Supporting Older Adults in Using Complex User Interfaces with Augmented Reality

Junhan Kong, Anhong Guo, Jeffrey P. Bigham
Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA
junhank@andrew.cmu.edu, {anhongg, jbigham}@cs.cmu.edu

ABSTRACT
Using complex interfaces have shown to be challenging for older adults. Existing tutorial systems can be cumbersome, and sometimes difficulty to use. To solve this problem, we present a system to support older adults in using visual interfaces by providing step-by-step visual guidance with augmented reality. Using the Apple ARKit platform, our system detects the interface in a phone camera view, and provides visual guidance for users to access the interface following a generated sequence of interactions based on pre-specified tasks and prior knowledge of the interface.

ACM Classification Keywords
H.5.2 Information interfaces and presentation: User Interfaces - *Input devices and strategies*; K.4.2 Computers and Society: Social Issues - *Assistive technologies*.

Author Keywords
Augmented reality; interfaces; older adults; cognitive support.

INTRODUCTION
Older adults have been shown to face challenges in using interfaces that are complex or unfamiliar to them [5], such as ticket vending machines, medical devices, etc. As preference for trial-and-error decreases with age, having older adults explore interfaces by themselves would be less favorable. In the meantime, although the use of instruction manual is more preferred among older adults, many find it difficult to use [7]. Furthermore, as design principles to accommodate older adults (e.g., larger fonts and more contrast) [2] are only followed in a limited set of interfaces, interacting with interfaces that are complex or unfamiliar to them could be a challenging task.

However, re-designing existing interfaces for older adults would take tremendous effort and are generally associated with huge costs. Surveys have also shown that predictability of usage and same patterns of interaction are considered crucial in making interfaces universally accessible [6]. Thus, designing a consistent guidance system across different interfaces to assist use of unfamiliar interfaces for older adults could potentially improve user experience and be favorable.

ASSETS'19, October 28–30, 2019, Pittsburgh, PA, USA

DOI: http://dx.doi.org/10.475/123_4

Many prior work proposed solutions in making smart mobile devices and web applications more accessible to older adults, such as Sato *et al.*'s work [8] on using voice augmentation to support elderly web users. Some prior work also approached the problem of interface accessibility for older adults by designing for a specific type of interface, e.g. ticket vending machines [9], yet they do not generalize. When it comes to a new interface or control users have never encountered, learning is still required, which increases the cognitive burden.

In this demo, we present a system to support older adults in using complex user interfaces by providing step-by-step visual guidance in an augmented reality setting. Instead of having to explore an unfamiliar interface through trial and error, users can simply answer a few questions to specify the needed input parameters for a task, then follow the visual guidance to complete it. Using Apple ARKit [1], our system detects the interface state in the user's camera view, generates an action sequence for the user to complete based on prior knowledge of the interface, and displays corresponding visual indicators on the phone screen to guide the users in performing the actions. As the application can be generalizable to a large variety of 2D and potentially 3D visual interfaces, the step-by-step guidance would be similar across interfaces, reducing the burden of learning for older adults.

IMPLEMENTATION
Our system provides cognitive support for older adults in using visual interfaces in a step-by-step manner, through allowing users to pre-specify tasks, recognizing state of the interfaces and where the users are in the process, then overlaying on top of interfaces in the AR scene to provide visual guidance.

Knowledge of existing interfaces can be provided by manufacturers, or extracted in a way similar to that in StateLens [4], which analyzes point-of-view videos to extract reference images for different interface screens and user interactions through a crowd-computer vision pipeline, thus building a state diagram for an interface. Users can first specify tasks using a method similar to the StateLens voice agent. Then, the system uses the user's pre-specified task to compute an interaction sequence, e.g., "press button A on the first screen", and "swipe left on the second screen." The system then detects the current interface and any state changes, indicating the user's progress in completing the task. With information about the current state and user's progress, it displays visual indicators in the AR scene to inform the user about the current interface and serve as a feedforward to guide the user to perform the target interaction, such as clicking a button.

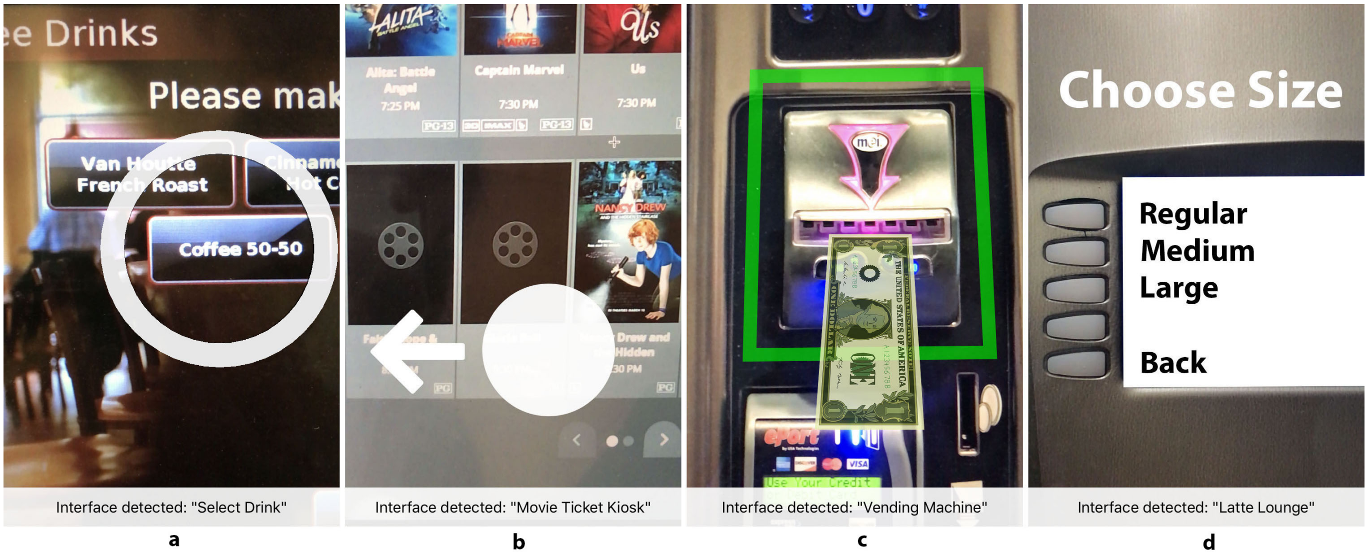

Figure 1. Design of visual guidance for different types of interactions: *(a)* On a coffee machine interface, a floating circle is displayed around the target button "coffee 50-50"; *(b)* On a movie ticket kiosk, an animated circle moving towards the target swipe direction is displayed to guide the user to swipe to the next page of movie list; *(c)* On a snack vending machine interface, an animation of inserting bills is displayed with the area highlighted, in order to guide the user to complete the payment; *(d)* On a text-heavy coffee machine, an overlay with bigger fonts and higher contrast is displayed on top of the original interface to make it more readable.

Recognizing User Interaction

Generally, the system recognizes user interactions by matching the current phone camera view with reference images using the Apple ARKit. For interfaces with dynamic layout and contents, e.g. ATM machines, a user interaction can be easily inferred from a change of screen display. For interfaces that are mostly static, e.g. microwave control panels whose change of status is only reflected in a small LCD screen, this could be more difficult. We are exploring solving this problem by focusing on the area with change of text in the view, and inferring the current status of interface from the text using a method similar to the VizLens LCD display reader [3], and potentially acquiring more knowledge on functionalities of different parts in the interface layout while gathering information about them. User interaction on this type of interfaces can also be potentially captured by combining changes of the visual interface with voice feedback heard by the phone microphone, e.g. a beeping sound from the microwave when a button is pressed. In this way, the system knows the user's performed interactions and thus the user's process in completing the whole task, and can then provide guidance specific to this current step.

Providing Visual Guidance

We designed different visual indicators to guide users for different types of interactions and with different interface elements, including buttons, toggles, sliders, and other physical interactions. The system also provides information support to make the interface more readable for older adults.

Buttons and toggles that require a simple click by the user are among the most widely used interface elements, both for static and dynamic interfaces. For the click interaction, an animated circle is displayed around the target button or toggle to click in the AR scene (Figure 1a).

User interactions in some interfaces, especially touchscreen interfaces, may also involve gestures. A slider (or swipe action) requires users to tap on certain parts of the interface and then move their finger. For the dragging or swiping interaction, an animated circle repeatedly appears at the initial tap location and then moves towards the target direction (Figure 1b).

Some tasks would also require actions other than interacting with interface elements, such as inserting bills or swiping credit card on a vending machine. In these scenarios, the target area is highlighted in the scene to raise the user's attention (Figure 1c), and then text will be displayed on the phone screen to clearly specify what the user needs to do, for example, "insert a $1 bill" or "swipe your credit card."

Aside from providing feedforward to guide users to perform interactions, the system also overlays on top of the existing interface in the AR scene to transform the interface in reality into one that's more easily readable and understandable by older adults. This involves using bigger fonts and higher contrast (Figure 1d), covering irrelevant sections to allow users to focus on the parts to interact with, adding descriptive icons or text descriptions, etc.

FUTURE WORK

We will explore ways to refine the system design and improve the accuracy in identifying user interactions, potentially by capturing non-visual feedback from interfaces such as beeping and by incorporating hand detection to track finger location and thus infer performed interaction. To improve usability, we will further iterate on the visual design to make it more intuitive and easily understandable for older adults. We also plan to conduct user studies with older adults to understand the effectiveness and performance of our system over traditional approaches such as tutorial manuals.

REFERENCES

1. Apple Inc. 2019. Get ready for the latest advances in augmented reality. (2019). https://developer.apple.com/augmented-reality/
2. Miranda A Farage, Kenneth W Miller, Funmi Ajayi, and Deborah Hutchins. 2012. Design principles to accommodate older adults. *Global journal of health science* 4, 2 (2012), 2.
3. Anhong Guo, Xiang 'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P. Bigham. 2016. VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 651–664. DOI: http://dx.doi.org/10.1145/2984511.2984518
4. Anhong Guo, Junhan Kong, Michael Rivera, Fangzheng Xu, and Jeffrey P Bigham. 2019. StateLens: A Solution to Making Existing Dynamic Touchscreens Accessible. In *Proceedings of the 32th Annual Symposium on User Interface Software and Technology (UIST '19)*. ACM.
5. Sri Kurniawan. 2006. An Exploratory Study of How Older Women Use Mobile Phones. In *Proceedings of the 8th International Conference on Ubiquitous Computing (UbiComp'06)*. Springer-Verlag, Berlin, Heidelberg, 105–122. DOI:http://dx.doi.org/10.1007/11853565_7
6. Michael Leitner, Özge Subasi, Norman Höller, Arjan Geven, and Manfred Tscheligi. 2009. User Requirement Analysis for a Railway Ticketing Portal with Emphasis on Semantic Accessibility for Older Users. In *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A) (W4A '09)*. ACM, New York, NY, USA, 114–122. DOI: http://dx.doi.org/10.1145/1535654.1535683
7. Rock Leung, Charlotte Tang, Shathel Haddad, Joanna Mcgrenere, Peter Graf, and Vilia Ingriany. 2012. How Older Adults Learn to Use Mobile Devices: Survey and Field Investigations. *ACM Trans. Access. Comput.* 4, 3, Article 11 (Dec. 2012), 33 pages. DOI: http://dx.doi.org/10.1145/2399193.2399195
8. Daisuke Sato, Masatomo Kobayashi, Hironobu Takagi, Chieko Asakawa, and Jiro Tanaka. 2011. How Voice Augmentation Supports Elderly Web Users. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11)*. ACM, New York, NY, USA, 155–162. DOI: http://dx.doi.org/10.1145/2049536.2049565
9. Michael Sengpiel. 2016. Teach or Design? How Older Adults&Rsquo; Use of Ticket Vending Machines Could Be More Effective. *ACM Trans. Access. Comput.* 9, 1, Article 2 (Oct. 2016), 27 pages. DOI: http://dx.doi.org/10.1145/2935619