

City University of New York (CUNY)

## CUNY Academic Works

---

Publications and Research

New York City College of Technology

---

2020

### Object Recognition and Voice Assistant with Augmented Reality

Juan Estrella

*CUNY New York City College of Technology*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/ny\\_pubs/969](https://academicworks.cuny.edu/ny_pubs/969)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)



# Object Recognition and Voice Assistant with Augmented Reality

Student: Juan Estrella | Advisor: Carranza

Computer Engineering Technology Department

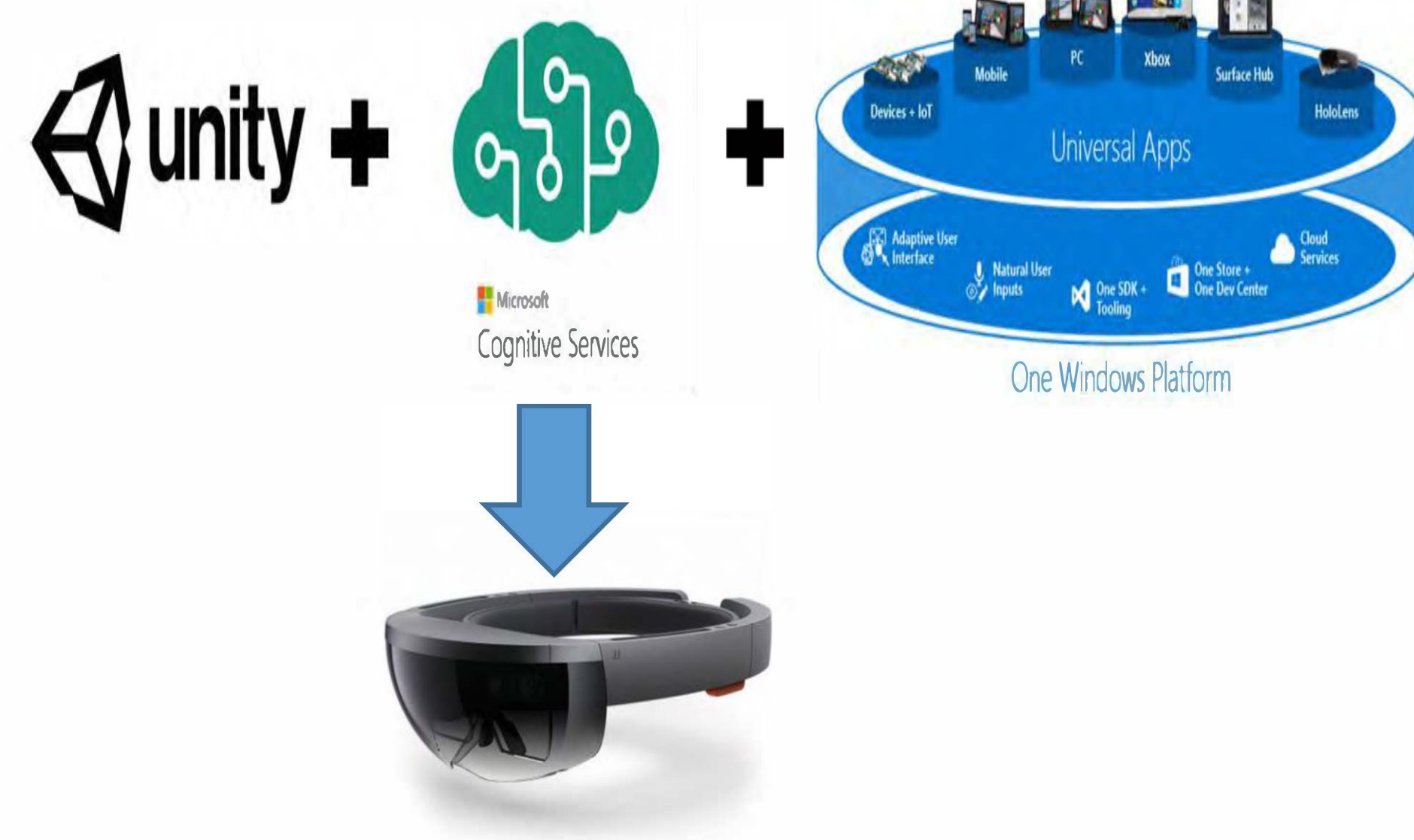
## Abstract

Our research project aims to provide a visually impaired person with a superimposed map that will guide the individual to the desired destination through a voice controlled virtual assistant application that integrates **Augmented Reality (AR)** with **Artificial Intelligence (AI)**. **Computer Vision** and **Natural Language Processing** (subfields of AI) will be combined to identify the spatial environment and then create a graphic enhancement that provides the most direct route to the specific destination. These technologies will be incorporated into the Microsoft Hololens, which will be controlled by the user.

## Materials and Methods

To build this app, we will integrate the following technologies:

- **Unity:** A video game engine to create the required virtual elements.
- **Hololens:** A fully untethered holographic computer that will blend optics and sensors to deliver 3D holograms that will be pinned to the real world.
- **Azure Cognitive Services:** Cloud services that are available to help developers build intelligence-based applications. Cognitive Services makes AI easily accessible to all developers, regardless of their technical expertise.
- Because the Hololens operates on a UWP (Universal Windows Platform), a proprietary software tool kit, the use of Windows will be required since the development libraries of Windows do not run on MacOS or any other Unix system.



## Introduction

- AI's versatility significantly benefits society in various realms, such as marketing, manufacturing, security, and healthcare. It is particularly advantageous in the healthcare industry, enabling faster diagnosis than that of humans in addition to self driving-services for the disabled.
- AI employs **Natural Language Processing**, the ability of computers to understand words through **Speech Recognition** or identification of the spoken words uttered by the user.
- When AI is combined with AR, a technology that superimposes virtual content, the user's real-world environment is enhanced, and the user is able to interact with this environment through certain gestures with the use of a Microsoft Hololens.
- The Microsoft Hololens relies on the "Airtap" gesture, which allows the user to select virtual objects and interact with them.
- To achieve this interaction, developers must surmount the challenge of finding a gateway that allows for the incorporation of AI into AR apps.

## Discussion

- AI and AR seem ideally suited to one another.
- In fact, AI relies on AR to display the desired visuals that enhance the user's environment.
- AI plays a big role in the building of intelligent **adaptive interfaces, object recognition and tracking, and gestural input.**
- Manipulation of the virtual environment through **tracking and voice commands** is the project's future aim.
- Visually Impaired people have difficulty getting to a specified location because of the challenge of trying to find their way in a strange environment. Additionally, human helpers may not always be available. However, the use of a Holographic Guider with real-time destination routing and voice capabilities increase the likelihood that the visually impaired will arrive at their desired destinations with efficiency.

## Our Work

- We will present an **AR application** in which an online **AI Bot** processes the user's **voice commands** to control the behavior of a **virtual avatar** to guide the user to a specific indoor destination such as a room.
- The **virtual avatar** is a **GPS based Bot** that will find the best route for the visually impaired user to follow to get to the specified destination.
- The **Bot** uses **Wayfinding**, which includes all the ways in which people orient themselves in physical space and navigate from place to place.
- We explored different alternatives for implementing the **AI Bot: Google's Cloud Speech Service, Microsoft Azure's Cognitive Services, and IBM's Watson API.** We decided to use **Microsoft Azure's Cognitive Services** because it was more supported by **Unity** than the other services.

## Results and Future Work

- Our app will respond to commands such as "Take me to the Bathroom," "Find Kitchen," "Locate Dining Room," and "Find Living Room"
- The integration of Azure's Cognitive Services with Unity consists of a script that triggers specific animations based on the structured data returned by the AI Bot, as a result of converting speech sentences to structured text.
- Azure's Cognitive Services are well supported and able to integrate with Unity and the Hololens. The combination is vigorous.
- The disadvantage is that **Azure's Cognitive Services** require an Internet connection.

## References

- Magnimind. "Invaluable Societal Benefits Of AI." Medium, Becoming Human: Artificial Intelligence Magazine, 17 June 2019, becominghuman.ai/invaluable-societal-benefits-of-ai-2ed62f7a653f.
- Kelley, Steven. "Seeing AI: Artificial Intelligence for Blind and Visually Impaired Users – Page 1235 – VisionAware." Seeing AI: Artificial Intelligence for Blind and Visually Impaired Users, Mar. 2020, visionaware.org/everyday-living/helpful-products/using-apps/seeing-ai-app/1235.

## Acknowledgements

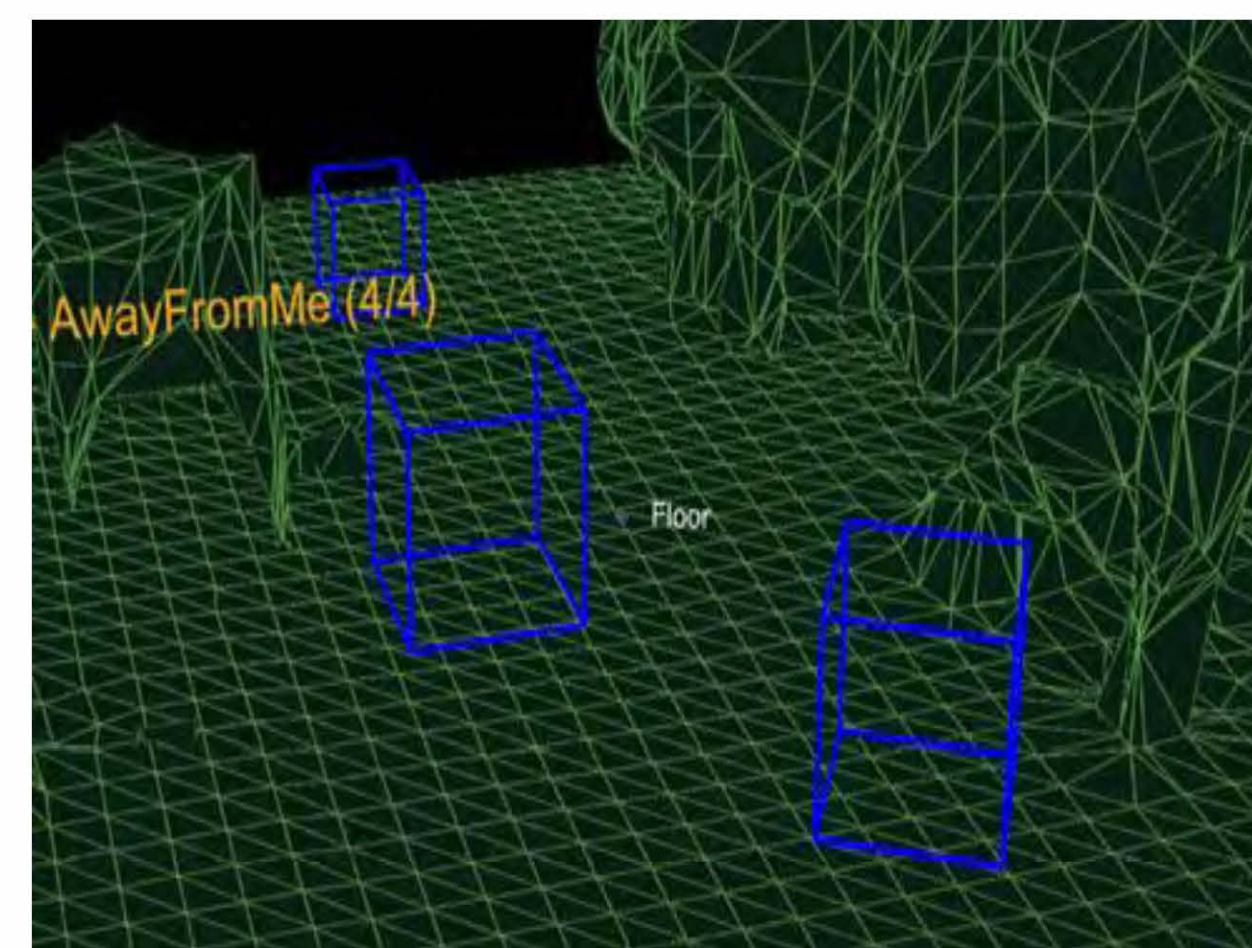
- Undergraduate Research Program
- CUNY Emerging Scholars Program
- GRTI 20 Grant, Project 4 "The AREngEd Project: Augmented Reality for Engineering Education"

Fall, 2020

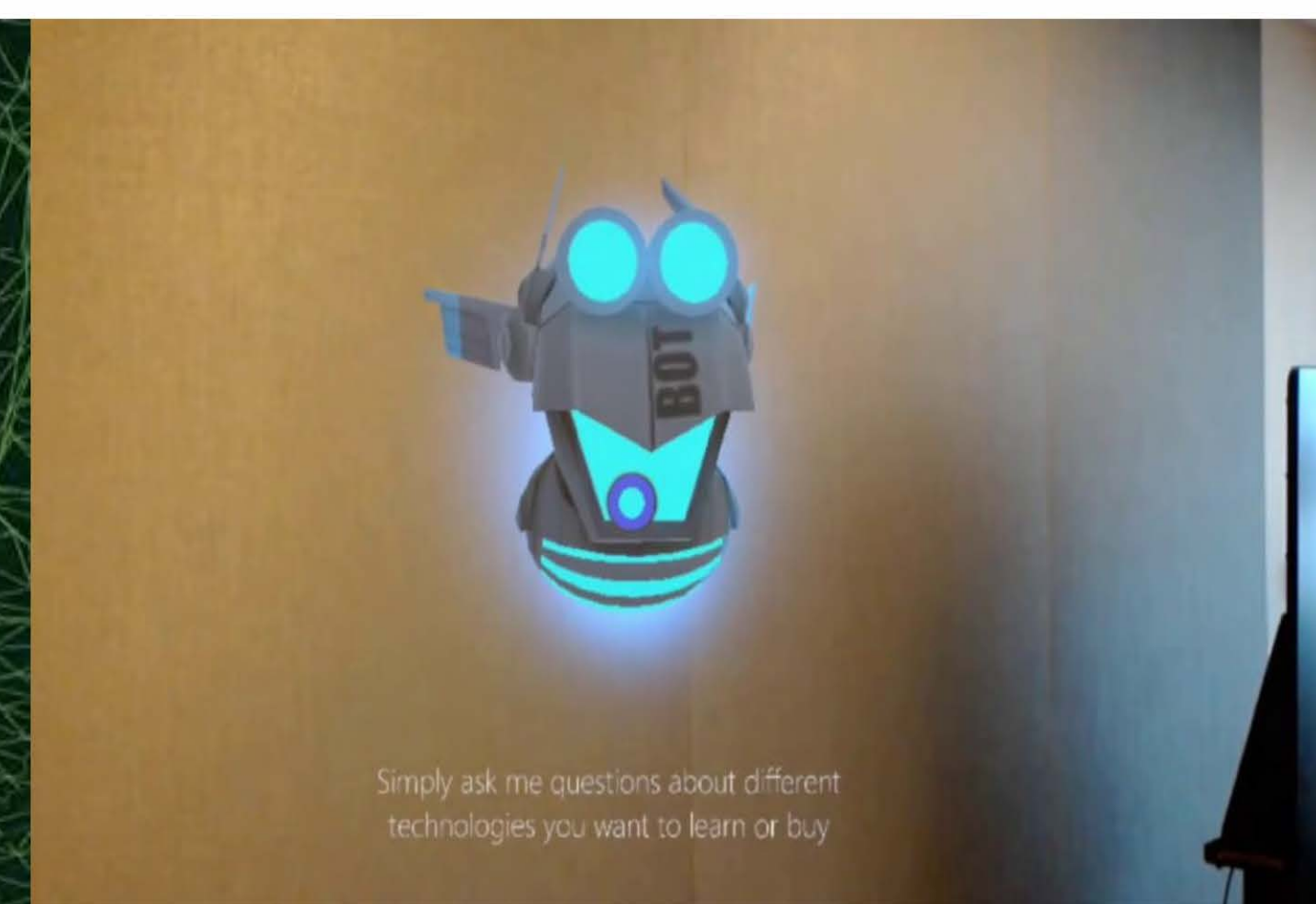
Hololens User



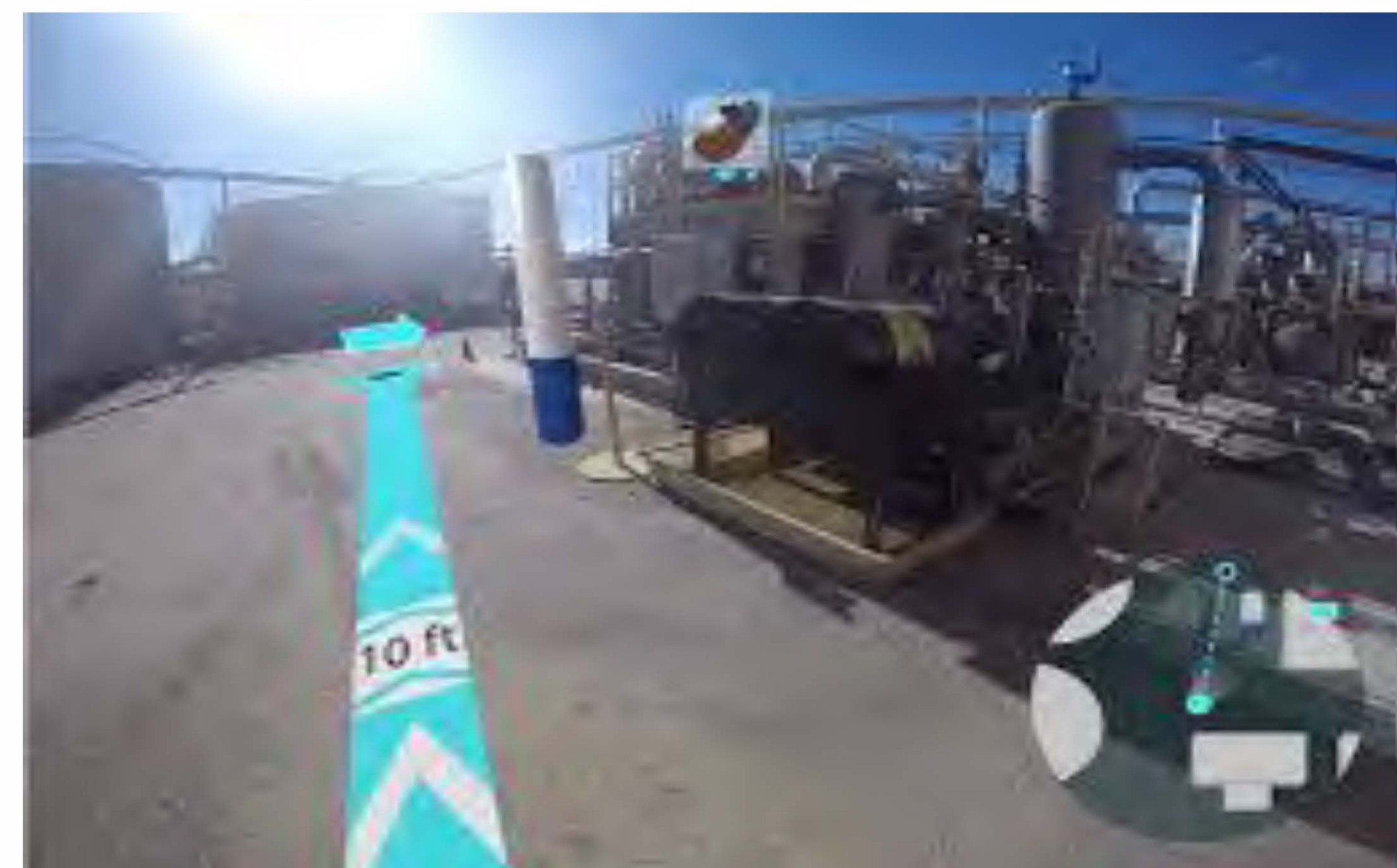
Hololens using "Spatial Awareness" to scan environment



Virtual AI Bot



Virtual Waypoint Indicator to guide the Visually Impaired



Tagged Objects for the Visual Impaired to See

