ISMAR 2025

Siren Song: Acoustic Attacks on Pose Estimation in XR Headsets

Zijian Huang[1], Yicheng Zhang[2,3], Sophie Chen[1],

Nael Abu-Ghazaleh[2], Jiasi Chen[1]

yzhang95@gmu.edu

[1]*University of Michigan, Ann Arbor*

[2]*University of California, Riverside*

[3]*George Mason University*

Motivation

- XR has applications across many areas:

But…what if pose estimation is under attack?

- **Importance of pose estimation**:

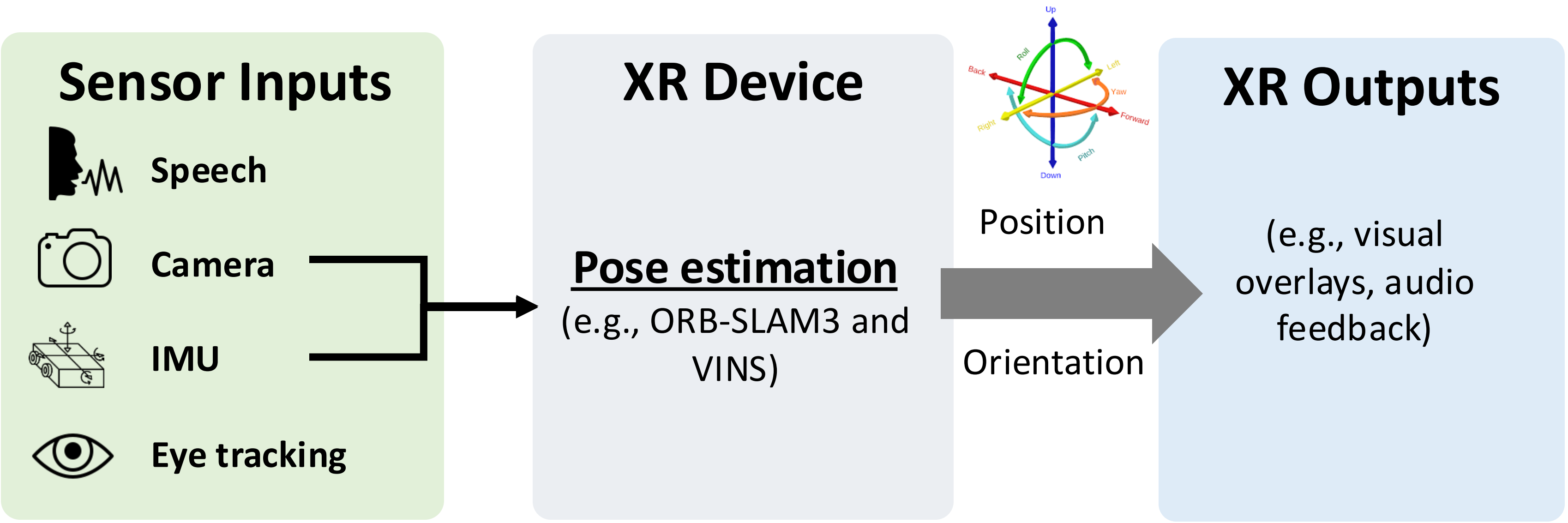
Rendering

User Interaction

VR sickness

Background: Pose estimation in XR

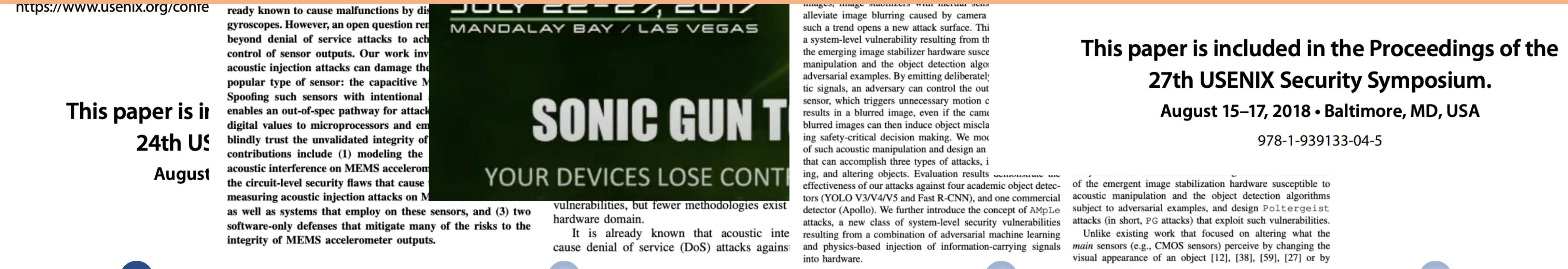
- XR devices fuse IMU and camera data to estimate the headset's location and orientation.



Background: Acoustic Attack

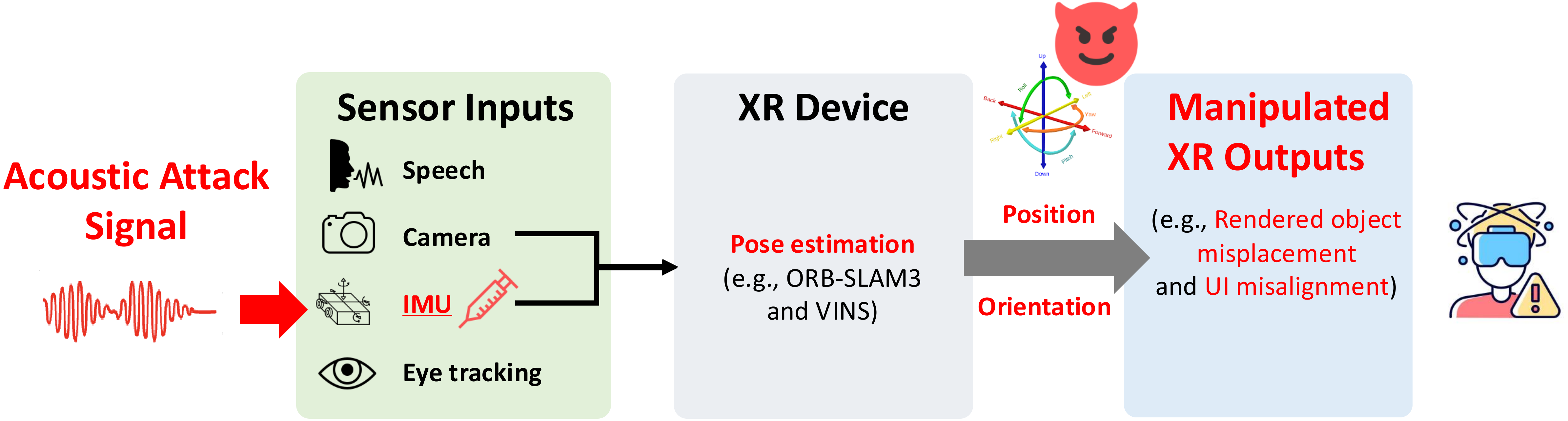
- Plenty of acoustic attacks on drones, vehicles, and phones.

- **Key Challenges:**
  - Noninvasive attack setup.
  - XR-specific pipeline.
  - Impact on user experience.

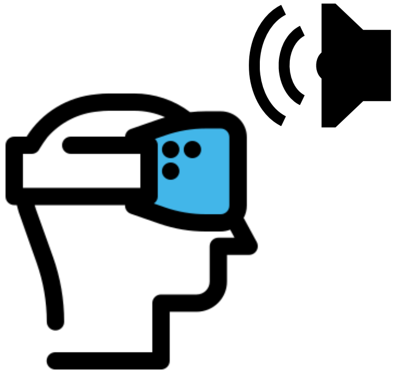


Background: Acoustic Attack

- Acoustic signals trigger MEMS resonance, corrupting IMU motion data.



Threat model

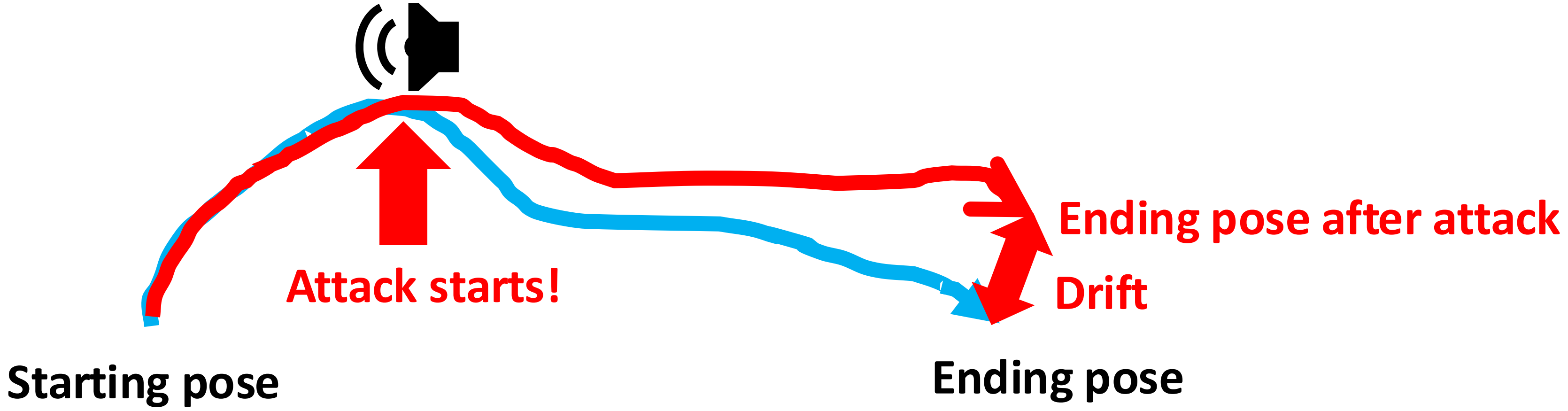


- **Attacker's capability:**
  - Play sound near victim — no physical or sensor access.
- **Attacker's resources:**
  - Profile same-model sensors to find resonant frequencies.
- **Attacker's goal:** benefit attacker (cheat) or harm victim (disrupt / induce errors).

Attacks on open-source systems

- Tested systems: ILLIXR [1] and ORB-SLAM3 [2].

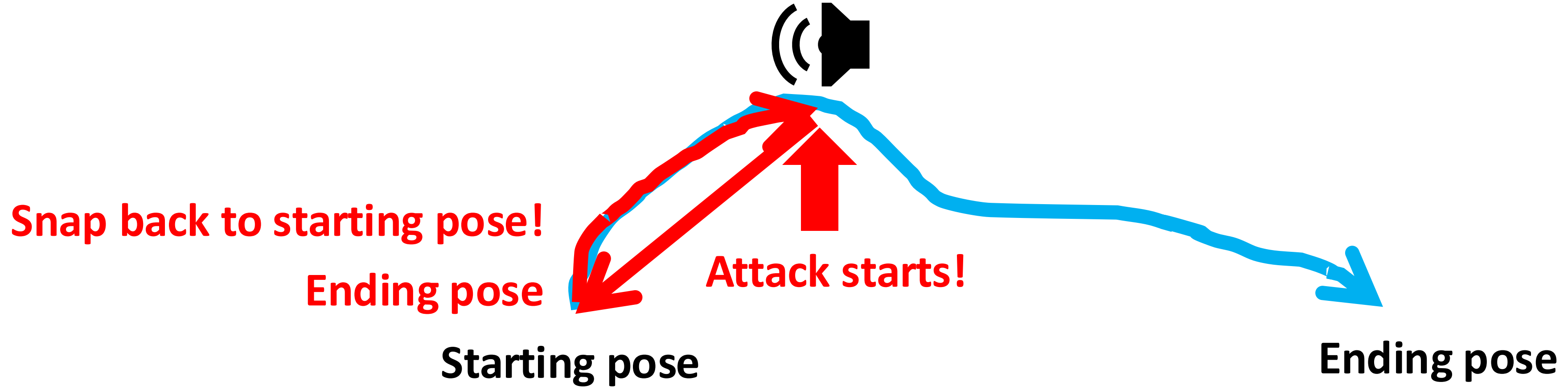
- **Slow drifting attack — when IMU bias is <u>small</u>.**
  - Small and slow pose drift.



[1] Huzaifa, Muhammad, et al. "ILLIXR: Enabling end-to-end extended reality research." *2021 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, 2021.
[2] Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam." *IEEE transactions on robotics* 37.6 (2021): 1874-1890.

Attacks on open-source systems

- **Slow drifting attack — when IMU bias is <u>small</u>.**
  - Steady and small pose bias.
- **Snapback attack — when IMU bias is <u>large</u>.**
  - Sudden pose "snap" to starting point.

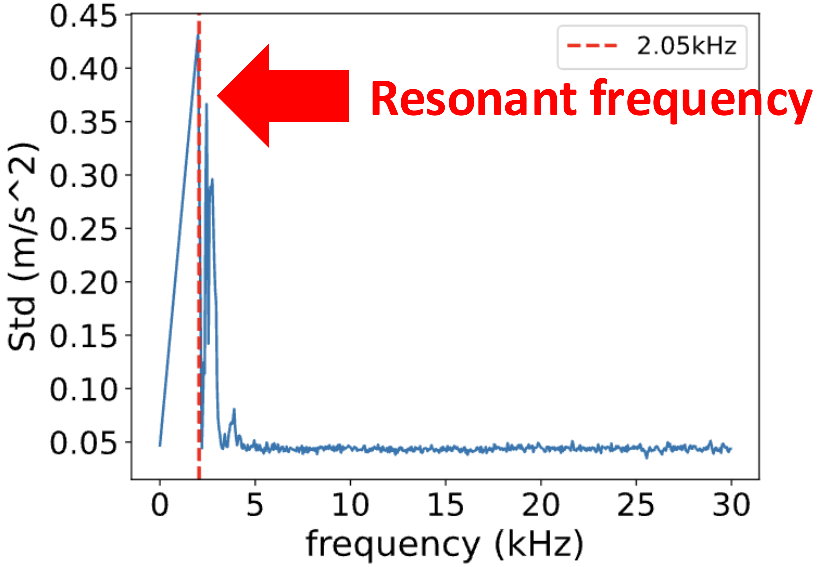


[1] Huzaifa, Muhammad, et al. "ILLIXR: Enabling end-to-end extended reality research." *2021 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, 2021.
[2] Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam." *IEEE transactions on robotics* 37.6 (2021): 1874-1890.

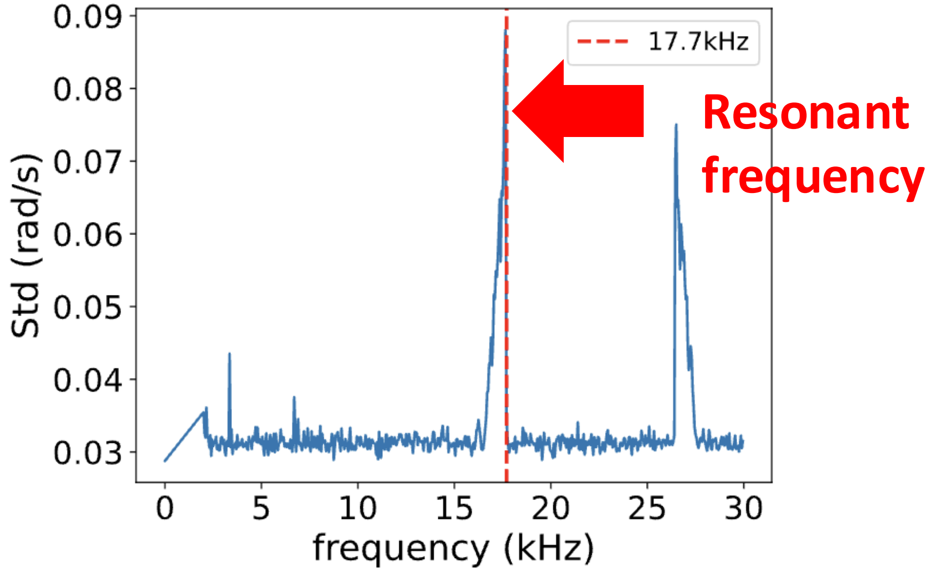
Attacks on Microsoft HoloLens

- **Resonant frequencies of the HoloLens IMU.**
- **Snapback attack (Success rate: > 90%).**

The accelerometer is vulnerable at 2.05 kHz.



Gyroscope is vulnerable at 17.7 kHz.



Attack 1: Manipulating user input (harm to user)

- **Attack motivation:** Many XR games map head motion to game controls.

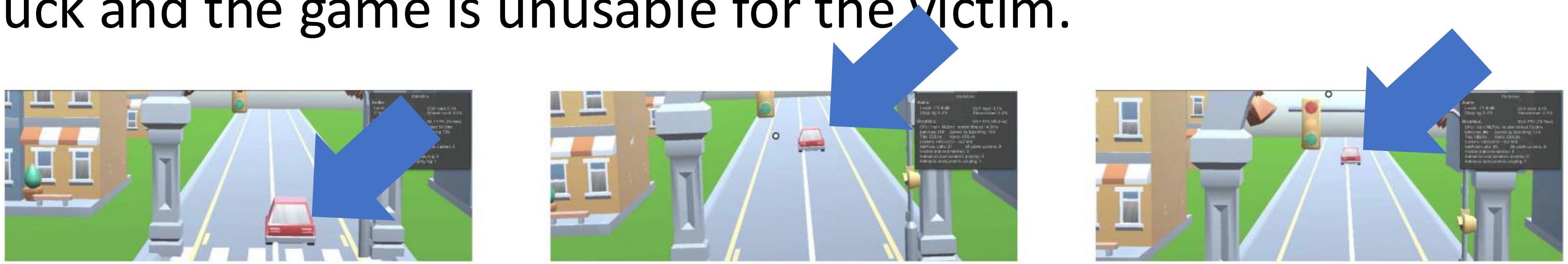
Attack 1: Manipulating user input (harm to user)

- **Attack motivation:** Many XR games map head motion to game controls.
- **Method (attack design):** We place a HoloLens on a remote-controlled car and inject acoustic signals targeting the IMU.

Attack 1: Manipulating user input (harm to user)

- **Outcome (attack result):**
  - The car initially moves but then shifts slightly.
  - Due to snapback, its estimated position resets to the start — the virtual car is stuck and the game is unusable for the victim.

Benign



(a) Benign, time=0 (b) Benign, time=1 (c) Benign, time=2

Car moving direction

Attack 2: Clickjacking (harm to user)

- **Motivation:** World-anchored UI (e.g., virtual keyboard) drifting breaks input and can destroy user actions.

Benign:
User can correctly type words on virtual keyboard.

The board is anchored in the fixed location.

Attack 2: Clickjacking (harm to user)

- **Outcome:** The virtual keyboard drifts or disappears

Attack:
user mis-clicks or cannot type.

Virtual keyboard drifts up and down, making user mis-click.

Attack 2: Clickjacking (harm to user)

- Small user study: goals & setup.
  - **RQ1 (Usability degradation) :** How does the acoustic attack affect usability and task performance?
  - **RQ2 (Impact of the acoustic signal):** Under what conditions is the acoustic signal more comfortable?

- **Participants & environment:** 5 volunteers, HoloLens 2, click-jacking app, 30-minute sessions.
  - RQ1: A timed typing task (three 3-letter words).
  - RQ2: Earmuffs or listening to music.

Attack 2: Clickjacking (harm to user)

- **Performance drop:** task completion fell from 100% (benign) to ~46.7% (attack).
- **Slower & harder:** average completion time rises (increasing from 6s to 16s); all participants said typing was harder under attack.
- **Practical note:** listening to music made the attack more tolerable and still effective — <u>stealthier</u> attack vector.

Conclusion

- First to study **acoustic attacks on XR pose estimation**.
- Identified **two novel effects** on XR pose estimation pipelines.
- Designed and demonstrated **four proof-of-concept attacks** on XR headsets.
- Conducted **user study** showing real usability degradation.
- Discussed **defense strategies**: hardware filtering, software re-localization.

Paper

Thank you!

Any questions?

Yicheng Zhang

yzhang95@gmu.edu

https://yichez.site

Personal Page