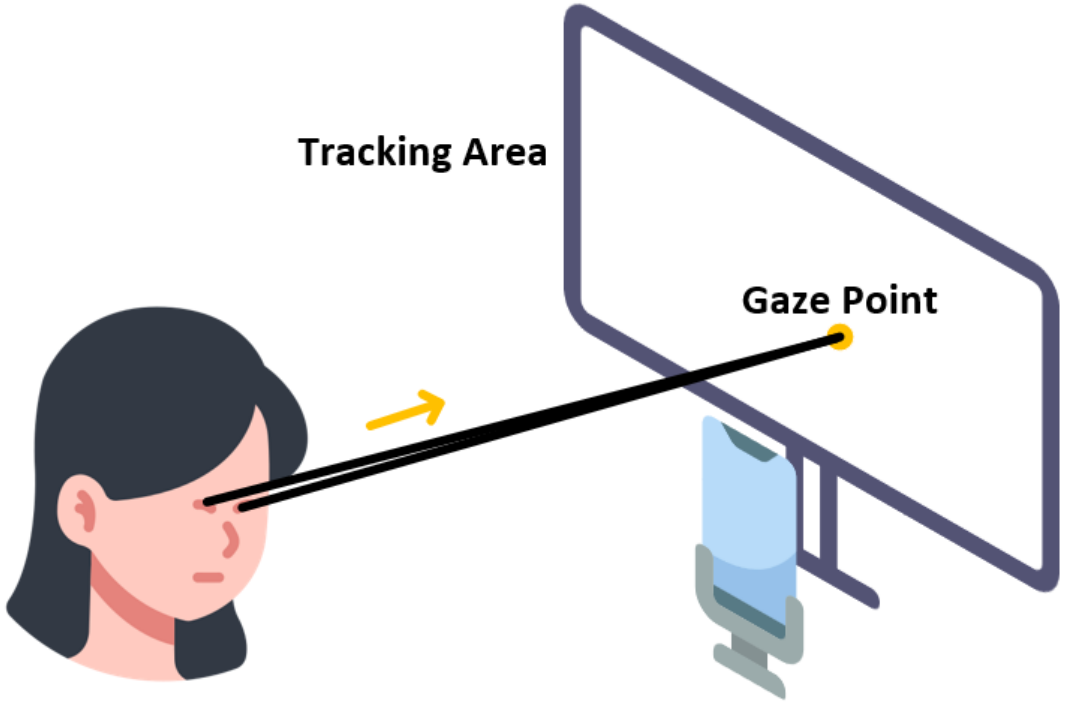
ASGaze:
Gaze Tracking on Any Surface with your Phone

Jiani Cao[1], Chengdong Lin[1], Yang Liu[2], Zhenjiang Li[1]

City University of Hong Kong[1], University of Cambridge[2]

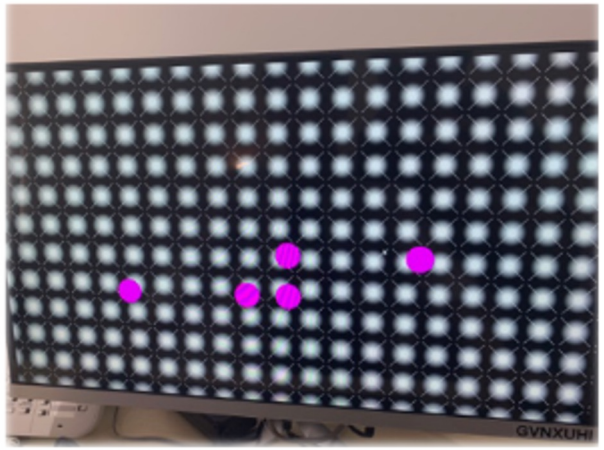
Definition

- What is gaze tracking?

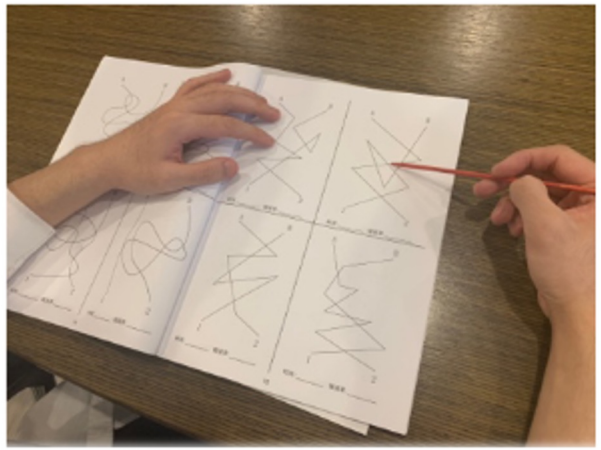


Motivation

Treatment and recovery of reading-disorder disease



Electronic Surface
(e.g., computer, phone)



Non-electronic Surface
(e.g., paper, whiteboard)

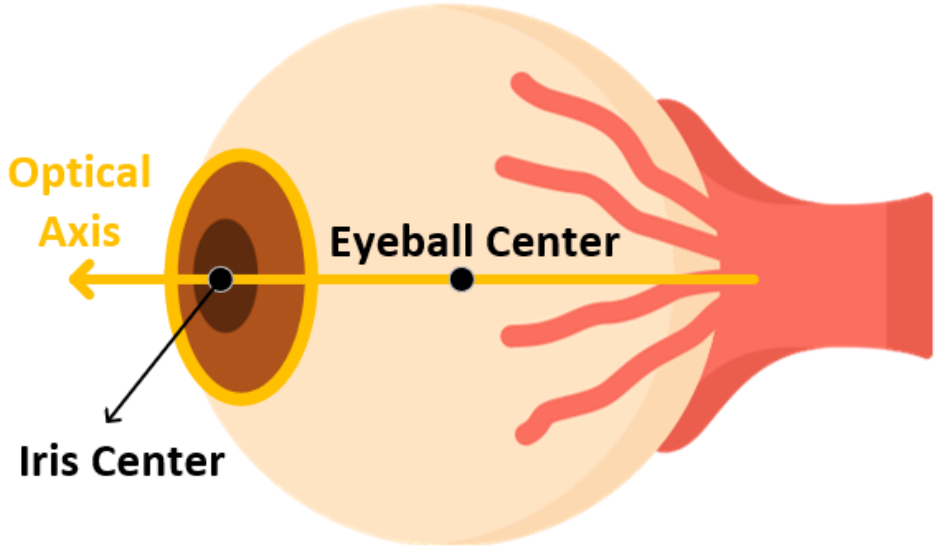
Doctors cannot fully understand the effectiveness of recovery.

Motivation

Our goal: an accurate, low-cost gaze tracker on any surface.

Existing Solutions

- Model-based approach

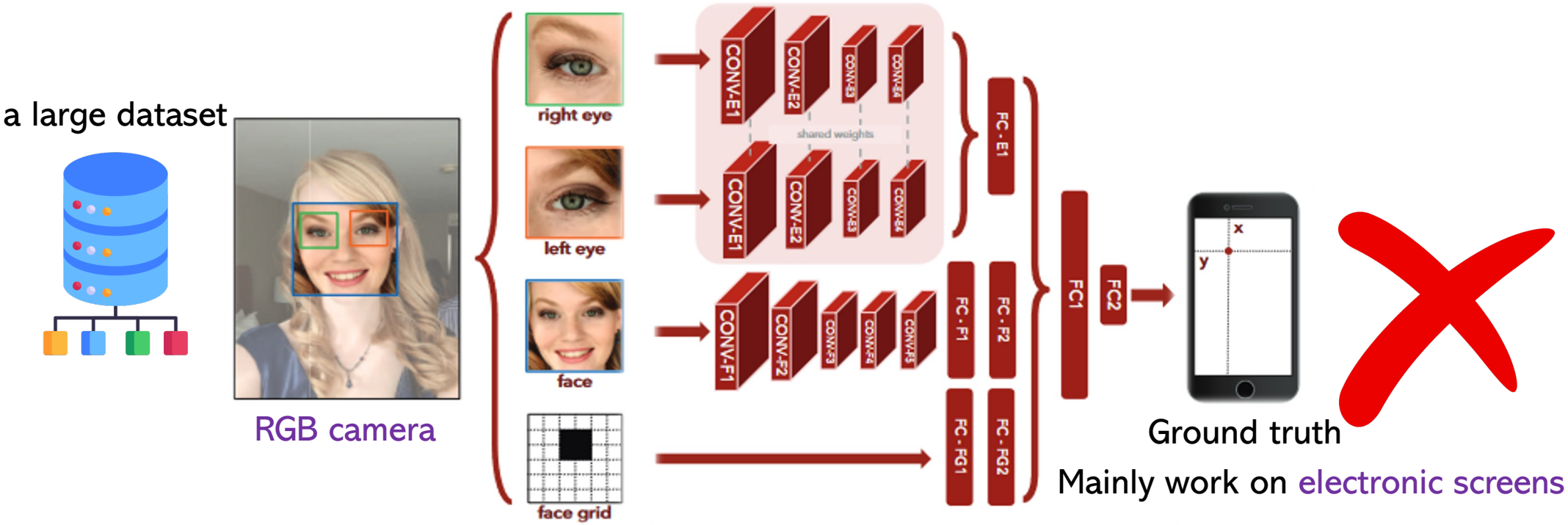


IR source, high-resolution camera

Expensive, e.g., 800~10,000 USD

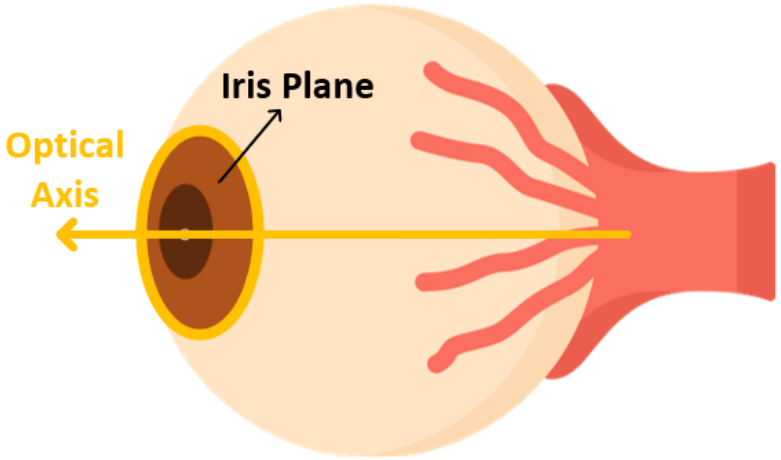
Existing Solutions

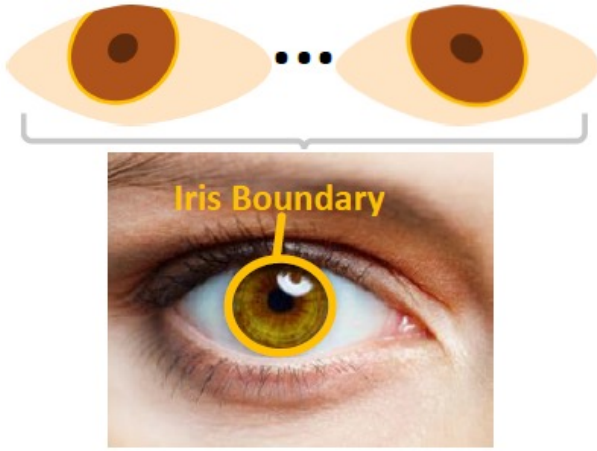
- Low-cost appearance-based approach



Mainly work on electronic screens

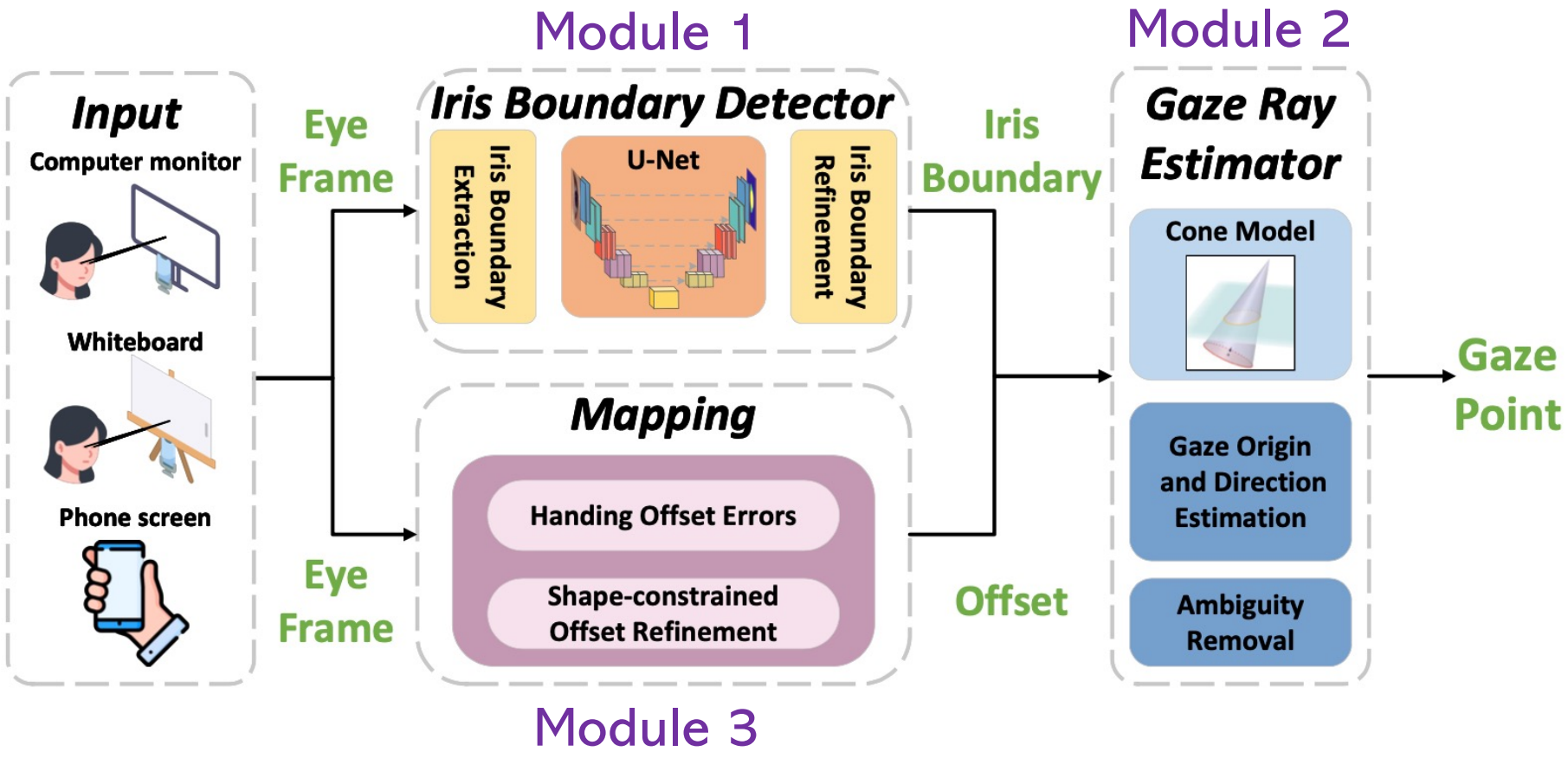
Tracking Principle





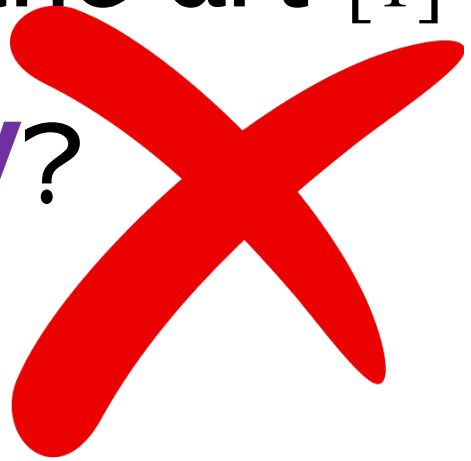
- Optical axis cannot be obtained directly
- Infer it from the shape of iris boundary

Our System: ASGaze



Module -1: Iris Boundary Detector

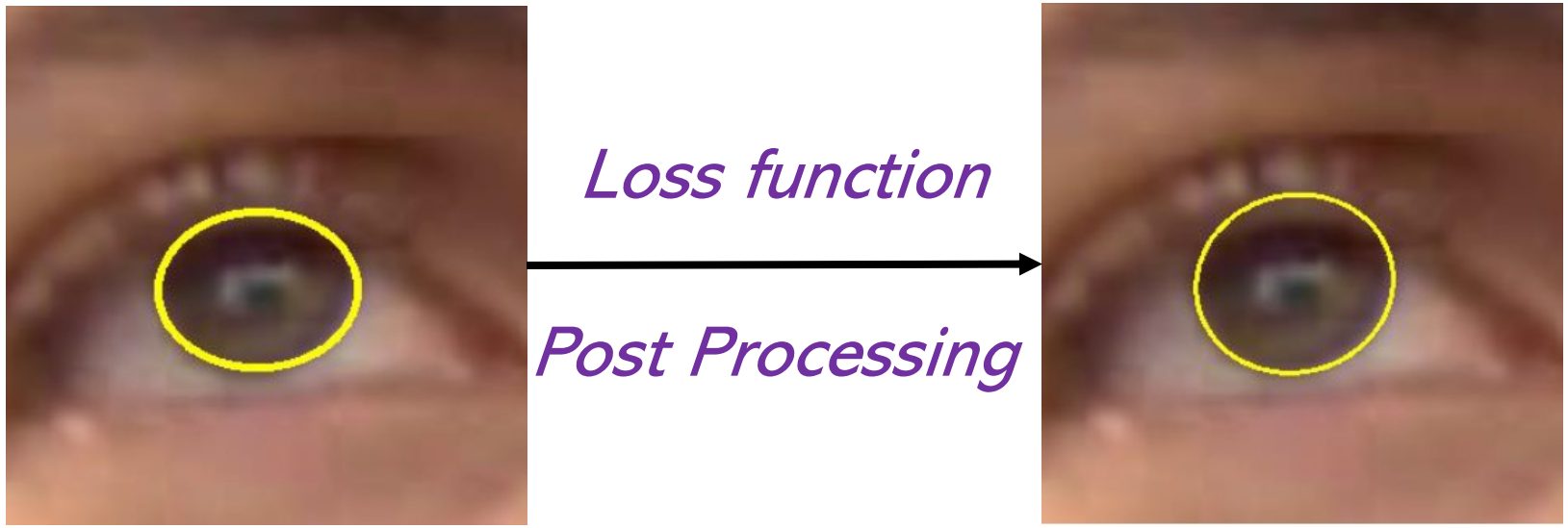
Can we use the state-of-the-art [1] design directly?



[1] C. Lin, X. Li, Z. Li, and J. Hou. Finding stars from fireworks: Improving non-cooperative iris tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022.

Module -1: Iris Boundary Detector

- The thickness (uncertainties) of iris boundary is not thin enough



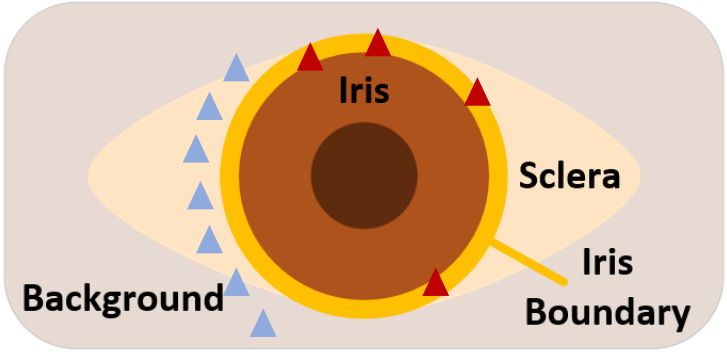
state-of-the-art

Ours

Class-imbalanced!

Module -1: Iris Boundary Detector

Loss function



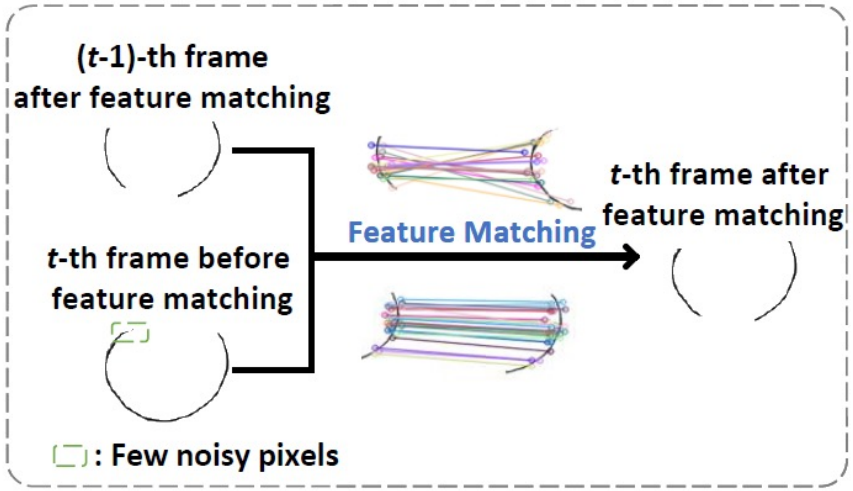
- L_1: $L_{ML} = -\sum_k \sum_i (1 - p_k(i))^\gamma \times I_k(i) \log(p_k(i)),$ Avoid missing pixel

$$\begin{cases} w \cdot L_{ML}, & x(i) \in \{iris\_boundary\}, \\ L_{ML}, & x(i) \notin \{iris\_boundary\}, \end{cases}$$

- L_2: $\sum_k \sum_i D(i) \times p_k(i),$ Remove noisy pixel
- L_3: $1 - 2\frac{\sum_k \beta_k \sum_i (I_k(i) \times p_k(i))}{\sum_k \beta_k \sum_i (I_k(i) + p_k(i))}$ Classification

overall loss: $L_1 + (1-\alpha)L_2 + \alpha L_3$

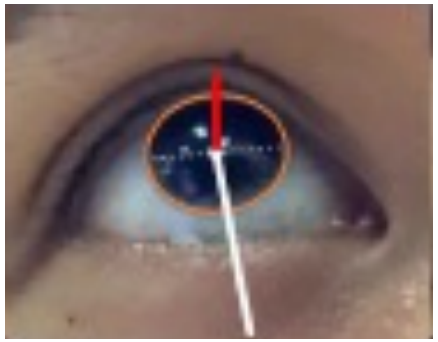
Post Processing



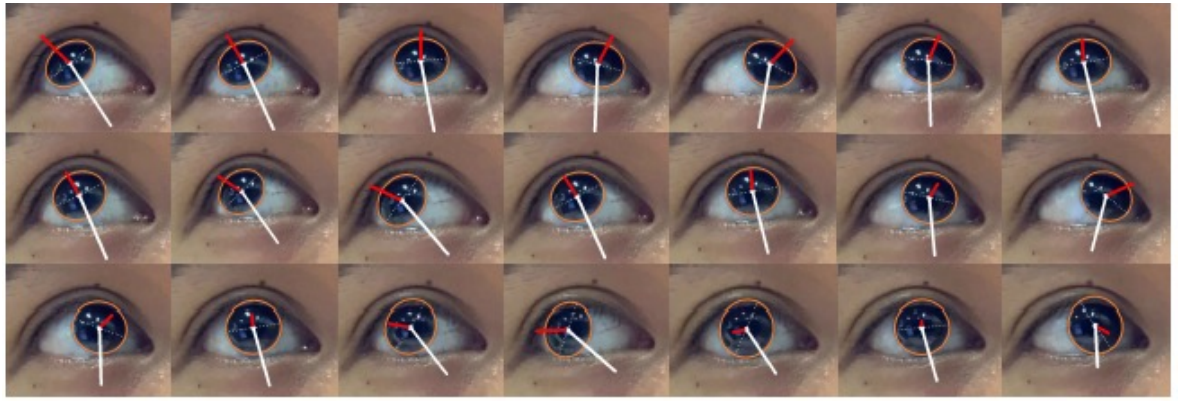
- Do feature matching
- Un-matched pixels are removed

Module -2: Gaze Ray Estimator

- 2D ellipse parameters → 3D gaze ray
  - key problem: ambiguity

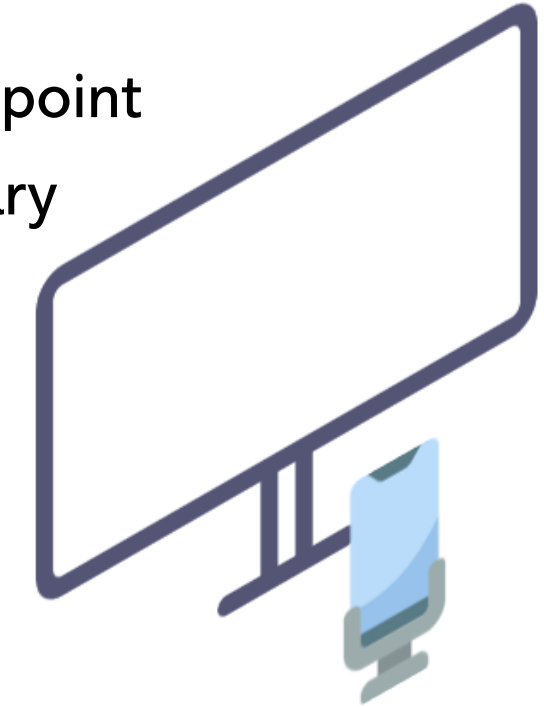


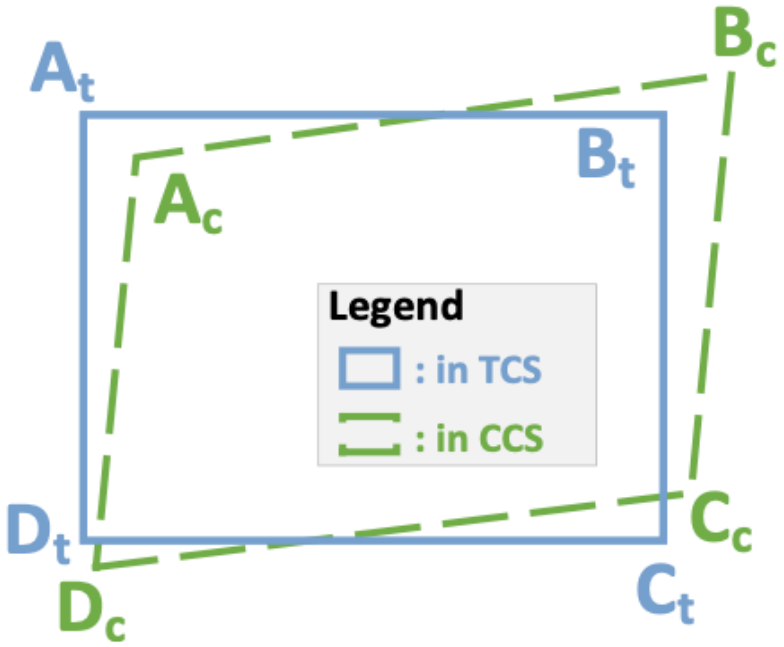
  - solution:
    - we choose the gaze direction that accumulates the least rotation change



Module -3: Mapping

- ✗ GT of gaze point
- ✓ Iris boundary

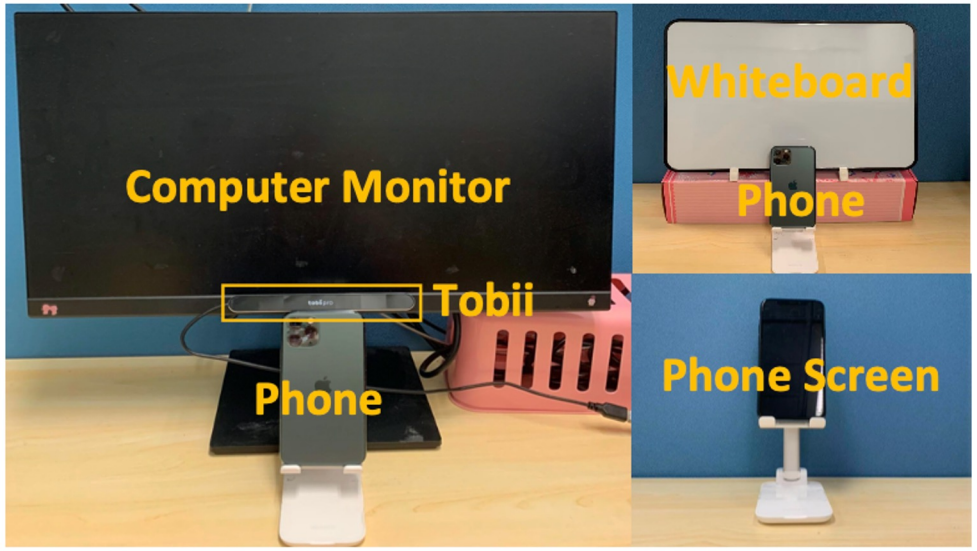




We only need user to stare at four known points.

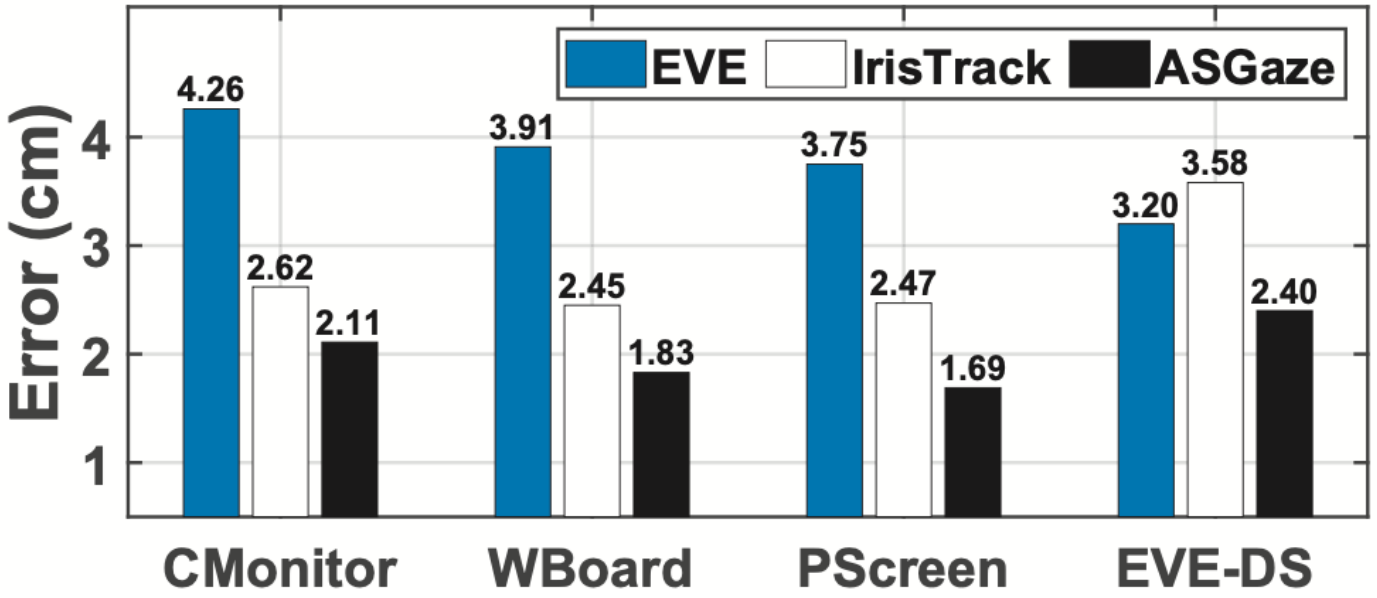
Experimental Setup

- Participants: 8 volunteers
- Tracking device:
  - RGB camera of iPhone 11 pro
- Tracking surfaces:
  - computer monitor
  - whiteboard
  - phone screen
  - public dataset



Overall Performance

- Compare with:
  - IrisTrack [1]
  - EVE [2]



EVE: 3.20~4.26 cm

IrisTrack: 2.45~3.58 cm

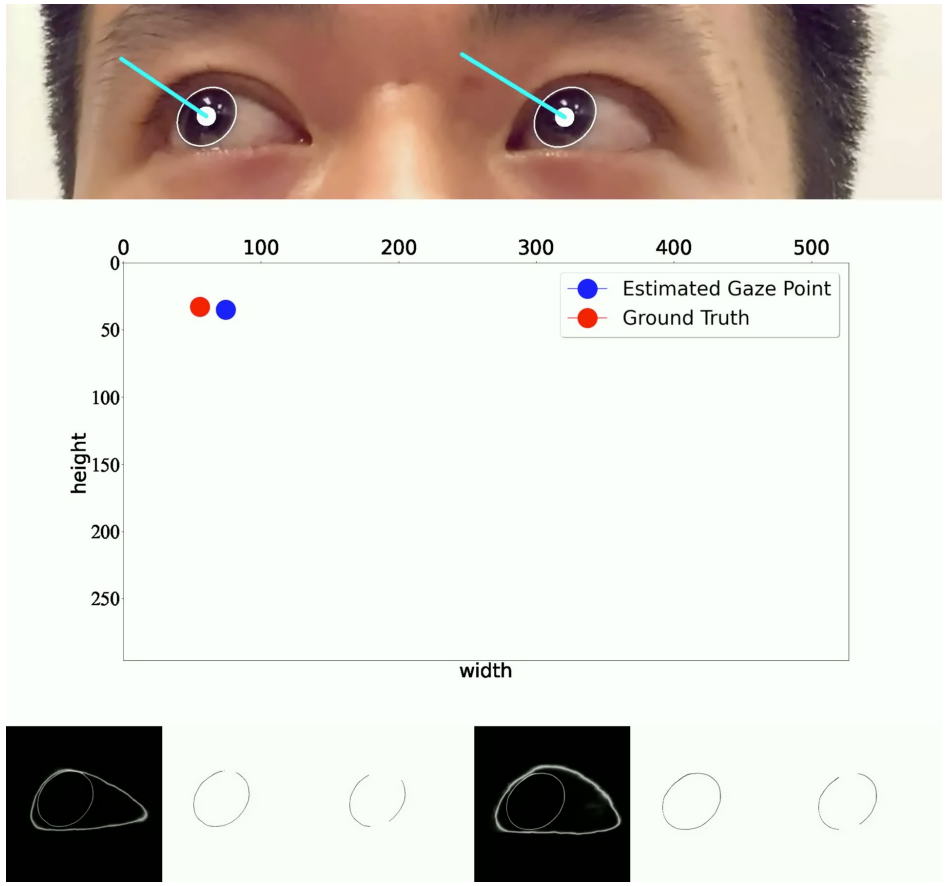
Ours: 1.69~2.40 cm

[1] C. Lin, X. Li, Z. Li, and J. Hou. Finding stars from fireworks: Improving non-cooperative iris tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022.
[2] S. Park, E. Aksan, X. Zhang, and O. Hilliges. Towards end-to-end video-based eye-tracking. In *Proc. of Springer ECCV*, 2020.

Demo

Project: https://asgaze.github.io/

Code: https://github.com/Jiani-CAO/ASGaze



Conclusion 1,2,3

1. One goal:
   - Gaze tracking using a common RGB camera

2. Two aspects:
   - Accurate tracking
   - Tracking on any surface

3. Three modules:
   - Iris boundary detector
   - Gaze ray estimator
   - Mapping

Thank you
Q&A

ASGaze:
Gaze Tracking on Any Surface with your Phone

Jiani Cao[1], Chengdong Lin[1], Yang Liu[2], Zhenjiang Li[1]

City University of Hong Kong[1], University of Cambridge[2]