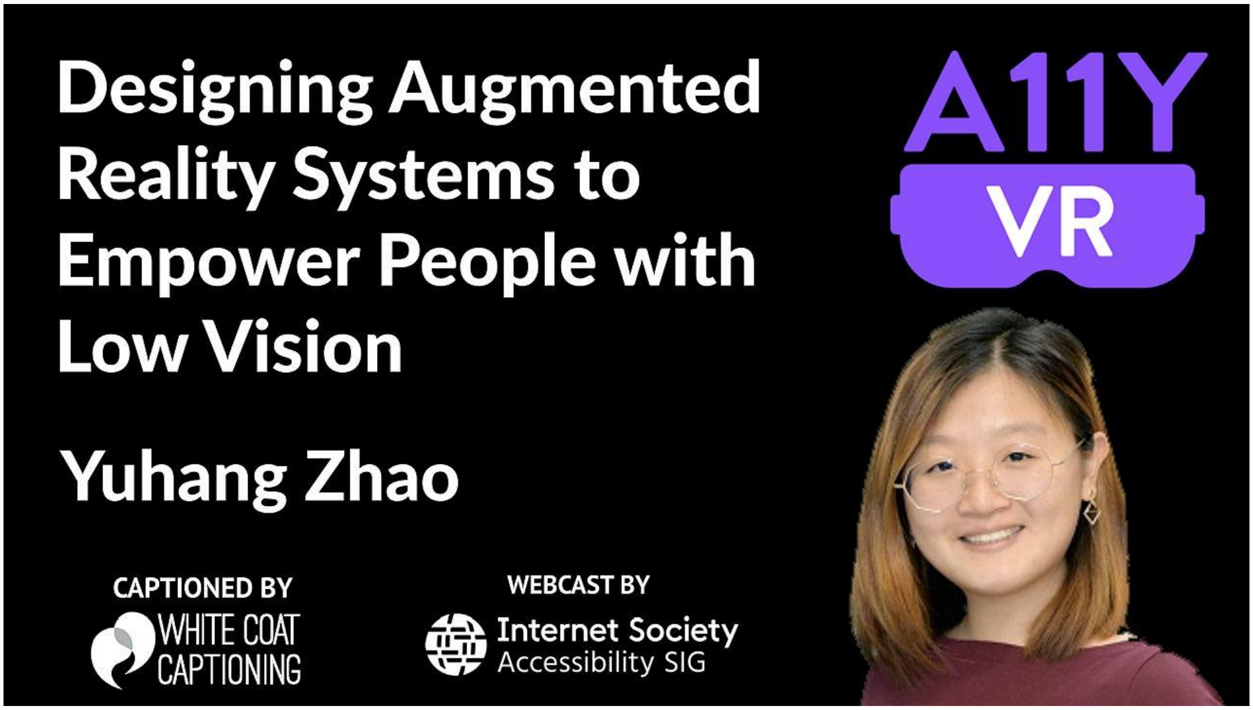

Designing Augmented Reality Systems to Empower People with Low Vision - Yuhang Zhao

Accessibility VR Meetup – Oct 22 2020

YUHANG ZHAO: Thank you, Thomas. Hello, everyone. Glad to be here today. This is actually my first talk in virtual reality. Very exciting experience.

Today I want to talk about my research on how I designed augmented reality systems to better serve and empower people with low vision who have impairments, but also have usable vision. First of all I want to talk a little bit background about what is low vision. By definition, low vision includes all kinds of vision impairments that fall short of blindness but can't be corrected by contacts and lenses, and it's very complex.

For example, this is the typical vision that a sighted person sees, or through their eyeglasses, but, if you have low vision, for example, if you have a cataract, then your vision can be very blurry. This blurriness can't be corrected by eyeglasses and there are many low-vision conditions. People can have central vision loss or lose their peripheral vision. These are just some examples and there are many different conditions. Also, a low-vision person can have multiple conditions at the same time. All of these bring a lot of difficulties to people's daily life.

The more interesting thing is, if you look at these images, you may also notice that low vision people, they still have a lot of functional vision to use. The fact is, most of them, they still prefer and actively use their vision in many important daily activities. This actually has made low-vision research more challenging, because this will not only involve technology in the vision, but also a deep understanding about different low-vision people's visual perceptual abilities.

In accessibility research there has been a lot of work that has been done for people who are completely blind, or people with vision impairments in general. Most of these work, they are focused on providing audio or haptic feedback to compensate for the impaired vision for people with visual impairments, but have overlooked the fact that people with low vision still have functional vision to use. So there is, actually, a significant lack of knowledge in the research field on how low-vision people use their vision in their daily life, and how can we create suitable technologies to enhance their visual abilities in their daily activities.

But the good news is, augmented reality technology has presented a very unique opportunity to leverage the functional vision that low-vision people have. So, for example, with the camera automatically recognizing the surrounding environment, low-vision people now can perceive very well. AR technology can generate very immersively, thus enabling low-vision people to better perceive their surrounding environment completely differently for visual tasks.

So, my research goal is to leverage, to take this unique opportunity that AR technology has presented, and explore how we can leverage AR technology to enhance low-vision people's visual ability. There are two goals in my research. One is to understand and explore low-vision people's visual perceptions when they're using different types of AR devices. Also, in different types of daily tasks, and based on their needs and their experiences, to design and build different AR systems to enhance their visual ability directly, by designing different types of visual augmentations to facilitate different daily tasks for them.

Today, I'm going to give you two examples about how I do this in my research. I'll start with visual search. I started to explore this visual search task with an explorer study where I observed how low-vision people conduct a grocery shopping task. This is a very typical start of my research, which is to do an exploratory study, because it gives me a really good context to understand the user's experience. How I do this is I ask the participants to conduct this grocery shopping task. I give them a targeted product, and I ask them to find the nearby grocery store, to get into the store, and to find the specific product and buy it. During this process, I followed the participants, observed their behaviors, and asked follow-up questions about what they can see, why they behave like this, and what technologies they are using.

During this process, one most important finding I got is, while low vision people encountered a lot of difficulties, one of the most challenging and time-consuming tasks is finding the specific target from this crowded grocery store shelf, which is a visual search task. How they usually do it is, they get very close to the shelf, pick up each item one-by-one, and bring them very close to their eyes. Some of them also use a magnifier to look at the details, and then put it back and pull up another item. They do this repeatedly, repeat this process until they find the correct product, which is very low efficiency process. During my study, I didn't find any effective tools low-vision people use to facilitate this visual search task.

So, there is a really big gap there. Visual search is very important, and also a fundamental visual task in people's daily life. Researchers have already designed different systems, come up with different solutions to address this problem for the general visually impaired population.

For example, there have been systems that leverage computer-vision technology, and also crowdsourcing, to recognize different products for the user, and then generate different audio feedback to guide the user to the target. Also, there have been systems that enable a blind person to scan the barcode on each of the products, and then they will hear the audio description of the different products. But still, the user has to scan each item one-by-one until they find the target.

So, for all this prior research in accessibility, they mostly focused on people who were completely blind, and only provided audio feedback during this visual-search process. But, as I mentioned before, low-vision people are different, because they still have existing vision, and the visual channel is actually the most efficient sensory channel they use to perceive the surrounding environment, especially in this visual-search task. That's why it's different from the prior research. My goal is more, how to leverage people's functional vision and design, to help them find a target more quickly and easily, in this visual-search task.

So, based on that, I designed CueSee, which is a new approach for people with low vision to help them conduct a visual-search task via direct visual augmentation. If the user wears a head-mounted display that shows the surrounding environment, and the system also has a camera scanning the environment, then we can use computer vision to automatically recognize the different targets for the user, and then, the system can render different visual cues to attract and turn the user's attention, so that the low-vision person can easily follow these cues to find a target.

Now I want to show a video demonstrating how the system works. First the user can use speech commands to speak out the name of the target and assign it to the system. I apologize that I cannot add audio to this virtual space, so I'm just going to describe whatever the human is saying. She says, look for the pea snack, one of the products on the shelf. Then she'll look around the shelf until the system recognizes the target, and then the system cues, and she can use the cues to find the target. This is how the system works. But, to accomplish such a system, the most important question we want to address is, what visual cues should we provide there? What visual cues would be effective enough for people with low vision? This actually can be a very challenging problem. Because, on the one hand, we want the visual cues to be visible enough, but a lot of visual information, that can be easily seen by people who are sighted, are not that easy to be seen by people with low vision.

For example, this image shows the blue cursor, which is a very typical visual cue on holo-lens, it's used as a cursor. But this is actually very difficult for people with low vision to see, and in this visual-search scenario, we not only want the cue to be visible, but also want it to be obvious enough to be able to attract people's attention immediately. On the other hand we also want to make sure that the cues are not too distracting. That prevents the user from using their own vision to observe the surrounding environment. For example, if a user has peripheral vision loss, it's like they see through a tunnel. In that case, if a cue is relatively bigger, it can easily block the user's full field of view, and stop them from using their own vision. So, it is very important to think about how to balance the visibility, and also the distraction, when we design the visual cues. Of course, there are a lot of different low-vision conditions, as I mentioned at the beginning of the talk, which means we also need to consider other design alternatives to fulfill people's different needs.

So, with all of these requirements, what I did is, I first worked with an optometrist to deeply understand different low-vision conditions and their needs. I also looked into some cognitive psychology theories to understand what visual stimuli can effectively attract people's attention. By combining these different perspectives, I designed five different visual cues to orient and attract low-vision people's attention. For the sake of time I'm going to only give one example.

This one is called Flash. It's designed based on the cognitive psychology that onset, which means the appearance of something, can visually attract the people's attention. To amplify this for people with low vision, I put a frame around the target, and switch the color from gray to black, to create a continual experience to attract low vision people's attention. Also, a low-vision user can select and combine different visual cues to enhance their experience in the CueSee system.

Before we give a lot of time into the deep implementation, we have to figure out if these visual cues are effective enough for low-vision people. To do that I combined a webcam with an Oculus DK2. This is a very old version of Oculus, now they have a much better version. I created a mockup of 24 products on a grocery store shelf. You may have noticed that there are QR codes on all of these products. This is because for this specific study we only wanted to focus on the evaluation of the cues. We wanted to minimize the potential of recognition errors, to ease the computer vision process. `That's why we just recognize these QR codes in this system. We recruited 12 low-vision participants and asked them to conduct multiple versions of this task of finding the grocery on the store shelf. The participants evaluated the task with two conditions, one is CueSee with their preferred cues, and the other one is best correction. That means the user can use their eyeglasses, or magnifier, or whatever assistive tools they already have, or prefer to.

Now I want to show two videos to compare low-vision participants' experiences -- the ones they used, and CueSee. First, when they're not using CueSee, they're using their original vision and their preferred assistive tools, so you can see that they get very close to a shelf and look at each item one-by-one. This is actually a very long process. It's a totally linear-scanning process. One, they are trying to look for the target. And here he found the target. But, as a comparison, with CueSee now you can see he's wearing the head-mounted display and it actually took him a much shorter time to find the target. So, the comparisons between these two videos demonstrate the effectiveness of the cues. Now we all know that this study was conducted in a very ideal situation where we only recognized the QR codes.

So now, we have validated the effectiveness of the design, and then what if we built a high-fidelity system that actually used computer-vision technology? How would that system affect the user's experience? We do want to consider how much the computer vision technology, including its accuracy, and the speed, will be a factor in the user's experience. To answer that question, I built out the system of CueSee. First, I added a product detection component. I used a computer vision module which is a faster R-CNN model, which is a state-of-the-art computer vision object detection model, and I trained this model based on an RPC dataset which includes 200 different products. After the training, this algorithm actually got a relatively high accuracy, with a mean average precision of 0.85, which is a pretty good result in terms of computer vision. Also, clearly, a deep-learning module would not be fast enough to support a realtime interaction system. So, I added an object-tracking component to speed the system up. We also have a speech recognition system to enable a user to be able to interact with the system with speech, for example, assign a target to the system. Based on this implementation, we conducted a very similar study, but this time all the products do not have any QR codes on them. So, the system purely relies on the computer vision component of the system.

With the similar setup, we recruited another seven low-vision participants, and asked them to conduct a similar visual search task with CueSee, and also their best correction. During the study, we found that CueSee has significantly reduced the participants' search time. If you look at these bar charts, it shows the mean time across all the participants when they do the visual search task under the two conditions. The blue bar represents the time when they used the best correction, and the yellow represents CueSee. You can see that the time was reduced to one-third of the user's original time, which means that it's a really big improvement, and also in terms of the search accuracy. Unfortunately, the computer vision technology does make errors, so the final accuracy is not 100%. But, still, compared with the users' baseline correction, the system still improved the users' search accuracy from 83% to 95%.

CueSee is the first augmented reality system that automatically recognizes surrounding environment, and generates tailored visual feedback to enhance low-vision people's visual ability, and isdifferent from the conventional low-vision aids, for example, magnification, that are temporarily applied by the user to improve the view. The most important point that I want to emphasize here is that it's very important, when we design visual augmentations, to consider the user's context, for example, their different visual abilities, and also the specific tasks, as opposed to magnifying their field of view directly.

That's an example of a visual search task. But, of course, visual search is not the only difficulty that low-vision people have encountered in their daily life. During my research I've always been looking for the challenges that low-vision people have in their daily life. For example, before we actually buy stuff in a grocery store, people need to get to the store first, which will involve a lot of mobility and navigation tasks. Steering navigation is one of the most dangerous tasks that low-vision people face. Prior research has shown that low-vision people have a much higher rate of injuries and falls when walking down the stairs, as opposed to people who are sighted.

What makes the stairs so difficult for people with low vision, and what technologies can be designed to facilitate this task for people with low vision? To better understand low-vision people's experience in this task, I started with an exploratory study, observing 14 different low-vision participants when they walked on different types of stairs, for example, indoor stairs and outdoor stairs, and stairs made out of different materials, like wooden stairs, and so on. I found that, surprisingly, the white canes widely used by people who are blind are not that often used by people with low vision. For my study, I found only 4 out of 14 low vision participants used a white cane. Most of them used their functional vision when they walked on the stairs. One of the most helpful information for them are the contrast strips which are these high-contrast colors at the edge of the stairs. A lot of participants tried to look at these contact strips to identify the existence, and also the position, of each step on the stairs.

However, the problem is, not all the stairs are accessible, and not all contrast strips are accessible. Many contrast strips don't have enough contrast, and many stairs don't have any contrast strips at all, which made stair navigation very difficult for people with low vision. Based on that, how can we design technologies for people with low vision, considering that the contrast strips are the most helpful information that all low-vision people are looking for when they walk on stairs? What if we can design some mobile and personalized version of contrast strips, so that we can make the originally inaccessible stairs more accessible? Inspired by that idea, I designed StairLight which is a system based on projector technology.

The reason I chose a projector-based AR device, as opposed to AR glasses, is because the current AR glasses always have very limited field of view, which is very difficult for people to use when they are conducting navigation tasks. In a navigation task, people rely on their lower peripheral vision a lot, but it is not covered by any AR glasses. That's why I decided to design a projector-based system. The low-vision user holds this projector-based device just like holding a flashlight, it has a camera that automatically recognizes the objects, and they can increase the projection of the highlights to increase the stair visibility. The problem is what visual highlights should we generate here. Should we simply replicate these yellow contrast strips, and generate very similar yellow highlights on each of the stairs? Does each stair mean the same to a low-vision user?

To answer that question, in my exploratory study, I asked my low-vision participants how they perceived each stair, and how they perceived different contrast strips. Based on their answers, I found that not all the stairs are the same. Among all the steps, the first and last stairs are the most important for people with low vision to locate. After they can identify the first step, based on their muscle memory, they can probably just walk through the whole stairs until they encounter the end of the stairs. Also, not all contrast strips are helpful for people with low vision. Many of them mentioned that, if they see the same contrast strips on each of the stairs, it actually messed up with their depth perception, and distracted them a lot. Most of them only wanted contrast strips top and the bottom of the stairs.

Based on the users' experiences, I started to design the visual highlights for StairLight. While the design space is quite large, my design purely depends on the users' needs that I distilled from my exploratory study. Specifically, for this study, I had two findings. First, two design rules. First, I need to emphasize the first and the last stairs, and also we should probably minimize the distraction of the middle stairs. Based on these rules, I used a different thickness of the highlights to distinguish the first and the last stairs from the rest of the stairs. So you can see, on the left of the image is a thick yellow highlight on the first step. In contrast, for the middle stairs, I used the relatively thin stairs, so the thick highlights can better emphasize the first step and also the last step.

To enable a low-vision person to notice the existence, and also the exact position, of the first step, I designed different animations for the first stair, so it can better attract the low-vision person's attention. This slide shows two examples. One is a flash, and the other is a movement animation. Also I wanted to minimize a distraction on the middle stairs. To do that, I provided different color alternatives for the highlights for the middle stairs, including a yellow color, and also a blue color. Of course, the user can choose not to have any highlights on the middle stairs if they prefer to. Based on the design of these visual highlights, the user can select and combine the different visual highlights as they walk on different stairs while using StairLight.

This is a video showing how the user used StairLight. If they point the system at the stairs, the system can recognize the stairs. It's just like the interaction of a flashlight, and the system can recognize the most close three to four steps, points the visual highlights to the stairs, and leads the user to the front. This is a demonstration of how the system works.

Still, the first question we want to answer is: how does the design work? Whether the visual highlights are effective enough for people with low vision? The problem is, if we think about how we typically evaluate a system, we usually think about the accuracy, or the speed, to measure performance. The stair-navigation task is a little different because it's a very dangerous task, and we probably don't really care if the user can run on the stairs. The most important thing is we want to guarantee the users' safety. Safety is actually a very tricky factor to measure, because it's probably impossible. Like, we measure how many times a low-vision person will fall off the stairs when they are using our system> It's very impossible. So, how should we do that?

To reasonably measure people's safety, as a first step, I, instead, measured the user's psychological security. This is actually a term from psychology which means how a person measures her environment, perceives her environment, to be safe and out of threat. Of course, this psychological security is not the same as safety, right? If we think about the actual system, the system can have recognition errors. If the user is using our system, and they trusted the system a lot, and think they are safe, but the system recognized the surrounding information wrong, and projected the visual highlights at the wrong position, it actually puts the user in a more dangerous position. How can we address this problem?

To make the psychological security more close to people's actually safety, I designed an evaluation study, to generate a relatively more ideal situation, to make people's psychological closer to their actually safety. How I designed the security, I put two stationary projects at the top and bottom of the stairs, and I used a method called Wizard of Oz, which is very typical method of HCI. By observing the users' behavior, the researcher can adjust the system manually, based on the user's position on the stairs in our case. I also asked the participant to hold a smartphone in their hands, to simulate the experience of having a handheld device in their hand. That's what you see in the demo video. The user is holding a smartphone while walking on the stairs.

Based on this setup, I recruited 12 low-vision participants, and asked them to navigate on the stairs in two conditions including using StairLight, and also, as a baseline, I also asked them to use their typical walking method, for example, if they wanted to use their white cane, or if they wanted to hold on to the railings. After the study I asked the participants to report, give a 7-point Likert scale score, and this bar chart shows the participants' mean scores for the positions. The blue represents their typical walking method, and the yellow method represents when they used StairLight. The result showed that Stairlight can significantly improve the participants' psychological security. All the participants felt much more confident when they were working on the stairs with our system.

So, that's the evaluation of the design and, after that, of course, we will need to implement the system. This is still a work in progress. I combined a Kinect V2 to a projector, and also an IMU to track the users' hand motion so that,so no matter how the user's hand moves during the walking process, we can still project the corrected result based on the user's behavior. I also worked with an industrial designer to build this case, via 3D printing and laser cutting, so that a user can hold all of these devices at the same time. Based on this hardware setup, I first calibrated the Kinect with the projector using Microsoft RoomAlive SDK, and then, based on the deck map with Kinect, I could use edge detection, and also line and stair pattern recognition, to recognize the position of the different stair edges, and then project back the result to the stairs. We also combined hand rotation information from the IMU to correct the projection results. This image shows the result of how we projected the visual highlights to the real stairs.

The system actually achieves a framerate of 20 frames per second, which is pretty good if you render all the information on a digital display. The problem is, now we are projecting the virtual information to the real environment and also, this environment itself is very challenging because the surface-level changes happen very frequently. When a person walked with this actual system, they still can feel a little bit a lag from the system, because of the recognition time. As a next step to address this problem, I wanted to add a position-tracking component to the system to compensate for the recognition delay. How I want to do this is, first, I am considering adding an indoor localization component to track the user's position, and another option I also am thinking about is using a state-of-the-art motion-capture system, for example a state-of-the-art motion-capture system, for example, Percetion Neuron, a full-body tracking system that can hopefully help me extract the absolute position on the stairs, so that we can better refine the projection result for the system.

These are the two examples of my research on how I designed augmented-reality systems to enhance low-vision people's visual ability in different daily tasks.

AR is a very powerful technology, and I believe that there are actually a lot of things we can do in terms of accessibility by using AR to enhance people's ability. Besides my current research on enhancing people's ability in very fundamental level and individual-level tasks, for example, navigation, reading, visual search, these tasks, I think there are also a lot of things we can do. I'm very interested in a more complex task that involves multiple users,for example a social interaction, a work collaboration task, and I think it's very interesting to think what's the role AR will play, because that will also involve a lot of the perceptions of the bystanders, and also potential privacy and social-acceptance problems.

Besides people with low vision whether we can extend my AR systems to is different user groups. For example, people with different perceptual disabilities, people with autism, people with learning disabilities. This image is one of my projects, where we tried to use Magic Leap to recognize the facial expression of a person, to help children with autism to better recognize other people's facial expressions. You can see we generated different virtual emojis to take the place of the other people's original head position, so that kids with autism can better perceive these emojis, in order to better understand people's visual expression.

While AR is a very powerful technology, at the same time, this technology itself can also bring a lot of accessibility issues. Obviously, this technology primarily is very vision dominant, which is very difficult for people with visual impairments to use, which includes both people who are blind and low vision. In order to design systems to enhance people's visual ability, I'm also very interested in how we can shape this technology to make it more accessible and usable for people with diverse abilities.

For example, some of my research also looks into how we can design assistive technologies to enhance the accessibility of virtual-reality technology. One example is, I designed a haptic controller. I call it "cane-controller" because it simulates the real world cane interaction to enable blind people to navigate a virtual-reality environment. Another one focused on people with low vision is called CMVR, where I designed 14 different low-vision tools to enable low-vision people to better perceive different VR scenes.

While there has been this different work that I have done to enhmace the accessibility of virtual reality, I do believe there's a lot of work to do for augmented reality, because AR is not only about the virtual space. It's actually a combination of the virtual space and the physical space. So, what information from the virtual and physical space we should accessibly convey to people with disabilities, and how can we convey the relationship between these two worlds? And also, besides making the technology itself accessible, how can we maintain the immersive experience that AR and VR are supposed to provide for people? These are important and interesting questions that I want to look into in the future.

By saying that, because this is a very large field, of course I won't be able to work on this research by myself. So, I'm actually looking for PhD students who are interested in this direction to join me. I'm not very sure about the distribution of the audience, but if you are interested in accessibility AR/VR research, and if you want to consider a PhD position, please feel free to contact me.

Today, I talked about two of my research projects, CueSee and StairLight, which are all AR systems that can provide direct visual augmentations to help people with low vision to perceive and understand the surrounding environment. In general, I do myself believe in the future of augmented and virtual reality, and I believe that, maybe in the next 20 years, everyone will wear smartglasses, and the environment will be appended with projectors, so that all the physical surfaces are interactive. To achieve that goal in the future, I think it's very important for us to think about how to best leverage, and also shape, this technology to make it better serve and empower people with diverse abilities.

Yup, so that's all my talk. Now I'm open to answer questions!

THOMAS LOGAN: Thank you so much, Yuhang! That was awesome. Hey, everyone, we'll now open up the floor for questions. Those of you on the YouTube stream... several of you, please type your questions into the chat on YouTube, and Stacy will ask them here in the room. Just to kick off for those in the room, does anyone have a question for Yuhang?

Great presentation, thank you!

I'll start with one just to kick it off. Yuhang, I was curious for the -- you mentioned, you know, at the start there's lots of different types of low vision users or people -- there's all these different types of diagnoses and the ways that the field of vision can be affected. Starting with CueSee, and you were showing the black-and-gray flashing solution. Were there any differences between, like, preferences for people with certain categories of low vision for, like, one solution or the other? Was that part of the research?

YUHANG ZHAO: Yeah, it is, actually. So let me get back to my slides with all the cue examples. That's actually one of the research goals, as well, besides, we want to validate the effectiveness of the cues. This is a very good question. It's tricky because, even thoug there are many different low-vision condition categories, low-vision people have very mixed visual conditions. They can both have very blurred vision, and also lose part of their field of view. At the beginning we didn't expect we can find any specific patterns, that's not the major goal. But, during the study, we did find that people have preferred patterns. For example, the Guideline and the Spotlight, I didn't go into the details of the design. But here if you look at the Guideline, it's very simple, it's very subtle. We just connected the center of the target to the center of user's field of view, and the Spotlight, we make the background grayscale but keep the original color of the target. These two are mostly preferred by people with relatively good vision -- not really a specific visual condition, but people who have a better field of view, and better visual acuity, because these designs are subtle. It won't distract them a lot, but still they are able to notice these visual cues to find a target.

But, as opposed to these relatively subtle designs, if you look at the Sunrays, the third one in the first row. There are eight lines that extending from the target to their full field of view. And, actually, the flash one. These two are relatively dramatic cues, but they are very effective for people with ultra low vision, which means that they have very low visual acuity. A lot of people have blind spots in their field of vision, it's very hard for us to know which part of their field of view they actually have or do not have. With all these eight lines, as long as the user can locate two of them, they can follow the lines to the target. So, these two are very effective for people with severe low vision. Unfortunately, movement -- none of my participants selected it because it's just too much, too overwhelming for it. Especially this location -- this movement, this rotation, disabled them from seeing the details of the product, even though they know where it is, they still want to verify the position of the exact target that they are looking for. So, this one is not a very successful design. Of course, it doesn't mean that movement in general is not a good cue, because people's peripheral vision is very sensitive to different types of movement. But probably rotation is not the best way to convey it.

THOMAS LOGAN: Thank you so much. That was great. And I'm glad you went into the details of these. It's really interesting. Do we have any questions in the room? And you can type them into the chat.

DAVID: Thomas, it's David. Can you hear us?

THOMAS LOGAN: Go ahead.

DAVID: All right. When you looked at the Stairlight and talked about safety -- how did you discount people perhaps hadn't had sufficient training? [bad audio]

YUHANG ZHAO: I'm not sure I heard the question clearly. Are you're asking, when I evaluated the Stairlight, how do I deal with how to train the participants while using the system?

DAVID: No, it's more a case of had they actually used a long cane properly?

YUHANG ZHAO: Oh, whether they can use a --

DAVID: I'm sorry, I'm not sure my question is coming through very clearly. I work in the are of visual impairment, and one of the issues with stairs, we look at, does the person actually use them to identify the stairs? You mentioned there is a very small cohort who actually did. Is that correct?

YUHANG ZHAO: Yep.

DAVID: How did you discount these guys haven't been trained properly, or haven't had the sufficient training?

YUHANG ZHAO: So you're saying, how can I decide whether the participants had properly training of the white cane, whether that's the reason they didn't use the white cane. We didn't verify that, unfortunately. During our study, the first thing, when they come to our lab to conduct the study, is we asked them what tools they use, and four of them decided to use a white cane in the study, but a lot of them actually had the white cane, and we did ask them whether they know how to use a white cane. It actually depends, because a lot of low-vision people, it's not about whether they have the proper training of the white cane, it's whether they want to use the white cane.

I'm not saying that the white cane is not an effective tool for people with low vision. The fact is they probably are very effective. But the truth is, a lot of them, even though they know they can use the white cane, they don't want to. A lot of them, first of all, they have jobs and they don't want to expose their disability related to low vision, and also a lot of people mention that, if they are walking with a white cane, they can be targets of some potential crime. Many of them because they still have functional vision, so they see themselves as being able to use their vision, and they also don't want to use the white cane because of their own preferences. There is a lot of personal emotions and social acceptance there.

That's the reason StairLight is not really designed to replace the white cane. That's why, in our study, we also allowed people to use the white cane together with the Stairlight, if they wanted to. The major goal of this project is more about, can we provide another alternative if people do not want to use the white cane in their daily life.

THOMAS LOGAN: I had another question related to the stair model. You showed the dull yellow and blue as different colors to use for identifying the stairs. Same question related to your previous project, but was their preferences based off of low-vision diagnosis for certain colors?

YUHANG ZHAO: Yeah, that's a really good question. Very different from CueSee, where we find that people have very different preferences, for StairLight people have much more consistent preference patterns. Across most of the participants, it's the combination of using a bright yellow thick highlight on the first and the last stairs, and use yellow thick -- thin highlights on the middle of the stairs. It's this very basic combination. This is what the most users would prefer. Even though there are some users that do use the blue color. But, blue and dull yellow, they are challenging colors for some of the low-vision participants. That's why they want this -- they still prefer this relatively bright yellow color on the stairs.

One interesting thing is, even from the exploratory study, many people mentioned that they did not want highlights on the middle stairs. But when we had to help people decide what they actually want to use in the system,I think 13 of them decided to use highlights on the middle stairs. If you look at the demo, you can see -- if they see this dynamic process of the highlights appearing on the stairs, people feel very safe about seeing the feedback all the time. If they don't see them, they feel very insecure.

THOMAS LOGAN: Thank you. Any other questions in the audience?

JD DIAMOND: Yes, please. Very great presentation. Thank you very much. I wish if I were a student, I could join you. The example of the cane that you mentioned is kind of fascinating to me. Is that something that you've built out and what kind of mechanism, if so, would you use to build to execute that kind of VR cane?

YUHANG ZHAO: Oh, you mean the Canetroller at the end? That is actually a brake mechanism that people wear in front of their body. Right in front of the user is a brake mechanism, and there's a slider extended to the front and we attach --

[technical issue]

YUHANG ZHAO: So it's basically a controller that people can wear right in front of them. It's a brake mechanism. There, we put a 3D-printed slider extended from this brake mechanism. This slider connects to half of the real white cane. If the user sweeps this controller by holding this hard white cane, then if the virtual objects in the virtual environment are hit by the virtual cane that represented by the actual controller, then this brake mechanism will shut on, so the user will feel this force feedback, because this brake mechanism will stop them sweeping, and they'll get this force feedback just like they hit a real object. At the same time we generate spatial audio to simulate the different sounds when the cane hits on different materials, and we also attach a voice coil to our control so that it can generate different leverage and patterns to simulate one that sweeps on different virtual materials like glass, carpet, concrete and so on.

JD DIAMOND: Very cool -- thank you!

THOMAS LOGAN: I'm going to ask one more question, since I don't hear another one. I was curious about the speech recognition -- adding speech recognition to VR solutions is -- since it's not really built in to most of the platform -- or for some of the platforms it is, some of them it isn't. I was just curious what was your process to get, like, speech recognition for CueSee, or just when you want to use speech recognition, do you have a preferred technology for that?

YUHANG ZHAO: Honestly I didn't want to implement CueSee like using, for example, Unity, using a typical AR and VR engine. I just built this based on basic programming technology. because it's more of a custom device, right? Attach a webcam with Oculus Rift. I just see Oculus as an external display. I just render all the information on my computer, based on python, because we use a lot of deep-learning models. Python would be the easiest programming language to do that. We also use Google voice recognition engine for the speech component. But it's not really invented into Unity, if that's what you're curious about. In this prototype, we just used a basic programming language, and generated an image on the computer display, and added it to the Oculus display, adding it as an external display.

THOMAS LOGAN: Thank you, that makes sense.

THOMAS LOGAN: Thank you. We have time for a couple more questions and if we don't have questions we can move to our social component. Does anybody else have questions in the room?

All right, well. I want to say thank you so much to Yuhang Zhao for the presentation today. It's tons of amazing research and I really look forward to the continued work you do at University of Wisconsin. If anyone in here is looking for PhD students, yeah, I can definitely tell there's gonna be lots of exciting research continuing from you and your students. So thank you so much for the presentation today.

YUHANG ZHAO: Thank you.