

# Tech designed to aid visually impaired could benefit from human-AI collaboration

22 April 2022, by Jessica Hallman



Credit: Pixabay/CC0 Public Domain

Remote sighted assistance (RSA) technology—which connects visually impaired individuals with human agents through a live video call on their smartphones—helps people with low or no vision navigate tasks that require sight. But what happens when existing computer vision technology doesn't fully support an agent in fulfilling certain requests, such as reading instructions on a medicine bottle or recognizing flight information on an airport's digital screen?

According to researchers at the Penn State College of Information Sciences and Technology, there are some challenges that cannot be solved with existing [computer vision techniques](#). Instead, the researchers posit that they would be better addressed by humans and AI working together to improve the [technology](#) and enhance the experience for both visually impaired [users](#) and the agents who support them.

In a recent study presented at the 27th International Conference on Intelligent User Interfaces (IUI) in March, the researchers highlighted five emerging problems with RSA that

they say warrant new development in human-AI collaboration. Addressing these problems could advance computer [vision](#) research and initiate the next generation of RSA service, according to John M. Carroll, distinguished professor of information sciences and technology.

"We're interested in developing this particular paradigm because it is a collaborative activity involving sighted and non-sighted people, as well as computer vision capabilities," said Carroll. "We framed it in a very rich way where there are a lot of interesting issues of human-human interaction, human-technology interaction and [technology innovation](#)."

Remote sighted assistance technology is currently available through free applications that connect visually impaired users with sighted volunteers or as a paid service connecting them to sighted agents. The technology is deployed when a visually impaired person needs help with a daily task that requires sight—such as finding an empty table in a restaurant, reading a food package label or identifying what color an object is—and calls an agent using a live video function on their mobile device. The agent then sees the user's world through that lens, serving as their eyes to help them navigate their request.

But according to Syed Billah, assistant professor of IST and co-author on the paper, the support that agents provide is not easy.

"For example, creating a worldview by looking through the camera is mentally demanding for the agents," said Billah. "The good news is that part of this task can be offloaded to computers running a 3D reconstruction algorithm."

However, some of the support that agents provide—such as helping a visually impaired user navigate a [parking lot](#) or read a label on a bottle of medication—comes with higher stakes.

"To address these problems, there is room for improvement with the current computer vision technology," said Billah.

In their study, the researchers reviewed existing RSA technologies and interviewed users to understand technical and navigational challenges they face when using the service. They then identified a subset of challenges that could be addressed with existing computer vision technologies, and proposed design ideas for addressing them. They also identified five emerging problems that, due to their complexity, cannot be addressed by existing computer vision techniques.

The researchers believe these problems could lead to new opportunities to enhance the RSA design and experience by:

- Recognizing that objects commonly identified as obstacles by smartphone cameras may not be considered obstacles by visually impaired individuals, but instead are useful tools. For example, a wall bordering a sidewalk may be displayed as an obstacle in common navigational apps, but a visually impaired person walking with a cane may rely on it to navigate their steps.
- Helping users navigate their environment when a live camera feed may be lost during low cellular bandwidth, which frequently occurs in indoor settings.
- Recognizing content on digital LCD displays, such as [flight information](#) in an airport or temperature control panels in a hotel room.
- Recognizing texts on irregular surfaces. Often, important information is printed in ways that make it difficult for human agents assisting visually impaired individuals to read; for example, medication instructions on a curved pill bottle or a list of ingredients on a bag of chips.
- Predicting how out-of-frame people or objects will move. Agents must be able to quickly communicate environmental information in a user's public surroundings, for example other pedestrians or a moving car, to help the user avoid collision and keep the user safe. However, the

researchers found that it is currently difficult for agents to track these other people and objects, and nearly impossible to predict their trajectories.

The researchers hope that their study will improve the experience for both visually impaired users and agents.

"In the future we imagine that we can use computer vision to give the agent a very immersive experience and provide them with the mixed reality technology," said Rui Yu, doctoral student of IST. "And we will be able to directly help the users get some basic information about their environment based on [computer](#) vision technology."

Sooyeon Lee, former doctoral student at the College of IST and current postdoctoral researcher at Rochester Institute of Technology, and Jingyi Xie, doctoral student of informatics, also collaborated on the study, which was supported by the U.S. National Institutes of Health and the National Library of Medicine.

Provided by Pennsylvania State University

APA citation: Tech designed to aid visually impaired could benefit from human-AI collaboration (2022, April 22) retrieved 15 August 2022 from <https://techxplore.com/news/2022-04-tech-aid-visually-impaired-benefit.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*