Head Tracking for Confirming User Selection: An Accessible Human-Machine Interface

Zeyad Ghulam[1,2], Greg French[1], Hussein A. Abdullah[1,3]

[1]School of Engineering, University of Guelph | [2]KITE Research Institute, UHN | [3]Australian University

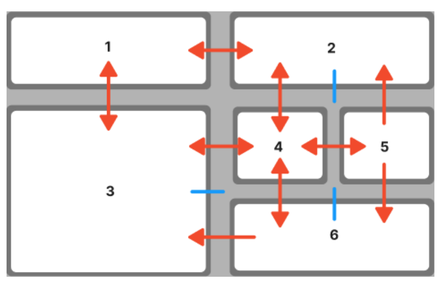
Introduction & Motivation

- Over 2.5 billion people globally rely on assistive technology, yet many with conditions like strokes, ALS, or cerebral palsy can't use hand-based systems (e.g., joysticks) or alternatives like eye-tracking and voice commands.
- The problem:
  - Joysticks need dexterity many lack
  - Eye-tracking demands perfect conditions + calibration
  - Voice control fails for non-verbal users
- We've pioneered a system using head tilts to navigate and mouth openings to confirm selections, it's a simple, hands-free and inclusive system.

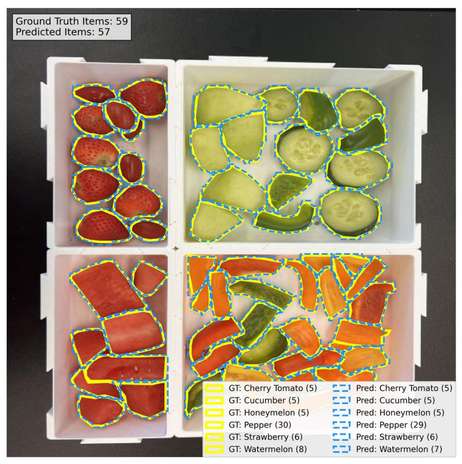
Methods

- **Hardware Setup:**
  - **Camera:** Intel RealSense D435i – captures 3D images with depth precision for reliable tracking
  - **Mount**: Fanuc F5 Robotic Arm – adjustable and sturdy for flexible positioning
  - **Items**: 3D-printed boxes (1x1, 2x1, 2x2 sizes) – modular and tangible for testing selections

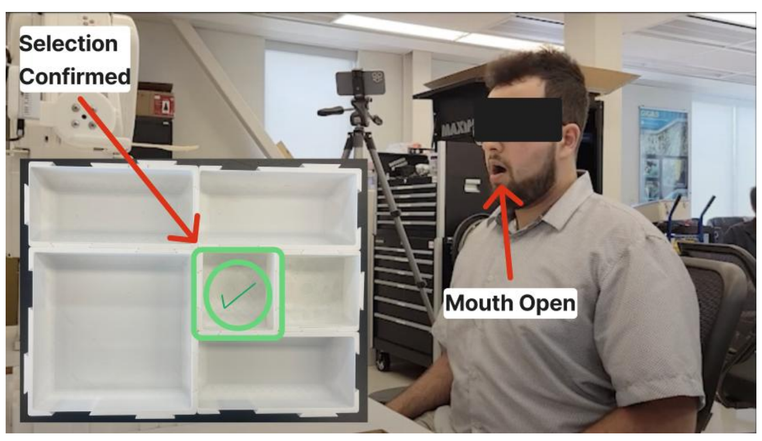




- **Software Tools**:
  - **Platform:** Python + ROS2 architecture
  - **Detection and Segmentation**: YOLOv9-seg – was chosen among different models including Detectron2 and Facebook's SAM. Used to track tray boxes as well as individual items found within the tray
  - **Facial Tracking**: MediaPipe – used to track the head and mouth with precision in real time



- **Test trial:**
  - 22 healthy participants (11 male, 11 female, average age 23) participated
  - A target location within the trays across each configuration was picked, directions (up, down, left, right) were called and the participants followed accordingly
  - 10 trials per configuration were conducted, each with a series of 5 head movements, and an open mouth command at the end to confirm the selection



Results

Head Tracking Accuracy (HTA)

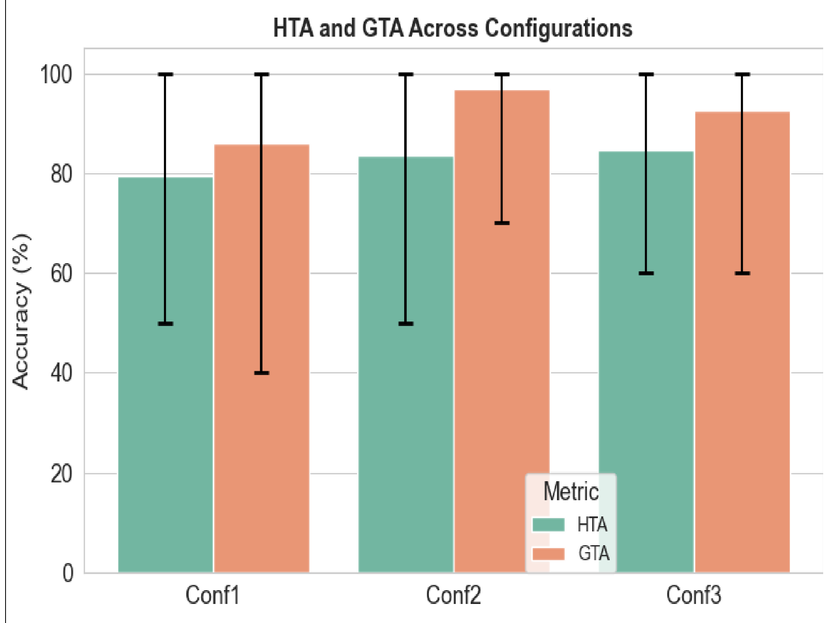
$$HTA_c = \left(\frac{\sum_{t=1}^{10} CorrectMovements_{t,c}}{TotalCommands_c}\right) \times 100\%$$

Where c = Config (1-3), t = trial number, with 5 commands per trial

Graph Traversability Accuracy (GTA)

$$GTA_c = \left(\frac{\sum ValidTransitions_c}{\sum AttemptedTransitions_c}\right) \times 100\%$$

A transition was valid only if the nodes traversed in legal movements



Head Transition Movements (HTM)

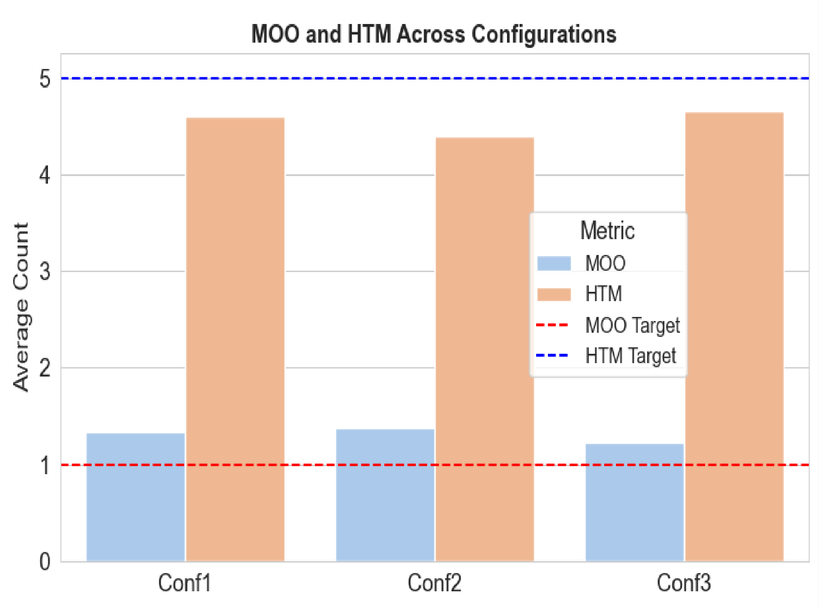
$$HTM_c = \frac{1}{N}\sum_{t=1}^{10} HeadMovementCount_{t,c}$$

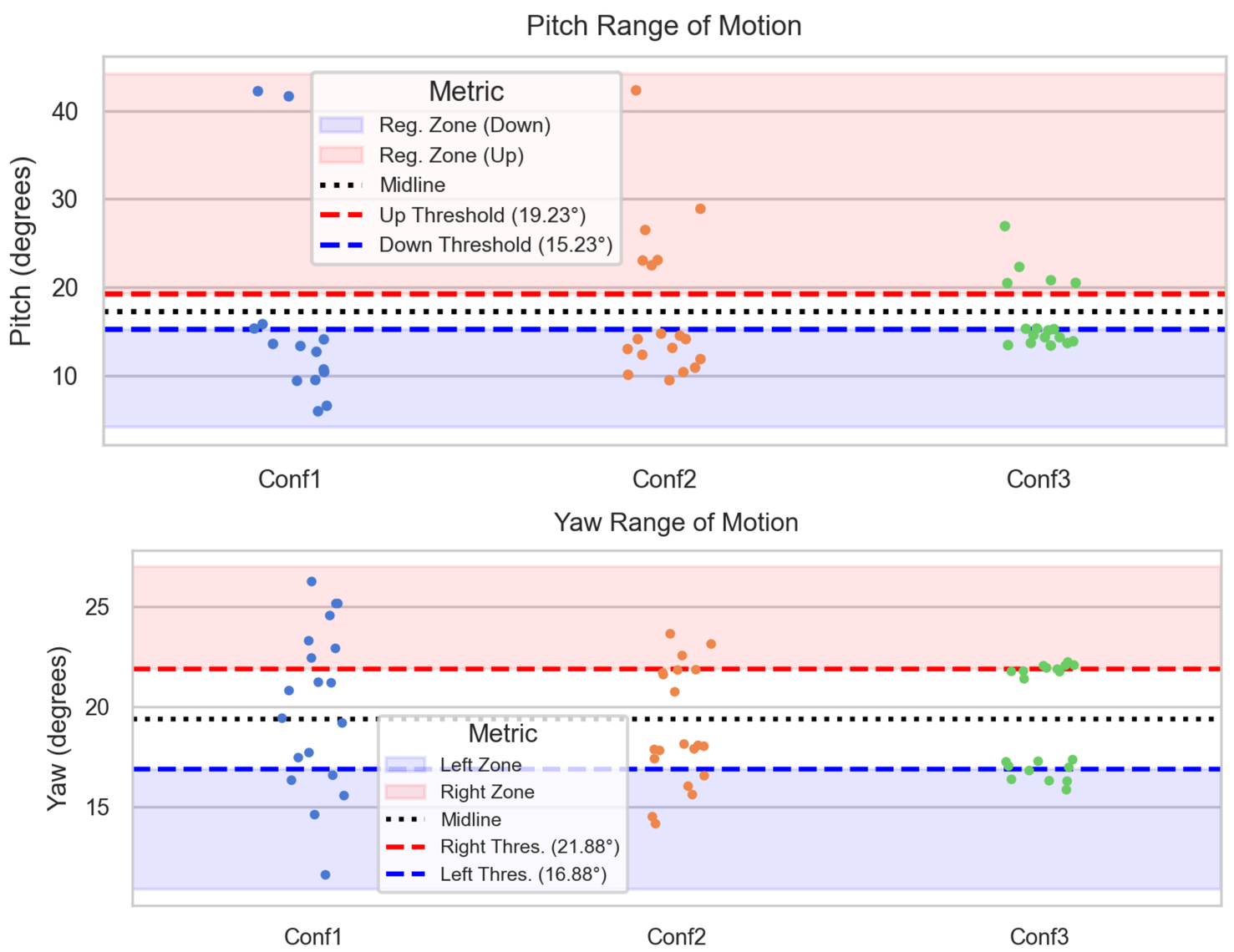
Where N = 10 trials per config. Ideal HTM = 5

Mouth Opening Occurrences (MOO)

$$MOO_c = \frac{1}{N}\sum_{t=1}^{10} MouthOpenCount_{t,c}$$

Ideal MOO = 1, reflects one mouth opening for final selection confirmation per trial





Discussion

- Results demonstrate effectiveness and usability of the head-tracking system for human-machine interaction. Key findings highlight both the system's reliability and the users' ability to adapt quickly to its mechanics.
  - **Learning effect and accuracy improvement**: Participants exhibited a clear learning curve. HTA increased from 79.5% in Config 1 to 84.7% in Config 3, while GTA rose from 85.9% to 92.6%
  - **Consistent performance**: The system maintained stable performance metrics throughout the study. MOO averaged approximately 1.3 per trials (target = 1). Similarly, HTM averaged 4.5 per trails (target = 5)
  - **Controlled movement and adaptation**: Participants' ROM in pitch and yaw became more controlled over time. By Config 3, pitch ROM stabilized at 16.7° and yaw at 19.3°, with reduced variability.
  - **User feedback and experiences**: Survey responses reflected high confidence in the system's Safety (9.69/10), Ease of Use (8.59/10) and Comfort (8.54/10).
  - **Future Work**: Evaluate the system with individuals who have upper-limb impairments. Also, implement user-specific calibration and develop a mechanism to adjust movement sensitivity aiming to further optimize performance and comfort.

zghulam@uoguelph.ca