Making 3D Printer Accessible for People with Visual Impairments by Reading Scrolling Text and Menus

Naoya Tagawa
Graduate School of Informatics, Osaka Metropolitan University
Japan
sc23087x@st.omu.ac.jp

Masakazu Iwamura
Graduate School of Informatics, Osaka Metropolitan University
Japan
masa.i@omu.ac.jp

Kazunori Minatani
National Center for University Entrance Examinations
Japan
minatani@rd.dnc.ac.jp

Koichi Kise
Graduate School of Informatics, Osaka Metropolitan University
Japan
kise@omu.ac.jp

Abstract

3D printing has immense potential to enhance the lives of people with visual impairments (PVI) by enabling them to understand shapes and other details through touch that words alone cannot convey. Several initiatives have made 3D tactile models accessible to PVI, yet these models are typically created by sighted individuals. Our goal is to empower PVI to create 3D tactile models independently, making 3D printers accessible to them. The biggest bottleneck for PVI in using 3D printers is the inability to read text on their display. Our work specifically focuses on making scrolling text and menus readable. Through a user study with 13 PVI (five blind and eight with low vision), we confirmed the effectiveness of the implemented functions over conventional smartphone apps and wearable devices.

CCS Concepts

• **Human-centered computing → Accessibility systems and tools**.

Keywords

3D printer, optical character reader (OCR), scrolling text, menus

ACM Reference Format:
Naoya Tagawa, Masakazu Iwamura, Kazunori Minatani, and Koichi Kise. 2024. Making 3D Printer Accessible for People with Visual Impairments by Reading Scrolling Text and Menus. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24), October 27–30, 2024, St. John's, NL, Canada.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3663548.3688517

ASSETS '24, October 27–30, 2024, St. John's, NL, Canada

ACM ISBN 979-8-4007-0677-6/24/10
https://doi.org/10.1145/3663548.3688517

1 Introduction

"Objects such as snowflakes, castles, and butterflies have become more than just words when explored as a 3D print." This impressive sentence appears in the very beginning of [10], which reports the activities of See3D [17], a non-profit organization that manages the printing and distribution of 3D-printed models for people who are blind. Similar activities include [2, 3]. 3D printing has immense potential to enhance the lives of people with visual impairments (PVI). 3D-printed tactile objects are invaluable sources of information for PVI, enabling them to comprehend shapes and other details through touch that words alone cannot convey [2, 17]. These objects can also serve as effective educational tools, which greatly benefit students with visual impairments [3, 12].

However, there is little literature (e.g., Minatani [12]) and few Web sites (e.g., Round Table [1]) on the use of 3D printers by PVI to print 3D models. There are several benefits for PVI using 3D printers [12]. The most significant is the ability to independently produce desired objects. 3D printing enables PVI to create the objects they want on their own, without depending on sighted people to interpret their descriptions, which can often be challenging and inaccurate. Another key benefit is the potential for personalized assistive devices. 3D printers allow individuals with visual and other impairments to design and create customized assistive technologies tailored to their specific needs. This is crucial as traditional assistive devices often do not meet individual requirements due to the relatively small disabled population.

Therefore, we aim to make 3D printers accessible for PVI. Methods that allow PVI to read text on home appliance displays have been extensively researched [4, 7–9, 11, 14–16, 18]. The major concerns of previous research include how to find the location of the text, how to take a good-quality photograph of the text, and how to read the text well. However, our work addresses completely different issues: scrolling text and menus. We implemented reading functions for them, which, to the best of our knowledge, is a novel contribution. Our method for scrolling text is simple yet effective, making it easily applicable to other OCR engines in smartphone apps and wearable devices. We implemented a prototype system for accessing a 3D printer console for its potential needs of PVI (e.g., [13]). A user study involving 13 PVI (five blind and eight with low

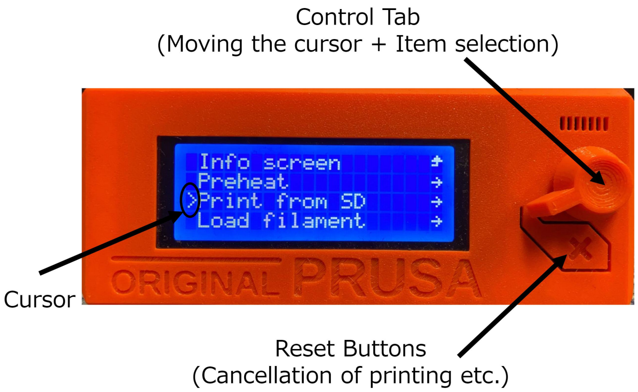

Figure 1: The console of 3D printer, which consists of a liquid crystal display (LCD) on the left and a control knob with a reset button directly under the control knob on the right.

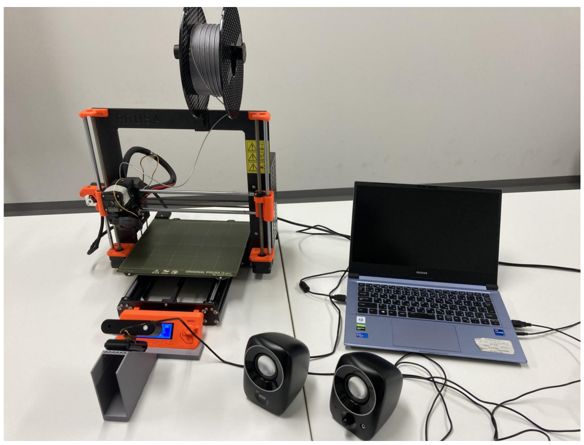

Figure 2: Overview of the proposed system, consisting of a camera mounted on the stand, a speaker, and a laptop computer.

vision) confirmed the effectiveness of the functions implemented over conventional smartphone apps and wearable devices.

The contributions of this paper are summarized below.

(1) To the best of our knowledge, we are the first to implement a function to read scrolling text. Our method is simple yet effective, making it easily applicable to other OCR engines in smartphone apps and wearable devices.
(2) To the best of our knowledge, we are the first to implement a function that reads the selected menu item.
(3) Through a user study with 13 PVI, we demonstrate that the proposed functions significantly enhance the ability of PVI to operate the 3D printer console.

2 Proposed System

The proposed system is tailored for Original Prusa i3 MK3, a consumer-grade 3D printer. The 3D printer has a console panel, as shown in Fig. 1, which consists of an LCD on the left and a control knob with a reset button directly under the control knob on the right. The knob can be turned and pushed. When turning the right knob clockwise or counterclockwise, a clicking sound is played and the cursor moves one step downward or upward. Pushing the knob makes a selection. The LCD displays white text on a blue background, showing four lines vertically and 20 characters horizontally. The LCD is surrounded by a black frame, which is encased in an orange plastic housing. Fig. 2 shows an overview of the proposed system attached to the 3D printer. The system consists of a web camera, a speaker, and a laptop computer. The web camera is mounted on the plastic stand printed by the 3D printer. The stand, printed using the 3D printer itself, ensures that the camera is positioned to capture the display head-on when fitted into the console.

2.1 Reading function for menu

Reading menus presents three main challenges:

(1) **Cursor position detection:** The text that should be read out is adjacent to the cursor. Therefore, the first step is to detect the cursor, and then identify the text adjacent to it. The system reads out the text only when the cursor moves. Therefore, when the movement of the cursor is detected, the system determines its position and reads out the item pointed to by the cursor. This process is realized by character recognition and its simple post-process to find the adjacent text. When reading out, the system speaks "The cursor points to" before the item pointed to by the cursor.
(2) **Vertical scroll detection:** If the number of menu items exceeds the number of vertical lines on the display, some menu items will extend beyond the visible screen area. When the cursor is at the top of the screen and the user attempts to move it further up, the screen will scroll down to reveal the new item, unless the cursor is already pointing to the first item in the menu list. The same principle applies when the screen scrolls upward. This process is realized by change detection.
(3) **Screen transition detection:** Transitioning to a new level in the menu hierarchy should be detected. Pressing the knob causes a transition to a new screen displaying more detailed information about the selected item. This process is also realized by change detection.

2.2 Reading function for horizontal scrolling text

Reading scrolling text poses a unique challenge since the entire text cannot be seen at once. Therefore, we need to recognize the visible portion of the text in each frame (say, partial text string) and concatenate them to form a complete message. We refer to this concatenated text across multiple frames as the *integrated text*. Although text recognition techniques are generally accurate, they are inherently prone to errors. Therefore, our method must be robust against recognition errors while avoiding the consumption of large computational resources. To address these requirements, we propose a simple and efficient method. Our method first aligns the partial text strings and then selects the character that occurs most frequently at each position, which form the integrated text. However, the partial text string sequentially arrives one by one.

Therefore, our method sequentially concatenates partial text strings by alternating between two key processes: text string alignment and text concatenation. In the text string alignment process, the method finds the most appropriate position to align the partial text string from the current frame with the integrated text. This ensures that the new partial text string is correctly positioned in relation to the already concatenated text. Once the text is aligned, the text concatenation process merges the new text string with the integrated text. To ensure robustness against recognition errors, we employ a voting mechanism during the merging process. That is, we prepare a voting table for each position and take the character that occurs most frequently as the representative of the position. This voting system helps to minimize the impact of individual recognition errors on the final integrated text. By iteratively applying these two processes, our method effectively reads scrolling text while maintaining efficiency and robustness.

3 User study

We conducted a user study with 13 PVI: five blind and eight with low vision, seven males and six females, whose ages ranged from 34 to 77, with an average age of 51.8 years. We started with a pre-study interview to collect participants' demographic information and understand the participants' experience with smartphone apps and wearable devices that help to obtain visual information. Then, we conducted two sessions to evaluate two reading functions of our prototype system comparing with the participants' preferred method from among their eyes for those with low vision and the smartphone apps and wearable devices they selected. Finally, we conducted a post-study interview to supplement our quantitative results and to gain further insight into user preferences and needs.

3.1 Study design

In the first session (menu selection tasks), we prepared three different operation tasks, which require 1-, 2-, and 3-step operations, respectively. Before each task, we informed the participants of the items to select and the order to follow. We asked participants to perform the tasks using the proposed method and the conventional method. As a combination of three tasks and two methods, there were six experiments in total. The order of these experiments was randomly shuffled for each participant. The time required to complete the operation was measured. If participants felt that they could not complete the task even with more time, or they felt they were so tired, we allowed them to quit the task prematurely to reduce the participant's workload. In such a case, we simply considered the time of quitting as the time taken to complete the task. For tasks spanning multiple screens, we immediately pointed out any wrongly selected item, asked the participant to return to the previous screen, and resumed the experiment.

In the second session (scrolling text), participants are asked to hear the file names on an SD card that gradually scroll horizontally over time using the proposed system and the conventional method separately. We asked a question to ask whether the function is useful. Initially, we had planned to ask participants to estimate the file names using two methods, respectively and compare the required times and accuracies. However, due to the extreme difficulty of using the conventional method for this task, we decided to abandon this plan.

Table 1: Session 1 results: mean and standard deviation of the time distributions in seconds. The p-value was calculated using a one-sided Mann-Whitney U test.

Task	Proposed	Conventional	p-value
	Mean ± SD	Mean ± SD	
Task 1	**25.5** ± 12.7	273.6 ± 302.5	0.01563
Task 2	**88.8** ± 81.2	357.8 ± 316.2	0.02012
Task 3	**153.7** ± 49.5	723.2 ± 563.0	0.001468

3.2 Result

In the first session (menu selection tasks), we compare the time taken to complete the tasks of selecting specific menu items. Table 1 shows the mean and standard deviation of the time distributions in both methods for each task for each method. The table indicates that the proposed method completed all tasks in a shorter time than the conventional method. This is because the proposed method always read out the item pointed to by the cursor, without requiring adjustment of the camera angle. On the other hand, the conventional method required the participants to adjust the camera angle by themselves. Furthermore, the smartphone apps participants used had difficulty in recognizing the cursor, hence requiring multiple attempts to recognize it before the app read out the cursor. Some participants judged that the smartphone app did not read out the cursor. Therefore, they tried to memorize the order of the items by saving what the app read in a memo app. However, there were exceptional cases: two participants had low vision and could see the text on the display with their own eyes, so they operated while looking at the display. In their cases, the conventional method (i.e., their own eyes) required a shorter time to complete the tasks than the proposed method. Table 1 also shows the p-values. Since the number of samples was small and normality could not be assumed, we conducted a Mann-Whitney U test to see if there was a significant difference between the two methods. As a result, the p-value was less than 0.05 for all tasks, which can be said that the time required to complete the tasks in the proposed system was significantly shorter than the conventional method.

In the second session (scrolling text), we present the results of an interview to see if the proposed system could be used to understand file names on an SD card that scroll horizontally. We asked the question, *"Do you think concatenating horizontally scrolling text into a single sentence would help you understand the file name that scrolls horizontally?"* Seven participants (53.8%) strongly agreed, three (23.1%) agreed, two (15.4%) answered neutral, and one (7.7%) disagreed.

4 Discussion

4.1 How to improve the proposed method

Feedback from participants in the user study provided valuable insight on how to improve the proposed method. The suggestions include reducing silent durations, improving response speed, and reading only the information pointed to by the cursor when navigating menus. Interestingly, these desired improvements align closely with the behavior of screen readers. It is uncertain whether this preference stems from the participants' familiarity with screen readers or because screen readers have been well-developed over time.

4.2 Applying the proposed method beyond 3D printers

The proposed method is tailored for a specific model of 3D printers, which may raise concerns about its applicability to other models or home appliances. Although custom setup is required for each target device, efforts can be made to reduce adjustments. The following functions need to be adapted for broader applicability. First, the character recognition engine in our system is trained on character images obtained from the 3D printer screen. To generalize the OCR part, we may rely on the latest techniques in scene text recognition. Second, differences in screen layout must be addressed. As the layout of the screen varies between models and appliances, manual customization may be necessary. However, labor can be reduced by modularizing functions. Advances in generative AI, such as ChatGPT [5] and Claude-3 [6], could automate this problem to some extent. For example, if the manual for the target model is distributed in PDF format, it would be possible to customize the system with minimal human intervention by loading the manual into a generative AI. We look forward to such a future where AI can significantly reduce the effort required for customization.

5 Conclusion

Our goal is to empower PVI to create 3D tactile models independently. This autonomy allows them to produce desired objects without relying on sighted individuals, overcoming the challenges and inaccuracies of verbal descriptions. The biggest bottleneck for PVI in using 3D printers independently is the inability to read text on their display. Therefore, this paper specifically focuses on making 3D printers accessible by addressing the challenge of reading scrolling text and menus. Our method for scrolling text is simple yet effective, making it easily applicable to other OCR engines in smartphone apps and wearable devices.

We conducted a user study with 13 PVI, comparing our prototype system with their preferred methods, including smartphone apps, wearable devices, and their own eyes if they have sight. In the experiment of the reading function for menus, most participants completed the tasks faster using our prototype system and preferred it, except for those who used their own eyes. In that for scrolling text, all participants, including those who used their own eyes, recognized the advantages of our prototype system and preferred it. Some participants expressed their desire to see the proposed reading function for scrolling text implemented in smartphone apps. Some suggested improvements to the user interface of the prototype system, such as eliminating silent durations and reducing the response time after user operations. Interestingly, most of their requests aimed to make the behavior of the system similar to that of screen readers.

Acknowledgments

This research was supported by JST RISTEX "Solution-Driven Co-creative R&D Program for SDGs (SOLVE for SDGs): Solution Creation Phase," Grant Number JPMJRX21I5.

References

[1] 3D model selection, design and printing by people who are blind or have low vision — Round Table [n. d.]. . Retrieved April 24, 2024 from https://printdisability.org/about-us/accessible-graphics/3d-printing/blvmakers/

[2] 3D4SDGs: 3D Models for People with Visual Impairment [n. d.]. . Retrieved July 1, 2024 from https://3d4sdgs.net/

[3] 3D4VIP: 3D Printing for Education for Visually Impaired Students [n. d.]. . Retrieved July 1, 2024 from https://www.blista.de/EU-Projekt-3D4VIP

[4] Chiraphat Boonnag, Piyalitt Ittichaiwong, Wanumaidah Saengmolee, Narongrid Seesawad, Amrest Chinkamol, Saendee Rattanasomrerk, Kanyakorn Veerakanjana, Kamonwan Thanontip, Warissara Limpornchitwilai, and Theerawit Wilaiprasitporn. [n. d.]. PACMAN: A Framework for Pulse Oximeter Digit Detection and Reading in a Low-Resource Setting. 10, 15 ([n. d.]), 13196–13204. https://doi.org/10.1109/jiot.2023.3262205

[5] ChatGPT [n. d.]. . Retrieved April 25, 2024 from https://chat.openai.com/

[6] Claude-3 [n. d.]. . Retrieved April 25, 2024 from https://claude.ai/

[7] E. Finnegan, M. Villarroel, C. Velardo, and L. Tarassenko. [n. d.]. Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors. 43, 6 ([n. d.]), 341–355. https://doi.org/10.1080/03091902.2019.1673844

[8] Giovanni Fusco, Ender Tekin, Nicholas A. Giudice, and James M. Coughlan. 2015. Appliance Displays: Accessibility Challenges and Proposed Solutions. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '15)*. ACM Press. https://doi.org/10.1145/2700648.2811392

[9] Giovanni Fusco, Ender Tekin, Richard E. Ladner, and James M. Coughlan. 2014. Using computer vision to access appliance displays. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers and Accessibility (ASSETS '14)*. ACM Press. https://doi.org/10.1145/2661334.2661404

[10] Caroline Karbowski. [n. d.]. See3D: 3D Printing for People Who Are Blind. 23, 1 ([n. d.]). https://doi.org/10.14448/jsesd.12.0006

[11] Thotreingam Kasar. [n. d.]. *Recognition of Seven-Segment Displays from Images of Digital Energy Meters*. Springer Singapore, 1–10. https://doi.org/10.1007/978-981-13-2514-4_1

[12] Kazunori Minatani. 2017. An Analysis and Proposal of 3D Printing Applications for the Visually Impaired. In *Harnessing the Power of Technology to Improve Lives, Proceedings of the 14th European Conference on the Advancement of Assistive Technology, AAATE Conf. 2017, Sheffield, UK, September 12-15, 2017 (Studies in Health Technology and Informatics, Vol. 242)*, Peter Cudd and Luc P. de Witte (Eds.). IOS Press, 918–921. https://doi.org/10.3233/978-1-61499-798-6-918

[13] Kazunori Minatani. 2017. *Finding 3D CAD Data Production Methods that Work for People with Visual Impairments*. Springer International Publishing, 548–554. https://doi.org/10.1007/978-3-319-58750-9_76

[14] Lucas P. Moreira. 2022. Automated Medical Device Display Reading Using Deep Learning Object Detection. https://doi.org/10.48550/ARXIV.2210.01325

[15] Tim Morris, Paul Blenkhorn, Luke Crossey, Quang Ngo, Martin Ross, David Werner, and Christina Wong. [n. d.]. Clearspeech: A Display Reader for the Visually Handicapped. 14, 4 ([n. d.]), 492–500. https://doi.org/10.1109/tnsre.2006.881538

[16] Irati Rasines, Pedro Iriondo, and Ibai Díez. [n. d.]. *Real-Time Display Recognition System for Visually Impaired*. Springer Berlin Heidelberg, 623–629. https://doi.org/10.1007/978-3-642-31534-3_91

[17] See3D [n. d.]. . Retrieved July 2, 2024 from https://see3d.org/

[18] Ender Tekin, James M. Coughlan, and Huiying Shen. [n. d.]. Real-time detection and reading of LED/LCD displays for visually impaired persons. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)* (2011-01). IEEE. https://doi.org/10.1109/wacv.2011.5711544