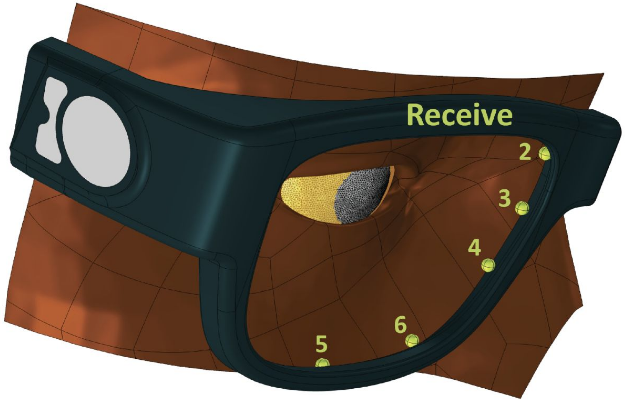
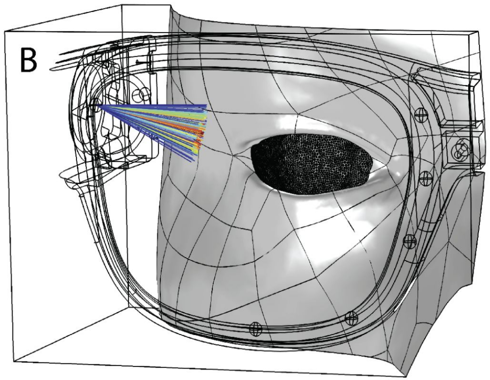
Ultrasound for Gaze Estimation—A Modeling and Empirical Study

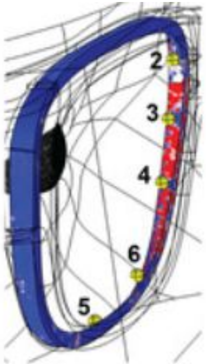
Aniket Fasate, Jeje Dennis, Madison O'Neil, Sofia Makowska

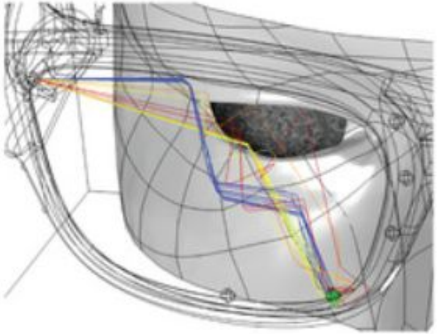


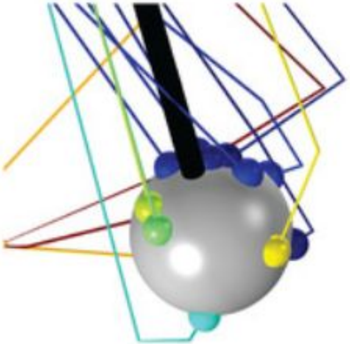


Novelty/Problem

- Most eye tracking sensors are camera based
  - Light sensitive which can result in inaccurate readings when the environment changes
  - Requires a lot of power to process at high speeds
- The alternative to these camera based sensors is utilizing ultrasound waves to track gaze via Capacitive Micromachined Ultrasonic Transducers (CMUTs)
  - CMUTS are a good size and resolution for VR and AR glasses







Approach to the problem

Simulation Modeling	Benchtop Experiments	Signal Processing	Machine Learning
Acoustic ray tracing for 1.7mHz transducers on AR glasses	a. Verify measurement accuracy b. Characterize transducers c. Gaze estimations	Extract ultrasound time of flight and amplitude data	Train nonlinear regression model

Sensors and their data

Sensors Used	Data Collected	Goal of Sensor Data
Ultrasound transducers (Capacitive Micromachined Ultrasonic Transducers - CMUTs) One transmitter + multiple receivers arranged around the eye area (glasses frame) Operated at ~1.74 MHz frequency	Time-of-flight: how long ultrasound takes to reflect off the eye Amplitude of reflections: signal strength varies with gaze angle Gaze labels from a mechanical goniometer simulating eye movements Tested with and without facial occlusions	Estimate eye gaze direction (horizontal and vertical angles) using ultrasound signal reflections Enable non-visual eye tracking — alternative to cameras or infrared-based methods Design a system that works in challenging environments: • Bright sunlight • Occlusion (e.g. glasses frame) • Low power requirements for wearable devices (AR/VR glasses)

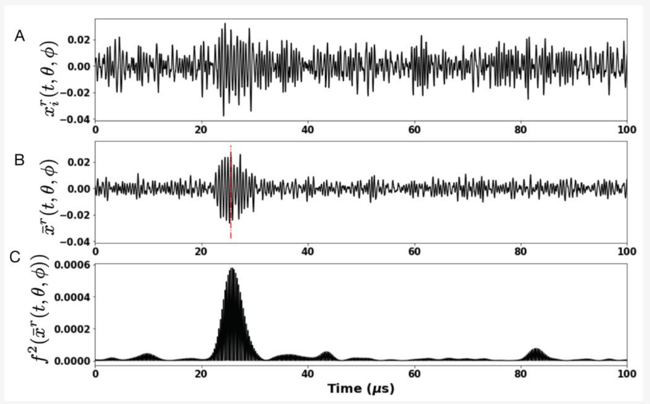
Methods, results, and conclusions

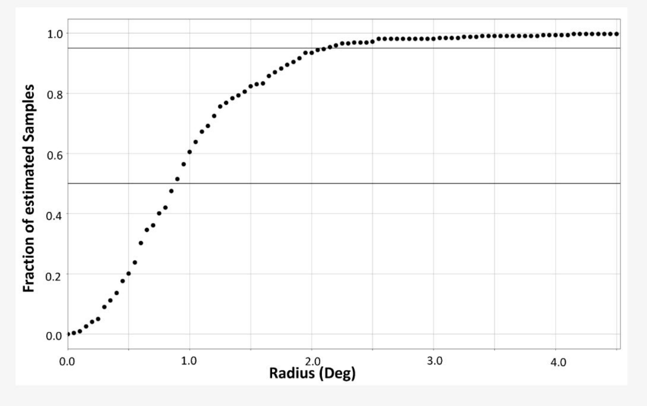
Signal Handling: The ultrasound signals are cleaned by averaging noisy traces and applying a butterworth band-pass filter (1.6-1.9 MHz). From these filtered signals, they extracted two key features time-of-flight and amplitude.

Modeling Approach: These featured signals are then fed into the regression models. The standout method was Gradient Boosted Regression Trees (GBRT), which outperformed a linear regression baseline in estimating horizontal and vertical gaze angles.

Result: The GBRT model achieved high accuracy in Gaze estimation with an RMSE error of ~0.97 degree. About 90% of gaze estimates were within 2 degree of ground truth, showing that both time-if-flight and amplitude features contribute meaningfully, even under occlusions.

Conclusions: This was the first experimental proof that ultrasound sensors can estimate gaze effectively. The approach is robust to ambient light, compact, adn low-power, making it promising for AR/VR devices.





Ideas for the final project

- The use of ultrasonic sensors with a single transmitter/multiple receivers can be used for high quality object detection in bright and occluded environments
- We can also use ultrasonic tracking as a control input depending on the application of the final project
- In order to reach the fidelity of the paper’s research we would need highly sensitive receivers and steady transmitters, ideally ones with low power consumption

Thank you! Questions?

Sensors and their data

Sensors Used	Data Collected	Goal of Sensor Data
Ultrasound transducers (Capacitive Micromachined Ultrasonic Transducers - CMUTs) One transmitter + multiple receivers arranged around the eye area (glasses frame) Operated at ~1.74 MHz frequency	Time-of-flight: how long ultrasound takes to reflect off the eye Amplitude of reflections: signal strength varies with gaze angle Gaze labels from a mechanical goniometer simulating eye movements Tested with and without facial occlusions	Estimate eye gaze direction (horizontal and vertical angles) using ultrasound signal reflections Enable non-visual eye tracking — alternative to cameras or infrared-based methods Design a system that works in challenging environments: ● Bright sunlight ● Occlusion (e.g. glasses frame) ● Low power requirements for wearable devices (AR/VR glasses)