 2.1. Image Accessibility

Screen readers cannot describe an image unless its metadata such as alt text are present. 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 [12–14,20–22]. Winters et al., for instance (Reference [14]) proposed an auditory display for social media that 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. [12] developed computer vision (CV) and natural language processing (NLP) modules to extract information about clothing images on an 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 [22], 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. [20] 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 [2,4,5,23,24]. Götzelmann et al. [23], for instance, presented a 3D-printed map to convey geographic information by touch. However, some worked on using computational methods to automatically generate tactile representations [3,25]. For example, Rodrigues et al. [25] proposed an algorithm for creating tactile representations of objects presented in an artwork varying in shape and depth. While it is promising, we focused on improving image accessibility on a touchscreen, which is widely adopted in personal devices such as smartphones and tablets, since it does not require additional hardware.

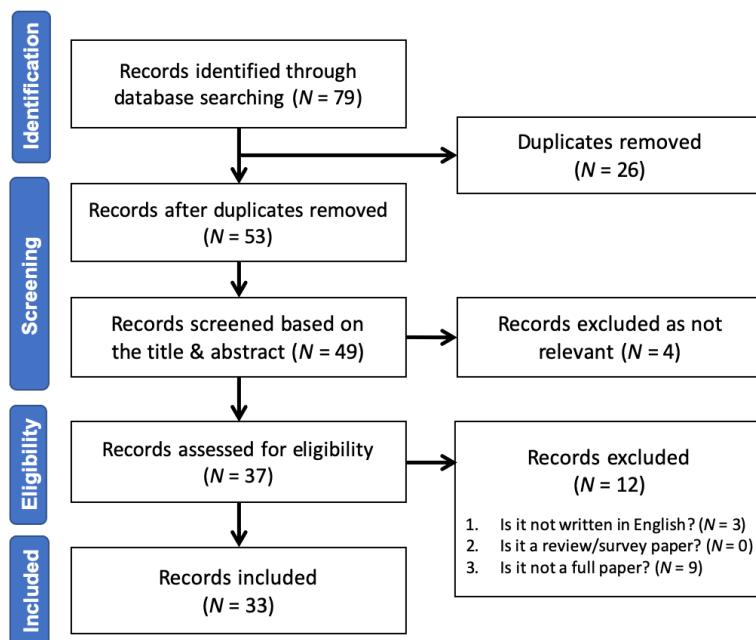
## 2.2. Touchscreen Accessibility

While we chose to focus on touchscreen-based image accessibility, touchscreen devices are innately inaccessible as they require accurate hand–eye coordination [26]. Thus, various studies have been conducted to improve touchscreen accessibility by providing tactile feedback using additional hardware devices [27–29]. *TouchCam*, for example, designed and implemented a camera-based wearable device that can be worn on a finger, which is used to access one'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 [6,30]. For instance, *TouchPlaces* [6] 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 [26,31–34]. *BrailleTouch* [32] and No-Look Notes [34], 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 (<https://www.apple.com/accessibility/vision/>, accessed on 15 April 2021) and Android's Talkback (<https://support.google.com/accessibility/android/answer/6283677?hl=en>). These screen readers read 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 they are readily available to a large number of end-users including BLV people as they have their own personal devices with touchscreens. In addition, as touchscreen devices offer screen reader functionality, they are 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 [15]. The process is shown in Figure 1.



**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 has been supported for BLV people?
- RQ3. How has image-related information been collected?
- RQ4. How has image-related information been delivered?
- RQ5. How have BLV people been 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). We then 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 the 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 or workshop and case study papers (N = 9).

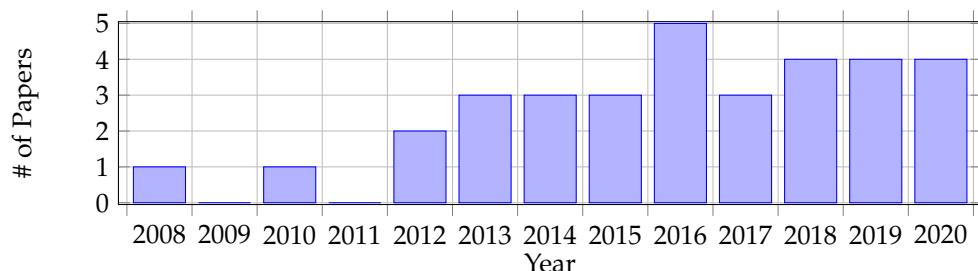
Then, papers were included if and only if the goal of the paper was improving the accessibility of any type of images for people with visual impairments.

#### 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 Section 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 since 2013. As for the country of authors' affiliations, United States had the dominant number of papers, which was 17. It was followed by Germany ( $N = 4$ ), Canada ( $N = 3$ ), Australia ( $N = 2$ ), and Japan ( $N = 2$ ). Other countries had only one paper each: 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. To be specific, approximately half of the papers studied images of maps ( $N = 10$ ) and graphs ( $N = 6$ ). Interestingly, while the accessibility of photographs for BLV was largely investigated in terms of web accessibility [21,38], only three out of 33 papers aimed to support photographs using touchscreen devices in particular. In addition, as touchscreen devices themselves have accessibility issues for people with visual impairments, requiring accurate hand–eye coordination [26], four papers focused on improving the accessibility of the touchscreen-based interface itself such as soft buttons [39–41] and gestures [42].

**Table 1.** Types of supported images for people who are blind or have low vision (BLV).

Image Type	Count	Papers
Map	10	[23,24,37,43–49]
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]

#### 4.3. Information Type

We have identified types of information provided to BLV to improve the accessibility of images (*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 was provided for map images. On the other hand, shape, size, boundary of objects, and 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 ( $N = 2$ ) were also present. Note that only two papers supported color information for photographs [35] and artwork [66]. Other information types include types of graph [53] and weather [24]. Five papers did not specify the information they provided.

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

Info Type	Count	Papers
Object Names	10	[23,24,43,46,49,55,62–64,66]
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	[23,46–49]
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]

#### 4.4. Information Preparation

In addition to types of information supported to improve image accessibility (*RQ3*), we have found that most studies have not specified how the visual information is collected or created (see Table 3). 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 one-third of the studies seem to manually create 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 visual information of images: image processing, optical character recognition (OCR), and computer vision ( $N = 5$ ). Meanwhile, two papers proposed a system where the image descriptions are provided by crowdworkers.

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

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]
Object Segmentation	2	[35,36]
Crowdsourcing	2	[60,61]
OCR	1	[61]
Not specified	18	[23,24,37,41,42,44–52,55,57,63,65]

#### 4.5. Interaction Types

We were also interested in how image-related information is delivered using touchscreen 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 were allowed to share image feeds from cameras with others so that they could get information about their surrounding physical objects such as touch panels on a microwave [60,61].

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

Input	Count	Papers
Touch	28	[24,35–37,39,42–46,48–59,61–64,64–66]
Gesture	5	[37,42,46,47,49]
Keyboard	4	[24,35–37]
Stylus	3	[37,40,61]
Mouse	2	[35,37]
Voice command	2	[46,49]
Camera	2	[60,61]
Physical UI	1	[41]

*Output Modalities.* As for output modalities, various types of feedback techniques were used (see Table 5). Approximately half of the studies used a single modality: audio only ( $N = 14$ ; including both speech and non-speech audio) or vibration only ( $N = 3$ ). On the other hand, others used multimodal feedback, where the combination of audio and vibration was most frequent ( $N = 6$ ), followed by audio with tactile feedback ( $N = 5$ ). The most 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 pitches of sound [24,35,36,42] or rhythms [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. [23] used a 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.

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

Output	Count	Papers
Speech	20	[24,36,41,42,45–49,51,53,54,58,60–63,65,66]
Non-speech Audio	14	[23,24,24,35–37,42,43,48,51,52,55,55,65]
Vibration	14	[39,40,43–45,47,50–55,57,59]
Tactile Feedback	9	[23,24,39,41,44,47,56,63,64]
Force Feedback	3	[39,40,56]

#### 4.6. Involvement of BLV

Finally, we checked if BLV, the target users, were involved in the system development and evaluation processes; see Table 6. We first examined if user evaluation was conducted regardless of whether target users were involved or not. As a result, we found that all studies but two had tested their system with human subjects. Most of them had a controlled lab study, where metrics related to task performance were collected for evaluation such as the number of correct responses and completion time. However, close to half of the studies had subjective assessments such as easiness and satisfaction in a Likert scales, or open-ended comments about their experience after using the systems.

Of the remaining 31 papers, three papers had user studies but with 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, seven studies used participatory design approaches during their design process. Moreover, some papers had BLV participate in their formative qualitative studies at an early stage of their system development to make their ideas concrete (i.e., survey, interview).

**Table 6.** Methodologies used in the studies and BLV's involvement in system design and evaluation. Note that the following three studies conducted user studies with blind-folded sighted participants [54–56].

Process	Methodology	Count	Papers
Evaluation	Controlled Lab Study	28	[36,37,39–44,46–63,66]
	Subjective Assessment	15	[24,36,41–43,47,50,51,53,58,61–65,65,66]
Design	Participatory Design	7	[24,50,61–65]
Ideation	Interview	3	[24,50,65]
	Survey	2	[24,65]
N/A	No user study	2	[23,35]

#### 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 focused on providing knowledge or information based on facts (e.g., maps, graphs) to users rather than offering improved user experience that BLV users can enjoy allowing subjective interpretations. Thus, we focused on supporting two types of images in particular that are rarely studied for screen reader users: artwork and comics. Here, we demonstrate how these two types of images can be supported and appreciated with improved accessibility: AccessArt [16–18] and AccessComics [19] (see Table 7 as well).

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

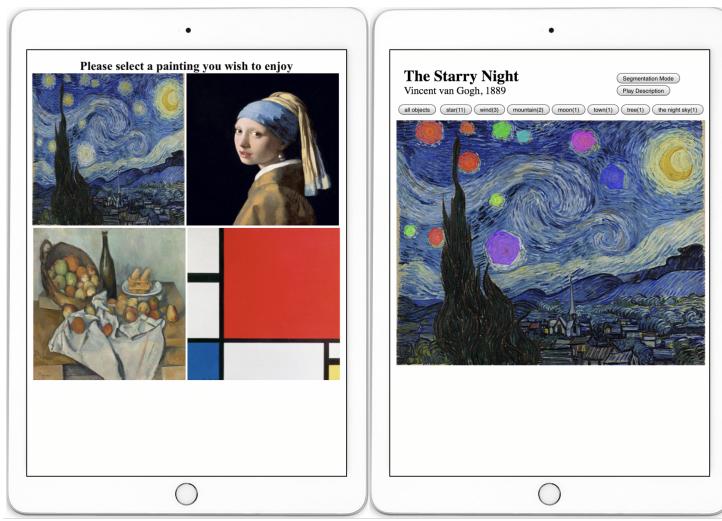
	AccessArt	AccessComics
Image Type	2D paintings	Digital comics
Information Type	Object names, spatial layout, shape, size, color	Panel number, background, character appearance and actions
Data Preparation	Manual, crowdsourced	Manual (visual descriptions) open dataset (panel, balloon, script)
Interaction	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

### 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 [2–5]. However, it is not feasible to make these replicas for every exhibited artwork. Thus, we began to design and implement touchscreen-based artwork exploration tool called *AccessArt*.

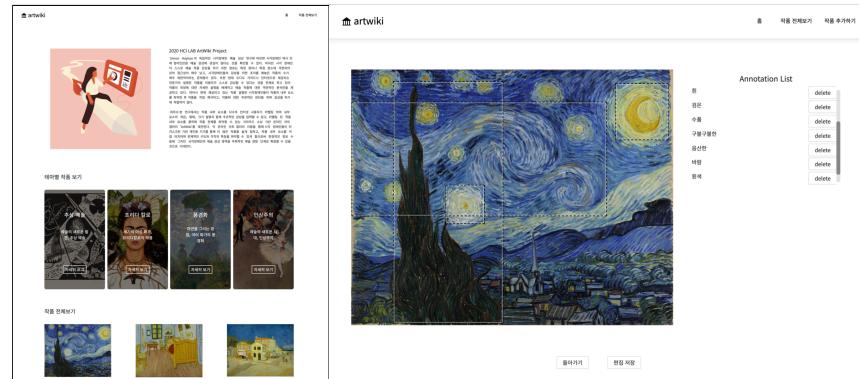
*AccessArt Ver1.* The very first version of AccessArt is shown in Figure 3, in which it had four paintings with varying genres: landscape, portrait, abstract, and still life [17]. As for the object-level labels, we segmented each object along with descriptions. Then we developed a web application that 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 can 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 provided positive feedback.

*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 were not expected have expertise in art [16]. We used Amazon Mechanical Turk (<https://www.mturk.com/>) for collecting object-label metadata for eight different paintings from an 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 the *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 3.** User interface prototype with two interaction modes: *Overview* (left) and *Exploration* (right). As for the *Exploration Mode* example, *star* is selected as object of interest, highlighted in various colors.

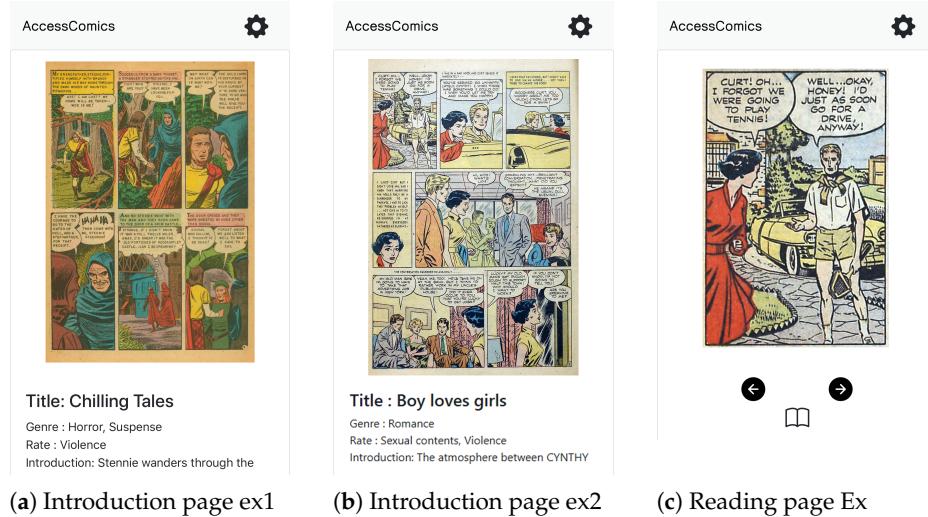
*AccessArt Ver3.* As a final step, we implemented an online platform (<https://artwiki2020.vercel.app/>) as shown in Figure 4. It is designed to allow anonymous users to freely volunteer to provide object segmentation and description, inspired by Wikipedia ([www.wikipedia.org](http://www.wikipedia.org)). While no user evaluation has been 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.



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

## 5.2. AccessComics for Comics Accessibility

Compared to artwork accessibility, fewer 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 the .pdf version of digital comics. Inspired by this study, our system, *AccessComics* [19], is designed to provide BLV users with overview (as shown in Figure 5a,b), various reading units (i.e., page, strip, panel), magnifier, text-to-speech, and autoplay. Moreover, we mapped different voices with different characters to offer a high sense of immersion in addition to improved accessibility similar to how Wang et al. [79] used a voice synthesis technique that can express various emotional states given scripts. Here, we briefly describe how the system is implemented.



**Figure 5.** Solution and error progression curve for different times and  $\varepsilon$  values, with  $n = m = 40$  and  $t_{max} = 1$ .

*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 5c, the following would be played using the audio channel starting with the panel number, followed by background and character-related visual descriptions.

## 6. Discussion

Here, we discuss the current state of research 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 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 been 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. Types of Information Supported for Different Image Types

As it has been found in prior work that BLV wish to get different types of information depending on the context [21], 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. On the other hand, shape, size, and line-length information were conveyed to users for geometric objects. However, little study has been done about other types of images, although specific locations or spatial

relationships of objects within an image such as photographs and touchscreen user interface are considered important [20,30]. To identify types of information that users are interested in for each type of image, adopting recommendation techniques [86,87] can be a solution for providing user-specific content based on users' preference, interests, and needs.

### 6.3. Limited Room for Subjective Interpretations

The majority of the studies have prioritized images that contain useful information (e.g., facts, knowledge) over images that can be interpreted subjectively, differently from one person to another, such as artwork, using touchscreen devices. Even for artwork images, many studies have focused on delivering encyclopedia-style explanations (i.e., title, artist, painting styles) [2–5,88]. AccessArt [16] was an exception, where they demonstrated if their artwork appreciation system can enable BLV people to make their own judgements and criticism about artwork they explored. We believe that more investigations are needed to improve the experience of enjoying the content of the images or of making decisions based on subjective judgements for BLV people (e.g., providing a summary of product reviews of others as a reference).

### 6.4. Automatic Retrieval of Metadata of Images for Scalability

The greatest number of studies that we have identified in our systematic review, all but seven out of 33 papers, assumed that image-related information is given. If not, researchers 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) (<https://www.w3.org/TR/WCAG21/>). Moreover, it is not feasible for a couple of researchers to generate metadata for individual images. Thus, automatic approaches such as machine learning techniques have been studied [12,13,89,90]. However, since the accuracy of descriptions produced by humans is not as high, we recommend the crowdsourcing approach for generating descriptions [18,20] if precise annotations are needed. This can serve as a human–AI collaboration for validating auto-generated annotations [13]. Eventually, these data can be used to train machine learning models for implementing a fully automated image description generation system [91–93].

### 6.5. Limited Input and Output Modalities of Touchscreen Devices

Unlike other assistive systems that require BLV to physically visit certain locations (e.g., [75,88]) or that require special hardware devices with tactile cues (e.g., [3,4]), 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 touchscreen devices can offer are limited to audio and vibration feedback. To provide more intuitive and rich feedback, more in-depth studies on how to ease the design of 2.5D or 3D models and how the cost and time for producing tactile representations of images can be reduced should be conducted. One way to do so is open-sourcing the process, as in Instructables (<https://www.instructables.com/>), which is an online community where people explore and share instructions for do-it-yourself projects. Meanwhile, we also recommend touchscreen-based approaches to make a larger number of images accessible to a greater number of BLV people.

### 6.6. 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 participate in the design process at an early state when developing a new technology [94]. 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 disabilities [95–97].

## 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. Furthermore, we found that only about one-third of the papers provide multimodal feedback of audio and haptic. Moreover, our findings show that ways to collect image descriptions was out of the scope of interest for most studies, suggesting that automatic retrievals of image-related information is one of the bottlenecks for making images accessible on 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. Blindness and Vision Impairment. Available online: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accessed on 29 March 2021).
2. Gyoşhev, S.; Karastoyanov, D.; Stoimenov, N.; Cantoni, V.; Lombardi, L.; Setti, A. Exploiting a graphical Braille display for art masterpieces. In Proceedings of the International Conference on Computers Helping People with Special Needs, Linz, Austria, 11–13 July 2018; pp. 237–245.
3. Cantoni, V.; Lombardi, L.; Setti, A.; Gyoşhev, S.; Karastoyanov, D.; Stoimenov, N. Art Masterpieces Accessibility for Blind and Visually Impaired People. In Proceedings of the International Conference on Computers Helping People with Special Needs, Linz, Austria, 11–13 July 2018; pp. 267–274.
4. Iranzo Bartolome, J.; Cavazos Quero, L.; Kim, S.; Um, M.Y.; Cho, J. Exploring Art with a Voice Controlled Multimodal Guide for Blind People. In Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction, Tempe, AZ, USA, 17–20 March 2019; pp. 383–390.
5. 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. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 346–348.
6. Kane, S.K.; Morris, M.R.; Wobbrock, J.O. Touchplates: Low-cost tactile overlays for visually impaired touch screen users. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, DC, USA, 21–23 October 2013; pp. 1–8.
7. Holloway, L.; Marriott, K.; Butler, M. Accessible maps for the blind: Comparing 3D printed models with tactile graphics. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–13.
8. 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 Trans. Graph. (TOG)* **2020**, *39*, 89. [[CrossRef](#)]
9. Paterson, M. ‘Seeing with the hands’: Blindness, touch and the Enlightenment spatial imaginary. *Br. J. Vis. Impair.* **2006**, *24*, 52–59. [[CrossRef](#)]
10. Heller, M.A. *The Psychology of Touch*; Psychology Press: East Sussex, UK, 2013.
11. Watanabe, T.; Yamaguchi, T.; Nakagawa, M. Development of software for automatic creation of embossed graphs. In Proceedings of the International Conference on Computers for Handicapped Persons, Linz, Austria, 11–13 July 2012; pp. 174–181.
12. 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. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 107–118.
13. Zhong, Y.; Matsubara, M.; Morishima, A. Identification of Important Images for Understanding Web Pages. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 3568–3574.

14. Winters, R.M.; Joshi, N.; Cutrell, E.; Morris, M.R. Strategies for auditory display of Social Media. *Ergon. Des.* **2019**, *27*, 11–15. [[CrossRef](#)]
15. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int. J. Surg.* **2010**, *8*, 336–341. [[CrossRef](#)] [[PubMed](#)]
16. Nahyun Kwon, Y.L.; Oh, U. Supporting a Crowd-powered Accessible Online Art Gallery for People with Visual Impairments: A Feasibility Study. *Univers. Access Inf. Soc.* **2021**, accepted.
17. Kwon, N.; Koh, Y.; Oh, U. Supporting object-level exploration of artworks by touch for people with visual impairments. In Proceedings of the Conference of 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2019), Pittsburgh, PA, USA, 28–30 October 2019; pp. 600–602. [[CrossRef](#)]
18. Soohyun Yeo, S.C.; Oh, U. Crowdsourcing Based Online Art Gallery for Visual Impairment People. *HCI Korea Annu. Conf.* **2021**, *33*, 531–535.
19. Lee, Y.; Joh, H.; Yoo, S.; Oh, U. AccessComics: An Accessible Digital Comic Book Reader for People with Visual Impairments. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 12–22 October 2018; Association for Computing Machinery: New York, NY, USA, 2018; ISBN 978-1-4503-5650-3.
20. Morris, M.R.; Johnson, J.; Bennett, C.L.; Cutrell, E. Rich representations of visual content for screen reader users. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; p. 59.
21. Stangl, A.; Morris, M.R.; Gurari, D. “Person, Shoes, Tree. Is the Person Naked?” What People with Vision Impairments Want in Image Descriptions. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13.
22. Goncu, C.; Marriott, K. Creating ebooks with accessible graphics content. In Proceedings of the 2015 ACM Symposium on Document Engineering, Lausanne, Switzerland, 8–11 September 2015; pp. 89–92.
23. Götzelmüller, T. CapMaps: Capacitive sensing 3D printed audio-tactile maps. In Proceedings of the Conference of 15th International Conference on Computers Helping People with Special Needs (ICCHP 2016), Linz, Austria, 13–15 July 2016; Volume 9759, pp. 146–152. [[CrossRef](#)]
24. 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 Stud.* **2013**, *8*, 93–110.
25. 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. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Orlando, FL, USA, 21–25 July 2018; pp. 243–253.
26. McGookin, D.; Brewster, S.; Jiang, W. Investigating touchscreen accessibility for people with visual impairments. In Proceedings of the 5th Nordic conference on Human-computer interaction: Building bridges, Lund, Sweden, 20–22 October 2008; pp. 298–307.
27. Vanderheiden, G.C. Use of audio-haptic interface techniques to allow nonvisual access to touchscreen appliances. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications Sage CA: Los Angeles, CA, USA, 1996; Volume 40, p. 1266.
28. Landau, S.; Wells, L. Merging tactile sensory input and audio data by means of the Talking Tactile Tablet. In Proceedings of the EuroHaptics, Dublin, Ireland, 6–9 July 2003; Volume 3, pp. 414–418.
29. Stearns, L.; Oh, U.; Findlater, L.; Froehlich, J.E. TouchCam: Realtime Recognition of Location-Specific On-Body Gestures to Support Users with Visual Impairments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *1*, 1–23. [[CrossRef](#)]
30. 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. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, Santa Barbara, CA, USA, 16–19 October 2011; pp. 273–282.
31. Kane, S.K.; Bigham, J.P.; Wobbrock, J.O. Slide rule: Making mobile touch screens accessible to blind people using multi-touch interaction techniques. In Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility, Halifax, NS, Canada, 13–15 October 2008; pp. 73–80.
32. Frey, B.; Southern, C.; Romero, M. BrailleTouch: Mobile texting for the visually impaired. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Orlando, FL, USA, 9–14 July 2011; pp. 19–25.
33. 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.
34. Bonner, M.N.; Brudvik, J.T.; Abowd, G.D.; Edwards, W.K. No-look notes: Accessible eyes-free multi-touch text entry. In Proceedings of the International Conference on Pervasive Computing, Helsinki, Finland, 17–20 May 2010; pp. 409–426.
35. Toffa, O.; Mignotte, M. A Hierarchical Visual Feature-Based Approach for Image Sonification. *IEEE Trans. Multimed.* **2021**, *23*, 706–715. [[CrossRef](#)]
36. Banf, M.; Blanz, V. Sonification of images for the visually impaired using a multi-level approach. In Proceedings of the Conference of 4th Augmented Human International Conference, Stuttgart, Germany, 7–8 March 2013; pp. 162–169. [[CrossRef](#)]
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. In Proceedings of the 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, 8514 LNCS, HCI International 2014, Crete, Greece, 22–27 June 2014; pp. 154–165. [[CrossRef](#)]

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*. [[CrossRef](#)]
39. Onishi, J.; Sakajiri, M.; Miura, T.; Ono, T. Fundamental study on tactile cognition through haptic feedback touchscreen. In Proceedings of the Conference of 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, Manchester, UK, 13–16 October 2013; pp. 4207–4212. [[CrossRef](#)]
40. Kyung, K.U.; Lee, J.Y.; Park, J. Haptic stylus and empirical studies on braille, button, and texture display. *J. Biomed. Biotechnol.* **2008**, *2008*. [[CrossRef](#)] [[PubMed](#)]
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. In Proceedings of the Conference of 20th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2018, Galway, Ireland, 22–24 October 2018; pp. 131–142. [[CrossRef](#)]
42. Oh, U.; Branham, S.; Findlater, L.; Kane, S. Audio-based feedback techniques for teaching touchscreen gestures. *ACM Trans. Access. Comput. (TACCESS)* **2015**, *7*. [[CrossRef](#)]
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. *Expert Syst. Appl.* **2014**, *41*, 7210–7222. [[CrossRef](#)]
44. Watanabe, T.; Kaga, H.; Yagi, T. Evaluation of virtual tactile dots on touchscreens in map reading: Perception of distance and direction. *J. Adv. Comput. Intell. Intell. Inform.* **2017**, *21*, 79–86. [[CrossRef](#)]
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. In Proceedings of the Conference of 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2016, Florence, Italy, 6–9 September 2016; pp. 83–89. [[CrossRef](#)]
46. Götzmann, T. LucentMaps: 3D printed audiovisual tactile maps for blind and visually impaired people. In Proceedings of the Conference of 18th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2016, Reno, NV, USA, 24–26 October 2016; pp. 81–90. [[CrossRef](#)]
47. Yatani, K.; Banovic, N.; Truong, K. SpaceSense: Representing geographical information to visually impaired people using spatial tactile feedback. In Proceedings of the Conference of 30th ACM Conference on Human Factors in Computing Systems, CHI 2012, Austin, TX, USA, 5–10 May 2012; pp. 415–424. [[CrossRef](#)]
48. Su, J.; Rosenzweig, A.; Goel, A.; De Lara, E.; Truong, K. Timbremap: Enabling the visually-impaired to use maps on touch-enabled devices. In Proceedings of the Conference of 12th International Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI2010, Lisbon, Portugal, 7–10 September 2010; pp. 17–26. [[CrossRef](#)]
49. Götzmann, T. Visually augmented audio-tactile graphics for visually impaired people. *ACM Trans. Access. Comput. (TACCESS)* **2018**, *11*, 1–31. [[CrossRef](#)]
50. Bateman, A.; Zhao, O.; Bajcsy, A.; Jennings, M.; Toth, B.; Cohen, A.; Horton, E.; Khattar, A.; Kuo, R.; Lee, F.; et al. A user-centered design and analysis of an electrostatic haptic touchscreen system for students with visual impairments. *Int. J. -Hum.-Comput. Stud.* **2018**, *109*, 102–111. [[CrossRef](#)]
51. Gorlewicz, J.; Tennison, J.; Uesbeck, P.; Richard, M.; Palani, H.; Stefk, A.; Smith, D.; Giudice, N. Design Guidelines and Recommendations for Multimodal, Touchscreen-based Graphics. *ACM Trans. Access. Comput. (TACCESS)* **2020**, *13*. [[CrossRef](#)]
52. Tekli, J.; Issa, Y.; Chbeir, R. Evaluating touch-screen vibration modality for blind users to access simple shapes and graphics. *Int. J. Hum. Comput. Stud.* **2018**, *110*, 115–133. [[CrossRef](#)]
53. Hahn, M.; Mueller, C.; Gorlewicz, J. The Comprehension of STEM Graphics via a Multisensory Tablet Electronic Device by Students with Visual Impairments. *J. Vis. Impair. Blind.* **2019**, *113*, 404–418. [[CrossRef](#)]
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. *Multisensory Res.* **2014**, *27*, 359–378. [[CrossRef](#)]
55. Silva, P.; Pappas, T.; Atkins, J.; West, J.; Hartmann, W. Acoustic-tactile rendering of visual information. In Proceedings of the Conference of Human Vision and Electronic Imaging XVII, Burlingame, CA, USA, 23–26 January 2012; Volume 8291. [[CrossRef](#)]
56. Hausberger, T.; Terzer, M.; Enneking, F.; Jonas, Z.; Kim, Y. SurfTics-Kinesthetic and tactile feedback on a touchscreen device. In Proceedings of the Conference of 7th IEEE World Haptics Conference (WHC 2017), Munich, Germany, 6–9 June 2017; pp. 472–477. [[CrossRef](#)]
57. Palani, H.; Fink, P.; Giudice, N. Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices. *Int. J. Hum.-Comput. Interact.* **2020**, *36*, 1393–1414. [[CrossRef](#)]
58. Tennison, J.; Gorlewicz, J. Non-visual Perception of Lines on a Multimodal Touchscreen Tablet. *ACM Trans. Appl. Percept. (TAP)* **2019**, *16*. [[CrossRef](#)]
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. In Proceedings of the Conference of AHFE International Conferences on Usability and User Experience and Human Factors and Assistive Technology, Orlando, FL, USA, 21–25 July 2018; Volume 794, pp. 837–847. [[CrossRef](#)]
60. Zhong, Y.; Lasecki, W.; Brady, E.; Bigham, J. Regionspeak: Quick comprehensive spatial descriptions of complex images for blind users. In Proceedings of the Conference of 33rd Annual CHI Conference on Human Factors in Computing Systems, CHI 2015, Seoul, Korea, 18–23 April 2015; Volume 2015, pp. 2353–2362. [[CrossRef](#)]

61. Guo, A.; Kong, J.; Rivera, M.; Xu, F.; Bigham, J. StateLens: A reverse engineering solution for making existing dynamic touchscreens accessible. In Proceedings of the Conference of 32nd Annual ACM Symposium on User Interface Software and Technology, UIST 2019, New Orleans, LA, USA, 20–23 October 2019; pp. 371–385. [CrossRef]
62. Goncu, C.; Madugalla, A.; Marinai, S.; Marriott, K. Accessible on-line floor plans. In Proceedings of the Conference of 24th International Conference on World Wide Web, WWW 2015, Florence, Italy, 18–22 May 2015; pp. 388–398. [CrossRef]
63. Madugalla, A.; Marriott, K.; Marinai, S.; Capobianco, S.; Goncu, C. Creating Accessible Online Floor Plans for Visually Impaired Readers. *ACM Trans. Access. Comput. (TACCESS)* **2020**, *13*. [CrossRef]
64. Taylor, B.; Dey, A.; Siewiorek, D.; Smailagic, A. Customizable 3D printed tactile maps as interactive overlays. In Proceedings of the Conference of 18th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2016, Reno, NV, USA, 24–26 October 2016; pp. 71–79. [CrossRef]
65. Grussenmeyer, W.; Folmer, E. AudioDraw: User preferences in non-visual diagram drawing for touchscreens. In Proceedings of the Conference of 13th Web for All Conference, Montreal, QC, Canada, 11–13 April 2016. [CrossRef]
66. Reinholt, K.; Guinness, D.; Kane, S. Eyedescribe: Combining eye gaze and speech to automatically create accessible touch screen artwork. In Proceedings of the 14th ACM International Conference on Interactive Surfaces and Spaces, ISS 2019, Daejeon, Korea, 10–13 November 2019; pp. 101–112. [CrossRef]
67. Jayant, C.; Ji, H.; White, S.; Bigham, J.P. Supporting blind photography. In Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, Dundee, UK, 24–26 October 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. *Harv. Educ. Rev.* **2013**, *83*, 67–86. [CrossRef]
69. Handa, K.; Dairoku, H.; Toriyama, Y. Investigation of priority needs in terms of museum service accessibility for visually impaired visitors. *Br. J. Vis. Impair.* **2010**, *28*, 221–234. [CrossRef]
70. Hayhoe, S. Why Do We Think That People who are Visually Impaired Don't Want to Know About the Visual Arts II? In Proceedings of the Workshop: Creative Sandpit@ the (Virtual) American Museum, Bath, UK, 9–10 November 2020; Creative Sandpit: Bath, UK, 2020.
71. Asakawa, S.; Guerreiro, J.; Ahmetovic, D.; Kitani, K.M.; Asakawa, C. The Present and Future of Museum Accessibility for People with Visual Impairments. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 382–384.
72. The Audio Guide by Metropolitan Museum of Art. Available online: <https://www.metmuseum.org/visit/audio-guide> (accessed on 20 April 2019).
73. Accessibility at Museum of Modern Art. Available online: <http://www.moma.org/learn/disabilities/sight> (accessed on 20 April 2019).
74. Art Beyond Sight Tours at Seattle Art Museum. Available online: <http://www.seattleartmuseum.org/visit/accessibility#tou> (accessed on 20 April 2019).
75. Out Loud, The Andy Warhol Museum's Inclusive Audio Guide. Available online: <https://itunes.apple.com/us/app/the-warhol-out-loud/id1103407119> (accessed on 15 April 2021).
76. Feldman, E.B. *Practical Art Criticism*; Pearson: London, UK, 1994.
77. Ponsard, C.; Fries, V. An accessible viewer for digital comic books. In Proceedings of the International Conference on Computers for Handicapped Persons, Linz, Austria, 9–11 July 2008; pp. 569–577.
78. Rayar, F.; Oriola, B.; Jouffrais, C. ALCOVE: An accessible comic reader for people with low vision. In Proceedings of the 25th International Conference on Intelligent User Interfaces, Cagliari, Italy, 17 March 2020; pp. 410–418.
79. Wang, Y.; Wang, W.; Liang, W.; Yu, L.F. Comic-guided speech synthesis. *ACM Trans. Graph. (TOG)* **2019**, *38*, 1–14. [CrossRef]
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. In Proceedings of the 2013 12th International Conference on Document Analysis and Recognition, Washington, DC, USA, 25–28 August 2013; pp. 1145–1149.
81. Figueiredo, F.; Benevenuto, F.; Almeida, J.M. The tube over time: Characterizing popularity growth of youtube videos. In Proceedings of the fourth ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011; pp. 745–754.
82. Hartley, J. Uses of YouTube: Digital literacy and the growth of knowledge. In *YouTube: Online Video and Participatory Culture*; Digital Media and Society: Cambridge, UK, 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. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, Eindhoven, The Netherlands, 6–10 July 2020; ACM: New York, NY, USA, 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. In Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, Dundee, UK, 24–26 October 2011; pp. 123–130.
85. Maneesaeng, N.; Punyabukkana, P.; Suchato, A. Accessible video-call application on android for the blind. *Lect. Notes Softw. Eng.* **2016**, *4*, 95. [CrossRef]
86. Stai, E.; Kafetzoglou, S.; Tsiropoulou, E.E.; Papavassiliou, S. A holistic approach for personalization, relevance feedback & recommendation in enriched multimedia content. *Multimed. Tools Appl.* **2018**, *77*, 283–326.

87. Pouli, V.; Kafetzoglou, S.; Tsiropoulos, E.E.; Dimitriou, A.; Papavassiliou, S. Personalized multimedia content retrieval through relevance feedback techniques for enhanced user experience. In Proceedings of the 2015 13th International Conference on Telecommunications (ConTEL), Graz, Austria, 13–15 July 2015; pp. 1–8.
88. 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. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 93. [[CrossRef](#)]
89. Wu, S.; Wieland, J.; Farivar, O.; Schiller, J. Automatic alt-text: Computer-generated image descriptions for blind users on a social network service. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, ON, USA, 25 February–1 March 2017; pp. 1180–1192.
90. Huang, X.L.; Ma, X.; Hu, F. Machine learning and intelligent communications. *Mob. Netw. Appl.* **2018**, *23*, 68–70. [[CrossRef](#)]
91. Vinyals, O.; Toshev, A.; Bengio, S.; Erhan, D. Show and tell: A neural image caption generator. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3156–3164.
92. Dai, B.; Fidler, S.; Urtasun, R.; Lin, D. Towards diverse and natural image descriptions via a conditional gan. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2970–2979.
93. Li, Y.; Ouyang, W.; Zhou, B.; Wang, K.; Wang, X. Scene graph generation from objects, phrases and region captions. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1261–1270.
94. Fischer, B.; Peine, A.; Östlund, B. The importance of user involvement: A systematic review of involving older users in technology design. *Gerontologist* **2020**, *60*, e513–e523. [[CrossRef](#)] [[PubMed](#)]
95. Schuler, D.; Namioka, A. *Participatory Design: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 1993.
96. Sanders, E.B. From user-centered to participatory design approaches. *Des. Soc. Sci. Mak. Connect.* **2002**, *1*, 1.
97. Sahib, N.G.; Stockman, T.; Tombros, A.; Metatla, O. Participatory design with blind users: A scenario-based approach. In *IFIP Conference on Human-Computer Interaction*; Springer: Berlin, Germany, 2013; pp. 685–701.