Electrical Engineering Design - ELECENG 4OI6

P.O.S.E

Project Proposal

Group 10

Ethan Yao	**yaoy41**	400119971
Jarvis Wang	**wangj230**	400122276
Jinge Li	**lij269**	400131282

Executive Summary

A smart home in its principle is to combine system, service, and management to provide people with an efficient, comfortable, safe, and accessible living environment. Most smart home devices are controlled by either an application on computing devices or voice commands, however, for the individuals who suffer from hearing and speech impairment, their ability to interact with the world is severely limited, let alone using the smart devices. According to WHO, over 5% of the world's population or 430 million people are suffering from hearing loss [1], their difficulties of interacting with household appliances should be addressed. P.O.S.E is a skeleton tracking-based smart home system that is aiming to improve the smart home device control experience and provide accessibility support for people who may have difficulties interacting with smart devices in traditional ways. The application's main usability is to allow users to control smart home devices by their body gestures. P.O.S.E brings together Skeleton Tracking, Raycast Detection, 3D Scanning, Unity Game Engine and IoT to create a real-time body gesture based smart home controlling system.

Objective and Motivation

Motivation

Modern smart home devices lack usability and consideration for individuals who suffer from hearing and speech impairment. To improve the smart home device control experience and provide accessibility support for the disability, a system that is capable of interpreting body gestures by using RGB and depth sensor are conceived.

Objective

Our objective is to allow a system to detect and recognize human body gestures and their spatial relationship with living spaces so that users can control their smart home devices by just pointing to the devices they desire to operate with their arms and fingers. An example usage will be turning on a ceiling light and adjusting its brightness by pointing to the device and twisting the hand.

Background

Nowadays, more and more companies are exploring and developing close-distance hand gestures detection systems. For example, in 2016, San Francisco-based Elliptic Labs released their EASY IoT software, which uses ultrasonic presence-detection to enable touch-less gesture capabilities. Which be embedded in devices such as smart thermostats, kitchen appliances, lighting controls, and security systems [2].

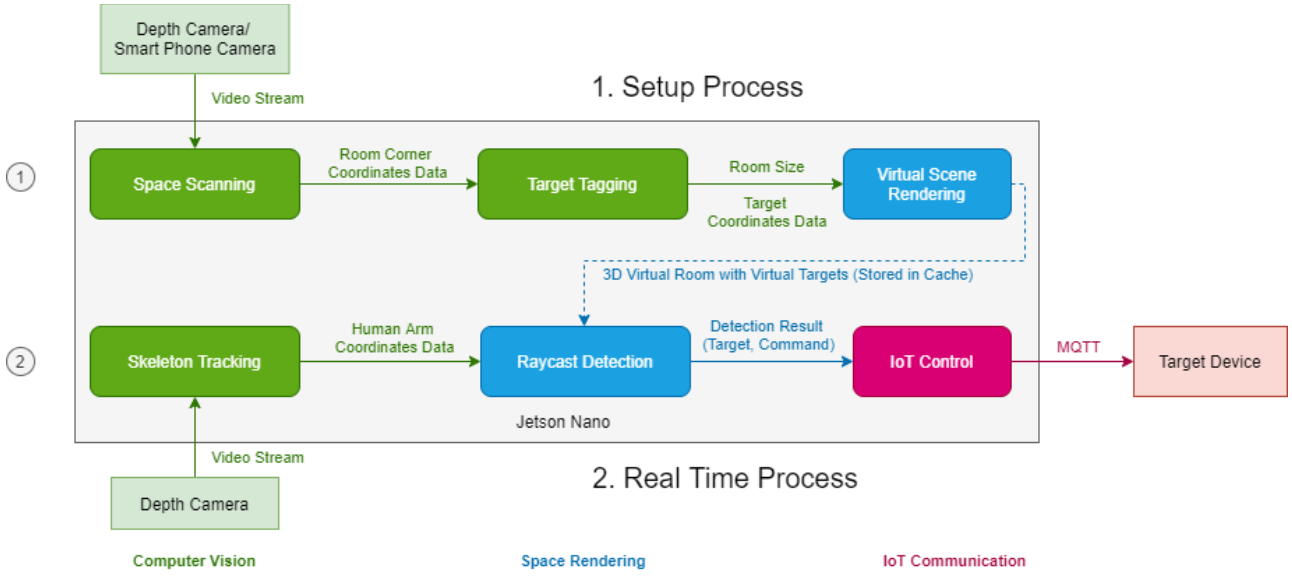
P.O.S.E expands this idea even future, users will not be constrained by the limited detection range of the traditional ultrasonic-based system, since P.O.S.E is able to identify users' body gestures and 3D geolocation anywhere in a room which immensely expands the targeted use cases and customers.

Step-by-Step Technical Approach

Three separate components will be used in a logical order to satisfy this system: computer vision, space rendering and IoT control.

- The computer vision part of this project will mainly focus on skeleton/body tracking and 3D scanning to retrieve the two 3d coordinates of the user's arm for the space rendering system to generate a pointer ray in real time and measure the size of the room and device's geolocation during the setup stage for later detection purposes.
- The space rendering system combines the control system and physics engine to render a virtual scene of the room during the setup stage and using raycast detection to detect real time collisions between the ray from user's arm and a virtual object generated by the relative position of the targeted device in the room.
- For the IoT control system, a low latency IoT protocol (MQTT) will be implemented to pass the data and control commands among the sensors, host machine and smart home devices so that the efficiency of the data stream can meet the performance requirements.

System Design



Resources

Hardware

- Depth camera: skeleton tracking, space scanning
- Host PC with a discrete graphics card: accelerating model training and 3D rendering
- Edge computing device: Raspberry Pi or Jetson Nano to run the entire system locally to ensure security and privacy
- Smart home devices: Light bulbs or switches to be controlled

Software

- Unity: Rendering virtual scene and performing raycast detection
- OpenPose [3]: Skeleton tracking models to train and use.
- PyTorch: Deep Learning framework to train space scanning and target tagging model
- Programming languages: C#, C++, Python

Manpower

- Two computer engineers and one mechatronics engineer

Budget

Item	Amount	Price
Intel Realsense L515 Depth Camera / Azure Kinect DK Camera	1	$500
Raspberry Pi 4B / Jetson Nano	1	$100
Smart Home Devices (Light Bulbs)	2	$20

Schedule of Activities and Milestones

Milestones

- **Milestone 1: Skeleton Tracking and Raycast Detection**
  - Develop and train skeleton tracking model to get real-time 2-point arm coordinates
  - The coordinates data will be sent to the rendering system to generate real time rays and perform raycast detection to a pre-built 3D object
- **Milestone 2: IoT Device Control and Target Tagging**
  - Develop an IoT control system with MQTT protocol to receive control command from the raycast application and forward to smart home devices like light bulbs
  - Develop and train a target tagging model to retrieve target device coordinates
- **Milestone 3: Space Scanning & Rendering**
  - Develop and train a room scanning model to get the room corner coordinates
  - Develop a scene rendering system to build a virtual scene containing the room and tagged devices based on the corner and device coordinates data
- **Milestone 4: System Integration and Testing**
  - Integrate the previous 3 milestones into one system and deploy on an edge computing device
  - Test the system with multiple smart home devices

Schedule of activities

		9/7/2021-9/19/2021	9/20/2021-10/3/2021	10/4/2021-10/17/2021	10/18/2021-10/31/2021	11/1/2021-11/14/2021	11/15/2021-11/28/2021	11/29/2021-12/12/2021	12/13/2021-12/26/2021
Tasks	Lead	Week 0	Week 2	Week 4	Week 6	Week 8	Week 10	Week 12	Week 14
Form Group and Decide Topic	All								
Project Proposal and Presentation	All								
Milestone 1 Completion	All								
Technology Review Meeting	All								
Milestone 2 Completion	All								
Progress Report	All								
		12/27/2021-1/9/2022	1/10/2022-1/23/2022	1/24/2022-2/6/2022	2/7/2022-2/20/2022	2/21/2022-3/6/2022	3/7/2022-3/20/2022	3/21/2022-4/3/2022	4/4/2022-4/17/2022

		Week 16	Week 18	Week 20	Week 22	Week 24	Week 26	Week 28	Week 30
Progress Presentation	All								
Milestone 3 Completion	All								
Milestone 4 Completion	All								
ECE Expo	All								

Risks and Alternative Plan

There are many difficulties in our project, and we are confident enough to resolve them along with the process of development. However, there are a couple uncertainties that we should predetermine an alternative plan to each of them.

- In real cases, the light condition varies from one room to another. Our sensor/camera may not be capable to work in some extreme conditions like a dark or black-painted room.
  Alternative plan: Multiple sensors with different functionalities may be used in this case to cover all light conditions. However, this will increase the budget of our product and eventually decrease the user's demand.
- For our current proposal, people with hand disabilities are not included in our potential users. However, as a device that is focused on helping people with disabilities, our usability should not be limited to users with healthy hands.
  Alternative plan: An add-on device that is designed for hand disabled person should be implemented. Such a device can put on users' disabled parts and send simulated arm movement signals to our system. So that hand disabled users can enjoy our product with an indiscriminate feeling like others.

Deliverables

Bronze target level: System has the basic body detection and device control functionalities based on virtual scene generated by user's manual measurement of their room and devices.

Silver target level: Based on bronze target level, the system can scan and render the virtual scene with error < 5%. In addition, user's arm movements should be detected with an accuracy above 90%.

Gold target level: Based on silver target level, in addition to body skeleton tracking, our system can detect user's gesture with an accuracy above 80%.

References

[1] "Deafness and hearing loss", 1 April 2021, https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss
[2] "Are gesture control devices the next big thing in home tech", 27 October 2016, https://www.builderonline.com/products/home-technology/are-gesture-control-devices-the-next-big-thing-in-home-tech_o
[3] "OpenPose", 2019, https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/