

Evaluating Author and User Experience for an Audio-Haptic System for Annotation of Physical Models

James M. Coughlan and Joshua Miele

The Smith-Kettlewell Eye Research Institute, San Francisco, CA 94115 USA
{coughlan, jam}@ski.org

ABSTRACT

We describe three usability studies involving a prototype system for creation and haptic exploration of labeled locations on 3D objects. The system uses a computer, webcam, and fiducial markers to associate a physical 3D object in the camera's view with a pre-defined digital map of labeled locations ("hotspots"), and to do real-time finger tracking, allowing a blind or visually impaired user to explore the object and hear individual labels spoken as each hotspot is touched. This paper describes: (a) a formative study with blind users exploring pre-annotated objects to assess system usability and accuracy; (b) a focus group of blind participants who used the system and, through structured and unstructured discussion, provided feedback on its practicality, possible applications, and real-world potential; and (c) a formative study in which a sighted adult used the system to add labels to on-screen images of objects, demonstrating the practicality of remote annotation of 3D models. These studies and related literature suggest potential for future iterations of the system to benefit blind and visually impaired users in educational, professional, and recreational contexts.

CCS Concepts

• Human-centered computing→Accessibility

Keywords

Blindness; low vision; 3D models; tactile graphics.

1. INTRODUCTION

Many people who are blind or visually impaired have insufficient access to a wide range of everyday objects needed for daily activities that require visual inspection on the part of the user. Such objects include printed documents, maps, infographics, appliances and 3D models (e.g., 3D-printed) used in STEM education, and are abundant in schools, the home and the workplace.

A common way of enhancing an object's accessibility is to affix braille labels to its surface. However, there is limited space available for braille labelling and it is inaccessible to those who don't read braille. A powerful alternative is the use of audio labels, which enable audio-haptic exploration of the object, whereby the user explores the surface naturally with the fingers and queries locations of interest. Audio labels can be associated with specific locations on tactile models and graphics in a variety of ways, including by placing touch-sensitive sensors at locations of interest, overlaying a tactile graphic on a touch-sensitive tablet, or using a camera-enabled "smart pen" that works with special materials [4]. While effective, these methods for producing audio labels require

SAMPLE: Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.
Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/12345.67890>

each object to be customized with special hardware and materials, which is costly and greatly limits their adoption.

By contrast, computer vision-based approaches, in which a camera tracks the user's hands as they explore an object, have the advantage of offering audio-haptic accessibility for *existing* objects with minimal or no customization. Several projects have employed this computer vision approach over the years, including Access Lens [3] for facilitating OCR, the Tactile Graphics Helper [1] to provide access to tactile graphics to students in an educational context, and VizLens [2] for making appliances accessible.

Our "CamIO" (short for "Camera Input-Output") project applies this computer vision approach to provide audio-haptic access to arbitrary 3D objects, including 3D models, tactile graphics, relief maps and appliances. It builds on our earlier work [5], but now uses a conventional camera (like Magic Touch [6]) instead of relying on a depth camera, and includes a new tool to create annotations for any rigid 3D object (not just 3D printed ones).

We describe recent developments, including improvements to the end-user software and a new tool to facilitate the creation of annotations specifying the information available for an object. Three user studies were performed, including a formative study to suggest possible UI improvements, a focus group study to assess the ways that people who are blind would want to use CamIO and another formative study to verify the usability and effectiveness of the annotation tool. The results of the user studies are discussed and highlight directions for future development of CamIO.

2. APPROACH

CamIO was conceived by one of the co-authors (who is blind) and has been iteratively improved based on ongoing user feedback with a user-centered research and development approach. Blindness rehabilitation professionals stressed the inconvenience of the Microsoft Kinect depth camera platform, so we modified our approach to accommodate a standard camera.

The new CamIO system consists of two main components: (a) the Explorer system (Fig. 1a) for end users who are blind or visually impaired, and (b) the Creator annotation tool that allows a sighted assistant to define the *hotspots* (locations of interest on the object) and the text labels associated with them. Both components are implemented on a standard laptop using its built-in camera; a preliminary version of Explorer also runs on Android.

Explorer uses computer vision to track the location of the user's *pointer finger* (usually the index finger), and whenever the fingertip nears a hotspot, the text label associated with the hotspot is announced using text-to-speech. Note that the user can explore the object naturally with one or both hands, but receives audio feedback on the location indicated by the pointing finger. Computer vision processing is simplified using a special black-and-white "fiducial" pattern board (Fig. 1a), which makes it easy to estimate the object's 3D pose (rotation and translation relative to the camera), and a multi-color fingertip marker (Fig. 1b).

Creator allows a sighted assistant to define digital annotations, which specify the hotspot locations and associated text labels for an object of interest. The assistant indicates hotspot locations by clicking on the appropriate location in two different images of the object taken from different vantage points and typing in the corresponding text labels. This process allows the annotations to be authored remotely, even without physical access to the object.



Figure 1. (a) CamIO system, showing laptop, object (relief map) and fiducial board. (b) Close-up of fingertip marker.

3. USER STUDIES

The first user study was a formative study we conducted with three completely blind volunteer participants (two male and one female, ages from 23-32). All three participants were able to use Explorer to independently read aloud the hotspots on a 3D-printed airplane model and a model of the Chrysler building, and expressed enthusiasm for the project. They suggested bug fixes and improvements, which were later implemented, including a new dwell gesture to trigger the announcement of additional hotspot information (when available).

Next, a focus group study was conducted to assess how blind adults are likely to use CamIO, which functions they want, and how they should be implemented. The study recruited five blind adults with ages ranging from 29-55 who had experience with braille reading, tactile graphics or haptic exploration and employment history as a blind adult. The participants were shown how to use Explorer, and used it for themselves to explore two objects: the 3D printed airplane and a Mackie 8-channel audio mixer.

Several dominant themes emerged from the study. (1) **Common use cases** include independent exploration of maps and models, access to appliances in the home or workplace, and educational applications including museum exhibits. (2) **New UI features**: better feedback was desired to help with proper positioning of the object relative to the camera, an audio signal indicating when the pointer finger exits a hotspot region, explicit feedback indicating when a secondary label is available, and the use of additional finger gestures. (3) **New authoring capabilities**: accessible authoring functions (including the ability to customize existing labels), the use of a stylus to define hotspot locations, and the use of crowdsourcing approaches to obtain annotations. (4) **Strengths**: the software-defined annotation approach that “[separates] accessibility from the original design of the object,” a “natural” and “responsive” interface that permits “fluid exploration without interruption” and the presence of multiple layers of information. (5) **Weaknesses**: visibility problems (especially maintaining the visibility of the fingertip marker), the incidence of incorrect feedback, and the need for annotations for each object of interest.

A final formative study focusing on annotation was conducted with one volunteer participant, an adult female with normal vision. We trained her to use Creator and then had her annotate two objects by herself. For each object she was given a list of five hotspots to annotate, including the label name for each hotspot. (The Creator tool was pre-loaded with the two views of each object.) She took approximately 5-10 minutes to complete the annotations for each object, which we verified were correct. Note that she was able to

complete the annotations independently without holding the objects or seeing them in person. She conveyed valuable feedback about Creator that we will address in the future, including her wish for a better understanding of how accurately she needed to click the mouse to specify the hotspot locations, and for a zoom feature to help her resolve fine spatial details.

4. CONCLUSIONS AND FUTURE WORK

We have described recent developments of the CamIO project, including improvements to Explorer, the new Creator tool, and user studies. The new CamIO approach has the advantage that it will be able to operate on any standard platform (both mobile and laptop) without the need for an external depth camera. The Creator tool allows a sighted person to perform remote annotation of an object without the need for physical access to it.

Overall, the user studies confirmed our belief that the CamIO project is likely to be of great interest to many people who are blind and that it has real-world potential. The studies suggest many avenues for future development, including UI improvements and new functionalities such as making authoring functions accessible.

Future work will include continuing to port the CamIO system to the smartphone/tablet platform. In particular, Creator will be ported to a browser platform, which will not only eliminate the inconvenience of downloading and installing the tool but also opens the possibility of crowdsourcing the annotation process, so that multiple people can contribute annotations for a single object.

We are also exploring possible applications of the CamIO approach for users who have low vision, who may prefer a visual interface that displays a close-up view of the current hotspot region (with the option of digitally rendering this view without the fingers obscuring it), instead of (or in addition to) an audio interface.

5. ACKNOWLEDGMENTS

This work was supported by NIH grants 5R01EY025332 and 2 R01 EY018890, and NIDILRR grant 90RE5024-01-00.

6. REFERENCES

- [1] Fusco, G., & Morash, V. S. (2015). The tactile graphics helper: providing audio clarification for tactile graphics using machine vision. *Proc. 17th Int. 'l ACM SIGACCESS Conf. on Computers & Accessibility* (pp. 97-106). ACM.
- [2] Guo, A., Chen, X. A., Qi, H., White, S., Ghosh, S., Asakawa, C., & Bigham, J. P. (2016). VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World. *Proc. 29th Annual Symposium on UI Software & Tech.* ACM.
- [3] Kane, S. K., Frey, B., & Wobbrock, J. O. (2013). Access lens: a gesture-based screen reader for real-world documents. *Proc. SIGCHI Conference on Human Factors in Computing Systems* (pp. 347-350). ACM.
- [4] Miele, J. (2010). Talking Tactile Apps for the Pulse Pen: STEM Binder. *25th Annual Int. 'l Tech. & Persons with Disabilities Conference (CSUN)*, March 2010.
- [5] Shen, H., Edwards, O., Miele, J., & Coughlan, J. M. (2013). CamIO: A 3D computer vision system enabling audio/haptic interaction with physical objects by blind users. *Proc. 15th Int. 'l ACM SIGACCESS Conf. on Computers and Accessibility* (p. 41). ACM.
- [6] Shi, L., McLachlan, R., Zhao, Y., & Azenkot, S. (2016). Magic Touch: Interacting with 3D Printed Graphics. *Proc. 18th Int. 'l ACM SIGACCESS Conf. on Computers and Accessibility* (pp. 329-330). ACM.