

Echo: A Human–Computer Collaborative Design of an Intelligent Object-Finding System for the Visually Impaired

Muxi Liu¹ and Sihan Zhao²

¹Product Design, China Academy of Art, Zhejiang, Hangzhou 311100, China

²Product Design, China University of Geosciences (Beijing), Beijing 100083, China

ABSTRACT

Locating everyday objects independently remains a persistent challenge for visually impaired individuals. Existing assistive tools often provide limited functionality, such as audio cues or object labeling, but rarely support both object recognition and retrieval. The key design challenge is developing an object-finding system that aligns with non-visual cognition and interaction habits. This study proposes Echo, an intelligent assistive system integrating AI-based visual recognition, RFID sensing, and multimodal feedback to support object localization and identification in home environments. The system adopts a human–computer collaborative framework consisting of two modules: semantic registration and real-time object finding. The interaction workflow follows three stages: fuzzy retrieval, dynamic guidance, and terminal confirmation. AI visual recognition identifies distinct objects and provides voice-guided navigation, while RFID tags enable rapid identification of visually similar items such as documents or medicine bottles. The research was conducted in three phases: user interviews and home observations (N = 12), prototype development, and usability evaluation with 20 visually impaired participants performing standardized object-finding tasks in a simulated home environment. Results indicate that Echo improved task efficiency and success rates. The tactile-dominant wristband interface with vibrotactile feedback reduced learning costs and supported spatial awareness during navigation. The findings highlight the value of human factors–oriented design in intelligent assistive technologies and demonstrate how human–computer collaboration can enhance autonomy for visually impaired users.

Keywords: Accessible design, Assistive design for the visually impaired, Intelligent object-finding system

INTRODUCTION

For visually impaired individuals, locating everyday objects—especially small but frequently used items—is a common yet challenging task. Items such as keys, phones, medication, or identification cards can require prolonged searching when misplaced and may even pose safety risks. Unlike sighted users who rely on visual scanning, visually impaired users depend on tactile exploration, memory, and auditory cues, which often become inefficient in cluttered or complex environments.

In recent years, assistive technologies have made progress in converting visual information into non-visual cues. Large language models have enabled more natural image-description assistance on smartphones, with applications such as Be My AI providing real-time environmental descriptions for visually impaired users. However, Lee (2025) found that these systems still face limitations, including weak contextual understanding that can lead to hallucinations or misinterpretations, as well as difficulties in accurately understanding users' intentions. These issues may reduce user trust and limit their practical value as assistive tools. Without assistive technologies, visually impaired users rely mainly on tactile exploration, memory, and auditory cues, which often makes object searching inefficient in cluttered environments.

This study focuses on a central question: how to enable visually impaired individuals to locate everyday objects more independently and efficiently in familiar environments such as homes and offices, thereby reducing reliance on family members or external assistance. To achieve this goal, the system proposed in this study emphasizes providing stable and reliable navigation cues and safety information during the object-finding process, while returning key decision-making and action pacing to the user. This approach supports a smoother interaction loop of “labeling – recognizing – locating – confirming.”

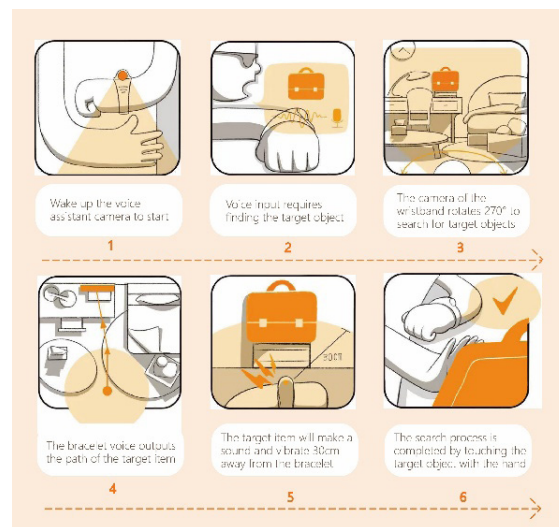


Figure 1: Workflow of the Echo intelligent object-finding system (marking–recognition–navigation–confirmation).

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The key challenge of assistive systems for visually impaired users is not only recognizing objects but also translating this information into executable actions in real environments. For example, SnapStick (2025) integrates a smart cane, bone-conduction headphones, and a vision–language model to support object recognition and scene description. While the system improves spatial awareness, existing solutions still face issues such as complexity, discomfort, and limited environmental perception.

In related assistive technologies, research is gradually shifting from simple “recognition or prompting” toward route guidance and hazard notification. Approaches based on RFID and other tagging technologies assign identities to objects, enabling systems to locate them via radio signals and guide users to target locations through wearable auditory or tactile feedback. Such approaches have demonstrated feasibility for indoor localization and navigation. For example, An RFID-Based Indoor Guiding System for Visually Impaired People (Radu et al., 2025) proposes an indoor guidance system based on passive ultra-high-frequency RFID technology to help visually impaired individuals reach predefined locations within public buildings. The system employs an unmanned ground vehicle for guidance and integrates RFID reading, voice commands, and audio instructions to support route recording and playback.

Vision-based recognition systems help identify surrounding objects but often lack reliable navigation guidance. Studies suggest that spatial information should be translated into actionable instructions for users. For example, Giudice et al. (2024) propose a landmark-based navigation approach that guides users step by step and reduces uncertainty.

NavGraph (Mascetti et al., 2025) proposes the Parsimonious Instructions principle, which balances timely guidance with avoiding information overload. Similarly, Tang et al. (2025) developed a multimodal wearable system that converts visual information into real-time action commands through bone-conduction audio and tactile feedback.

Similarly, the UNav system adopts step-by-step wayfinding instructions to guide visually impaired users during navigation (Yang et al., 2026).

Taken together, these studies suggest that organizing spatial and localization information into sequential, actionable instructions that users can follow is a key design direction for improving the usability of navigation systems for visually impaired individuals.

METHODOLOGY

This study adopted a mixed-methods research approach, combining qualitative interviews and experimental testing to systematically evaluate the usability and user experience of the Echo intelligent object-finding system. Following a user-centered design philosophy, the research was conducted in three stages. First, in-depth interviews and home observations were carried out to identify the pain points and behavioral patterns of visually impaired individuals when searching for objects in real domestic environments. Second, based on the needs analysis, a functional prototype of the Echo system was developed, integrating AI-based visual recognition, RFID localization, and non-visual interaction. The prototype consists of a smart wristband and a voice-controlled label printer. Finally, user testing was conducted in a standardized simulated home environment using a single-factor between-subjects experimental design. The independent variable was the type of assistive device (Echo system vs. a traditional beeping object finder with Braille labels), while the dependent variables included task completion time, success rate, cognitive load, and subjective usability ratings. The testing environment

included a living room, bedroom, and kitchen, with 20 everyday objects pre-set as target items. All participants completed a device familiarization training session prior to the experiment to control for familiarity effects.



Figure 2: Voice-controlled label printer used to generate Braille and RFID tags for object registration.

A total of 20 visually impaired participants were recruited for the study (aged 23–67, $M = 42.5$, $SD = 11.3$), including 8 blind participants, 7 with severe low vision, and 5 with moderate low vision; 12 were male and 8 were female. All participants had independent living experience and reported frequent difficulties in locating everyday objects. Participants were randomly assigned to either the experimental group or the control group (10 participants in each group). No significant differences were observed between the two groups in terms of age, gender, or level of visual impairment.



Figure 3: Functions of the Echo smart wristband, including AI recognition, RFID scanning, voice interaction, and vibrotactile feedback.

The Echo prototype system enables a complete workflow of “labeling – localization – navigation – confirmation” through the collaboration of the smart wristband and the voice label printer.

Table1: Echo smart bracelet hardware modules and technical specifications.

Component/Module	Primary Function	Technical Specifications & Design Features
AI Vision Recognition Camera	Actively scans the environment to identify object contours, salient colors, and surface materials	270° rotating field of view to reduce blind spots; equipped with a lightweight deep learning model for real-time object classification
RFID Positioning Module	Reads adhesive RFID tags for non-line-of-sight (NLOS) item localization	UHF RFID reader chip with an effective reading radius of ≥ 10 meters; capable of penetrating non-metallic doors and drawers; supports anti-collision algorithm for multiple tagsobstacles like cabinet
Interaction Buttons	Facilitates user input and provides operational feedback	Raised Braille physical buttons with layout conforming to tactile logic; integrates voice feedback and a differentiated vibration motor for dual-mode confirmation
Orientation Perception Design	Helps users quickly perceive the wearing direction and operation position of the bracelet	Asymmetrical beveled shell design allows users to distinguish top from bottom through touch; made with skin-friendly materials for comfortable wearing
Safety Monitoring Module	Monitors ambient temperature in real-time to prevent burns and accidents	Built-in thermal sensor triggers voice and vibration alerts when the ambient temperature exceeds a predefined safety threshold

Data collection included task completion time, success rate across 10 object-finding tasks, NASA-TLX cognitive workload ratings, and wristband interaction logs (e.g., voice commands, button operations, RFID scans, and AI recognition). Experimental group participants also completed semi-structured interviews on usability and perceived autonomy. All data were anonymized and analyzed.



Figure 4: Layout of the simulated home environment used for the object-finding experiment. Table 2. NASA-TLX Cognitive Load Scale.

Table 2: NASA-TLX cognitive load scale.

Dimension	Description	Rating (0-100)
Mental Demand	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, searching, etc.)?	___ / 100
Physical Demand	How much physical activity was required (e.g., moving, bending, reaching, etc.)?	___ / 100
Temporal Demand	How much time pressure did you feel due to the pace at which the task or task elements occurred?	___ / 100
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?	___ / 100
Performance	How successful do you think you were in accomplishing the goals of the task?	___ / 100
Frustration Level	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?	___ / 100

Table 3: Pairwise comparisons.

No.	Dimension Pair	Your Choice (mark \surd)
1	Mental Demand vs. Physical Demand	<input type="checkbox"/> Mental <input type="checkbox"/> Physical
2	Mental Demand vs. Temporal Demand	<input type="checkbox"/> Mental <input type="checkbox"/> Temporal
3	Mental Demand vs. Effort	<input type="checkbox"/> Mental <input type="checkbox"/> Effort
4	Mental Demand vs. Performance	<input type="checkbox"/> Mental <input type="checkbox"/> Performance
5	Mental Demand vs. Frustration	<input type="checkbox"/> Mental <input type="checkbox"/> Frustration
6	Physical Demand vs. Temporal Demand	<input type="checkbox"/> Physical <input type="checkbox"/> Temporal
7	Physical Demand vs. Effort	<input type="checkbox"/> Physical <input type="checkbox"/> Effort

(Continued)

Table 3: Continued.

No.	Dimension Pair	Your Choice (mark \surd)
8	Physical Demand vs. Performance	<input type="checkbox"/> Physical <input type="checkbox"/> Performance
9	Physical Demand vs. Frustration	<input type="checkbox"/> Physical <input type="checkbox"/> Frustration
10	Temporal Demand vs. Effort	<input type="checkbox"/> Temporal <input type="checkbox"/> Effort
11	Temporal Demand vs. Performance	<input type="checkbox"/> Temporal <input type="checkbox"/> Performance
12	Temporal Demand vs. Frustration	<input type="checkbox"/> Temporal <input type="checkbox"/> Frustration
13	Effort vs. Performance	<input type="checkbox"/> Effort <input type="checkbox"/> Performance
14	Effort vs. Frustration	<input type="checkbox"/> Effort <input type="checkbox"/> Frustration
15	Performance vs. Frustration	<input type="checkbox"/> Performance <input type="checkbox"/> Frustration

RESULTS

The experimental results indicate that the Echo system has a positive effect on improving the efficiency and accuracy of object-finding tasks. The experimental group achieved an average task completion time of 74.5 seconds ($SD = 16.8$), compared with 98.6 seconds ($SD = 24.3$) for the control group, $t(18) = 3.86$, $p < .01$, $d = 1.15$. In terms of task success rate, the experimental group reached 85% (85 successful tasks out of 100), while the control group achieved 78% (78 out of 100), $\chi^2(1) = 4.12$, $p < .05$.

For system usability evaluation, the experimental group obtained an average System Usability Scale (SUS) score of 79.8 ($SD = 8.2$), corresponding to a B+ grade, which indicates usability superior to more than 75% of systems. Among the SUS items, “the system is easy to use” (mean = 4.2/5) and “I learned to use the system quickly” (mean = 4.3/5) received relatively high scores, suggesting a low learning cost for users.

Regarding cognitive load, the experimental group reported a NASA-TLX overall workload score of 56.3 ($SD = 11.2$) during the first task, slightly higher than the control group’s 52.8 ($SD = 9.5$), $t(18) = 1.82$, $p = .08$ (not significant). However, with increased use, the workload in the experimental group gradually decreased, reaching 48.7 ($SD = 9.8$) by the fifth task, which was comparable to the control group’s 49.5 ($SD = 10.2$). In terms of specific dimensions, the experimental group initially reported slightly higher scores in effort (61.2 vs. 55.8) and frustration (52.3 vs. 48.6). However, participants rated their performance higher in the experimental condition (65.7) compared with the control group (57.9), indicating that users were able to perceive the performance benefits provided by the system.

In post-experiment interviews, eight participants affirmed the effectiveness of the multimodal feedback design. The combination of raised buttons and vibration made interaction more intuitive. One participant with severe low vision commented that “I can immediately tell which button it is just by touching the wristband, and the vibration feedback is also very clear.” Several participants also noted that the dual-channel feedback of touch and sound enhanced their spatial awareness. One blind participant described the experience as “like suddenly having a sound coordinate in the dark.”

However, six participants reported confusion about the boundaries of voice commands during their initial use of the system, such as uncertainty about how specifically they should describe the target object. This observation

is consistent with the relatively higher effort scores in the NASA-TLX results, suggesting that non-visual interaction design should more clearly communicate the system’s recognition scope and input expectations.

DISCUSSION

Overall, the results suggest that Echo improves the efficiency of object-finding tasks for visually impaired users and increases their ability to complete such tasks independently. Compared with the control group, the experimental group showed improvements in task completion time, first-attempt success rate, and reduced misidentification. Post-experiment interviews also indicated that participants felt more in control, reassured, and willing to complete tasks independently (Lee, 2025; Karamolegkou et al., 2025).

Based on the current framework, the Echo system has the potential to expand to broader user groups and collaborative scenarios (Gerald, 2025). In addition to visually impaired users, the system could also support everyday object management for sighted users. In this case, tactile route guidance could be replaced by map-based navigation or AR directional cues on mobile devices, while features such as final-stage confirmation and hazard alerts are retained. This would allow the system to evolve from an accessibility aid into a more universal tool for household object management and health-related reminders.

Second, high-risk search tasks could be outsourced to camera-equipped drones. After users specify the target object, the drone would conduct the search and retrieval and return the object once confirmed by the system. A family or remote caregiver interface could also be introduced to support users while preserving autonomy. The interface would provide summary notifications of key status information, such as medication adherence, and alert family members when risk events are detected.

CONCLUSION

In terms of theoretical contributions, this study extends a human–computer collaboration model for everyday object-finding scenarios that integrates “executable route guidance” with “verifiable confirmation.” This approach differs from many existing assistive technologies that either recognize objects but struggle to guide users effectively or provide auditory cues with unstable route guidance, representing only single-function capabilities. The study also validates the effectiveness of the “tactile-dominant with auditory support” design principle in non-visual interaction, and provides empirical evidence for evaluating user autonomy, a key human-factor dimension in intelligent assistive technologies.

In terms of practical contributions, the research proposes a reproducible “tagging–localization–navigation–confirmation” system framework for object finding (Gerald, 2025), including the specific design of a smart wristband and a voice-enabled label printer. While improving object-finding efficiency, the system also enhances users’ sense of safety in independent living through thermal hazard warnings and environmental monitoring mechanisms. These design solutions provide a practical technological pathway and direct reference for the development of accessible product design (“SnapStick,” 2025; Tang et al., 2025).

ACKNOWLEDGMENT

The authors would like to thank all visually impaired participants who generously shared their experiences and took part in the interviews and experimental evaluations. Their insights provided valuable guidance for the design and evaluation of the Echo system. We also appreciate the support and feedback from colleagues during the development and testing of the prototype.

REFERENCES

- Aloui, N. (2025). Towards spatial awareness: Real-time sensory augmentation with smart glasses for visually impaired individuals. *Electronics*, 14(17), 3365. <https://doi.org/10.3390/electronics14173365>
- Gerald, R. (2025). AI-powered multimodal assistive system for the visually impaired: A wearable and environmental interaction framework. *TechRxiv*. <https://doi.org/10.36227/techrxiv.174858149.98832198/v1>
- Giudice, N. A., Guenther, R., Kaplan, L., & Anderson, J. (2024). All the way there and back: Inertial-based, phone-in-pocket indoor wayfinding and backtracking apps for blind travelers. *ACM Transactions on Accessible Computing*, 17(4), 1–35. <https://doi.org/10.1145/3696005>
- Karamolegkou, A., Nikandrou, M., Pantazopoulos, G., Sanchez Villegas, D., Rust, P., Dhar, R., Hershovich, D., & Søgaard, A. (2025). Evaluating multimodal language models as visual assistants for visually impaired users. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics* (pp. 25949–25982). Association for Computational Linguistics. <https://aclanthology.org/2025.acl-long.1260/>
- Lee, S. (2025). Beyond visual perception: Insights from smartphone interaction of visually impaired users with large multimodal models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)* (pp. 1–20). Association for Computing Machinery. <https://doi.org/10.1145/3706598.3714210>
- Mascetti, S., Ahmetovic, D., Galimberti, G., & Coughlan, J. M. (2025). NavGraph: Enhancing blind travelers' navigation experience and safety. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 9(3), 117. <https://doi.org/10.1145/3749537>
- Radu, V., Crişan-Vida, M., & Stoicu-Tivadar, L. (2025). An RFID-based indoor guiding system for visually impaired people. *Information*, 16(3), 220. <https://doi.org/10.3390/info16030220>
- SnapStick: Merging AI and accessibility to enhance navigation for blind users. (2025). *Technologies*, 13(7), 297. <https://doi.org/10.3390/technologies13070297>
- Tang, J., Zhu, Y., Jiang, G., Xiao, L., Ren, W., Zhou, Y., Bi, H., Gu, Q., Yan, B., Zhang, J., Wu, X., Fan, Z., Gu, L., et al. (2025). Human-centred design and fabrication of a wearable multimodal visual assistance system. *Nature Machine Intelligence*. <https://doi.org/10.1038/s42256-025-01018-6>