VisionAI - Shopping Assistance for People with Vision Impairments

Category: Research

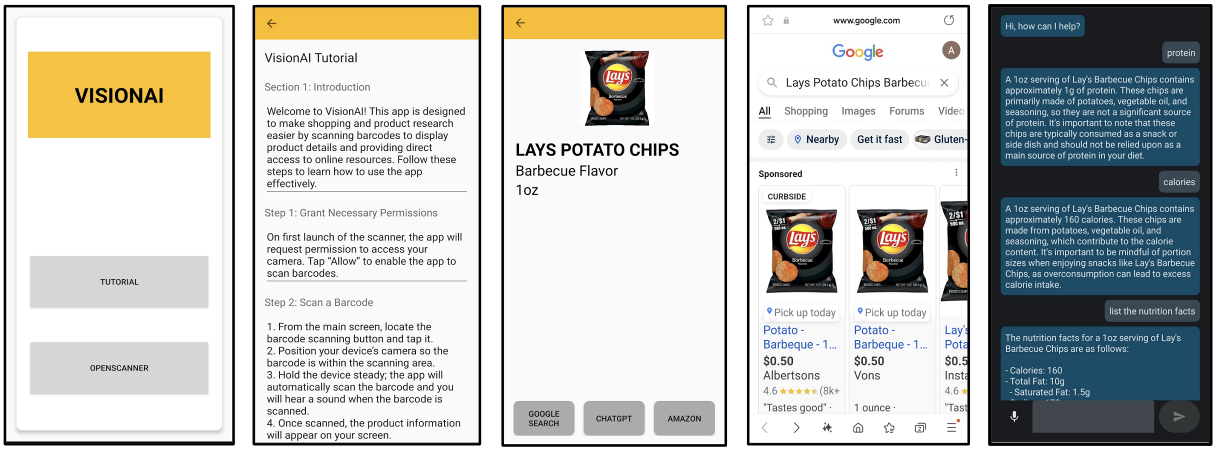

Figure 1: This figure illustrates the user interface of the VisionAI application, showing the main screen, tutorial, product information screen, Google search integration, and ChatGPT interaction from left to right.

Abstract

VisionAI is a mobile application designed to enhance the in-store shopping experience for individuals with vision impairments. Existing AI-supported apps either provide too much or too little information leading to frustration. Using the typical assistive technology design approach, we co-designed our app with a blind graduate student, focusing on the specific challenges with existing apps. While tailored to one individual's needs, we believe the resulting app addresses broader limitations in existing shopping applications used by people with vision impairments. Here we present the design process and feedback from our co-designer that indicates VisionAI's potential in empowering people and fostering a more inclusive shopping environment.

Index Terms: vision impairments, shopping, mobile app

1 Introduction

According to the American Foundation for the Blind[1], 3.89 million adults in the United States have difficulty seeing, even when wearing glasses, and 340,000 adults cannot see at all. However, technological advancements have made it possible for these people to perform many of the same tasks as fully-sighted people. Applications like Be My Eyes[2] and Aira [3] pair you with a sighted assistant while SeeingAI[4] and Google Lookout provide information about various products, scenes, text, and more. Screen readers provide assistance in navigating mobile devices in general. Individuals with visual impairments often encounter difficulties in reading product labels due to small print sizes, inadequate contrast between text and background, and other visual challenges. Magnifier applications are not always effective, as they may assist individuals with partial vision, but provide no benefit to those with little or no vision[5]. While specific research on accessible grocery shopping support remains limited [4], there exists a body of literature on the broader topic of accessibility of mobile applications. Ghidini et al. [3] discuss the importance of spoken or vibration feedback, and a minimalist design when developing apps for people with visual impairments. Balconi et al. [1] compared the emotional responses of sighted people to people with visual impairments when identifying products in a grocery store by measuring pulse volume amplitude (PVA) They found that people with vision impairments had heightened PVA, which suggested greater emotional arousal and cognitive effort than the control group.Finally, Bhatlawande et al[2] developed audio guidance for recognizing grocery items. However, how much information to provide once an object is recognized, is still an open question, one that we address in our work.

We present VisionAI, an Android app that streamlines grocery shopping for users with visual impairments. After scanning product labels, it provides relevant information and the ability to get more on-demand, empowering informed decisions without the information overload or inadequacy common in existing apps.

2 Existing Apps

Existing AI-based apps share similar goals with our system but have notable limitations as shopping assistants. Microsoft's SeeingAI offers diverse detection features but falls short as a shopping assistant. Our co-designer found it provides irrelevant information, delayed feedback, and inconsistent performance. There have been reports of issues with the barcode scanning feature, the predominant mechanism to support shopping. Be My Eyes connects users with sighted volunteers, but this reduces independence and poses challenges with volunteer availability and expertise. While the paid app Aira partially addresses inconsistent assistance, Be My Eyes still struggles with reliability and quality of support. Google Lookout offers various detection features but struggles with in-store shopping applications due to limited product information, recognition issues, and camera positioning challenges. While existing apps are helpful, they often frustrate users like our co-designer, by providing inappropriate amounts or types of information. For instance, SeeingAI doesn't just identify the product; it reads out all the ingredients, nutritional information, and any other text on the back of a potato chip bag. This overwhelming amount of information, presented all at once, is often not helpful in making a quick purchasing decision.

VisionAI addresses this issue by offering need-to-know informa-

[1] www.afb.org/research-and-initiatives/statistics/adults-vision-loss-nhis
[2] https://www.bemyeyes.com
[3] https://aira.io/aira-about-us/
[4] https://www.seeingai.com

tion and allowing users to access specific details as required.

3 Co-Design Process

Our development of VisionAI was guided by Anon (name anonymized for review), a blind PhD student and proficient screen reader user. We began with an in-depth exploration of challenges faced when using existing shopping apps, which informed our initial concept. This was followed by an iterative design process, collaborating closely with Anon to refine our application. We focused specifically on optimizing the item scanning feature to aid purchasing decisions, ensuring that each iteration addressed real user needs. This user-centric approach allowed us to create an app that directly tackles the shortcomings of current solutions, resulting in a more effective shopping assistant for Anon and potentially for others with vision impairments.

We held weekly meetings with Anon to review the app during development. These sessions involved hands-on testing, where Anon provided detailed feedback on the app's current state, identified areas for improvement, and suggested new features. This collaborative process led to important refinements. For example, early feedback prompted us to prioritize speech-to-text functionality over text-to-speech, as Anon already used an advanced screen reader. Crucially, we reduced the amount of product information displayed to prevent cognitive overload. Instead of presenting all package details upfront, we implemented a conversational interface using ChatGPT. This allows users to request specific information as needed, contrasting with existing apps that display all details without prioritizing importance.

4 System Overview

VisionAI aims not only to provide basic information, but to act as a comprehensive shopping assistant, akin to shopping with a friend or family member. The app allows people to scan barcodes to retrieve product information, search for that product online, and engage in a conversation with ChatGPT to get more information as needed.

Barcode Scanning and Search Our app uses Android's CameraX API[5], which is integrated with a PreviewView to display the camera's capture on the screen. The setup includes an ImageAnalysis instance that processes the camera feed to detect barcodes. The AnalyzeImage method employs Google's ML Kit Barcode Scanner. Upon detection of a barcode, the application performs two actions: (1) it plays a sound to notify the user that the scan was successful, and (2) it launches the next product information screen. The auditory feedback was essential as both Anon and Ghidini et al.[3] have emphasized the importance of spoken or vibration feedback.A Google Search button streamlines the search process by automatically using the information from the barcode scan as the search query. This approach minimizes the need for user input, making the search more efficient with fewer errors.

ChatGPT On the product information page, a button directs users to ChatGPT, passing the product name and description as parameters. This information is incorporated into a prompt, ensuring that ChatGPT tailors its responses to the specific product. Additionally, a dictation feature next to the ChatGPT input bar converts speech to text as an alternative to typing. We used ChatGPT-3.5-turbo-0125, an updated version of 3.5-turbo. We tested various temperatures to determine the optimal setting for the model, and given the need for factual and precise information, a temperature of 0.5 was selected, as it produced concise and relevant answers.

5 Design Reflection

VisionAI was developed iteratively through weekly meetings with Anon. During each meeting, Anon tested the application in its current state and provided feedback on its functionality, suggested improvements, and proposed new features. These sessions yielded several critical insights shared below.

Text-To-Speech and Speech-To-Text Initially, we considered integrating text-to-speech capabilities into VisionAI. However, Anon brought up the fact that that most visually impaired users already have advanced screen readers integrated into their devices, making additional text-to-speech functionality redundant. Instead, we focused on incorporating speech-to-text features to enhance user experience. This decision was influenced by the need to simplify input methods; users can now choose to type or speak their queries directly to ChatGPT, with a dictation button conveniently located next to the text input bar.

Less Information is Better Our initial design presented extensive product information to users immediately. Feedback from Anon highlighted that this approach, while well-intentioned, was overwhelming. Existing applications often bombard users with information, making it difficult to remember and prioritize for decision-making. Therefore, we iteratively reduced the amount of information displayed, allowing users to request additional details as needed. This change aimed to create a more manageable and user-friendly experience.

Adding Semantics to the Application During one testing session, Anon observed that the screen reader failed to recognize the input bar for ChatGPT queries. This issue underscored the importance of adding semantic tags to the application. By embedding these tags, we ensured that screen readers could accurately identify and interact with all elements of the application. This improvement not only enhanced usability but also provided consistency across different screen readers and user experiences.

6 Co-designer Evaluation

Anon assessed the cognitive demands of using VisionAI, evaluating learning, focus, decision-making, recall, and communication on a scale of 1 (minimal) to 5 (extensive). VisionAI scored 1 in all categories, outperforming SeeingAI (focus: 3, recall: 2) and Be My Eyes (learning: 2, recall: 2). These preliminary results suggest VisionAI may offer cognitive load improvements over existing apps, though a larger study is needed to confirm the significance of these findings.

Developing accessible technology like VisionAI is crucial for creating more inclusive shopping experiences, empowering individuals with visual impairments to shop independently and make informed decisions.

References

[1] M. Balconi, C. Acconito, and L. Angioletti. Emotional effects in object recognition by the visually impaired people in grocery shopping. *Sensors*, 22(21), 2022. 1

[2] S. Bhatlawande, A. Kumari, M. Agrawal, S. Shilaskar, J. Madake, and A. Raj. Audio guided aid for assisting visually challenged people for shopping of grocery items. *AIP Conf. Proc.*, 2717(1), 2023. 1

[3] E. Ghidini, W. D. L. Almeida, I. H. Manssour, and M. S. Silveira. Developing apps for visually impaired people: Lessons learned from practice. In *2016 49th Hawaii Intnl Conf. on System Sciences (HICSS)*, pages 5691–5700, 2016. 1, 2

[4] J. Kasowski, B. A. Johnson, R. Neydavood, A. Akkaraju, and M. Beyeler. A systematic review of extended reality (XR) for understanding and augmenting vision loss. *Journal of Vision*, 23(5), 2023. 1

[5] Y. Zhao, S. Szpiro, J. Knighten, and S. Azenkot. Cuesee: exploring visual cues for people with low vision to facilitate a visual search task. In *Proc. of the 2016 ACM Intnl Joint Conf. on Pervasive and Ubiquitous Computing*, UbiComp '16, page 73–84, New York, NY, USA, 2016. ACM. 1

[5] https://developer.android.com/media/camera/camerax