ID : 674

User Independent Gaze Estimation by Extracting Pupil Parameter and Its Mapping to the Gaze Angle

Sang Yoon Han*, Nam Ik Cho

jaegal88@ispl.snu.ac.kr, nicho@snu.ac.kr*

Seoul National University, Department of Electrical and Computer Engineering, INMC

Abstract Since gaze estimation plays a crucial role in recognizing human intentions, it has been researched for a long time, and its accuracy is ever increasing. However, due to the wide variation in eye shapes and focusing abilities between the individuals, accuracies of most algorithms vary depending on each person in the test group, especially when the initial calibration is not well performed. To alleviate the user-dependency, we attempt to derive features that are general for most people and use them as the input to a deep network instead of using the images as the input. Specifically, we use the pupil shape as the core feature because it is directly related to the 3D eyeball rotation, and thus the gaze direction. While existing deep learning methods learn the gaze point by extracting various features from the image, we focus on the mapping function from the eyeball rotation to the gaze point by using the pupil shape as the input. It is shown that the accuracy of gaze point estimation also becomes robust for the uncalibrated points by following the characteristics of the mapping function. Also, our gaze network learns the gaze difference to facilitate the re-calibration process to fix the calibration-drift problem that typically occurs with glass-type or head-mount devices.

MOTIVATION

- ✓ Finding point of gaze is important for the human-computer interaction especially for the user interface in AR/VR devices.
- ✓ Existing leraning based methods need a huge amount of training data and/or complicated adaptation learning with the user's own training data, which is not easy for non-expert users.

PROPOSED METHOD

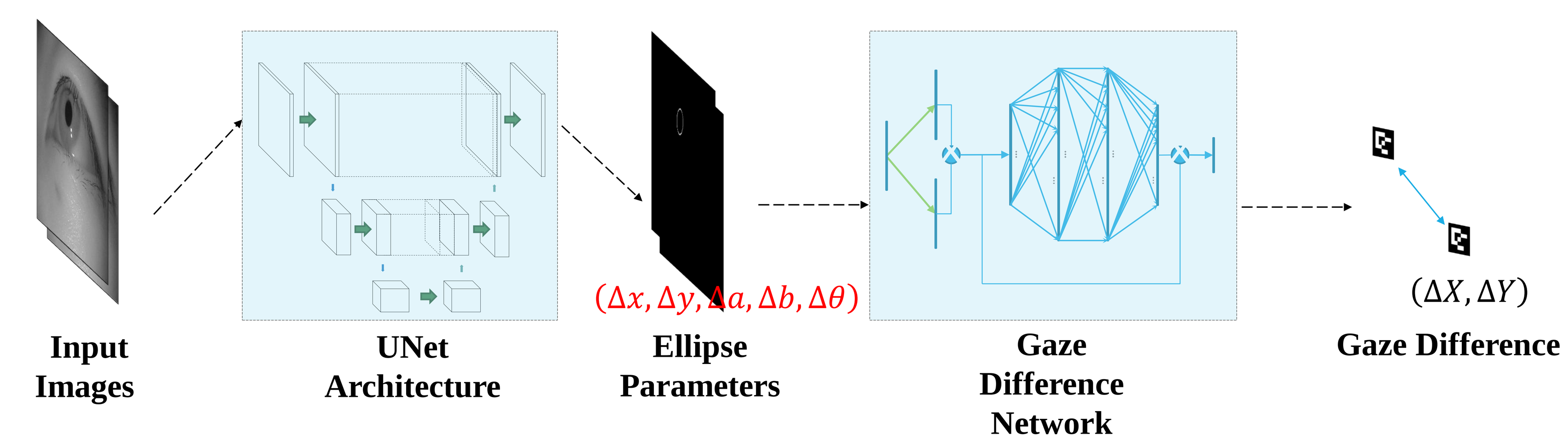
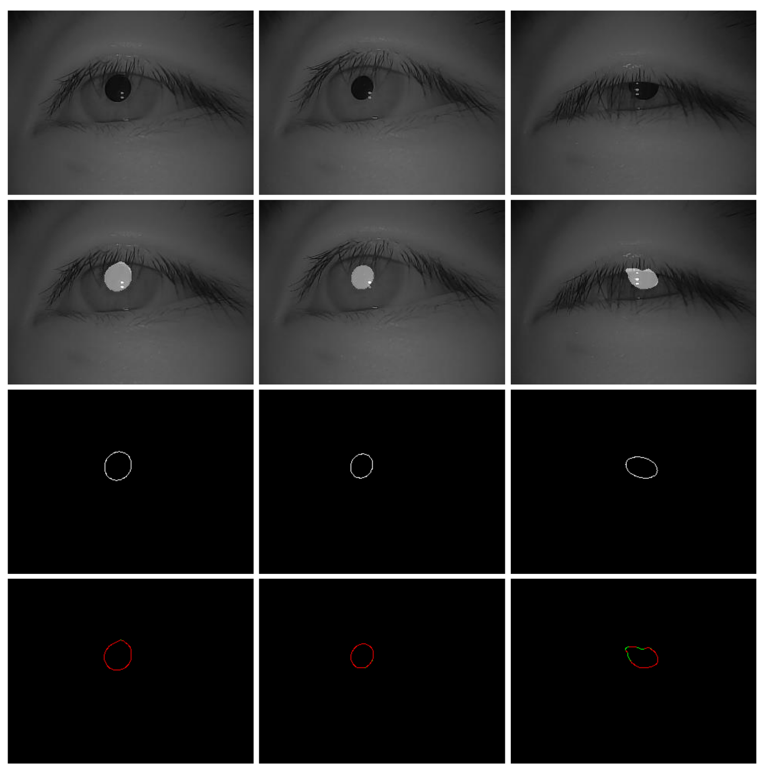


Figure 1. The workflow of ellipse fitting in our method using UNet architecture. The left part of the figure is from our preliminary work



$$\rho = \sum_{p \in \Omega_O} S_p$$

$$\phi = \frac{\sum_{p \in \Omega_{O_{edge}}} B^i}{\sum_{p \in \Omega_{O_{edge}}} 1}$$

$$\text{Confidence}: \varphi = \frac{\rho + \phi}{2} > 0.95$$

$$\text{where } B^i_{pqr} = \begin{cases} 1, & \text{if } p \in \Omega_{Ellipse} \\ 0, & \text{otherwise} \end{cases}$$

Figure 2. Pupil segmentation results and fitting ellipse to measure ellipse confidence score. (a),(b) are considered as pupil and (c) is not considered as a pupil.

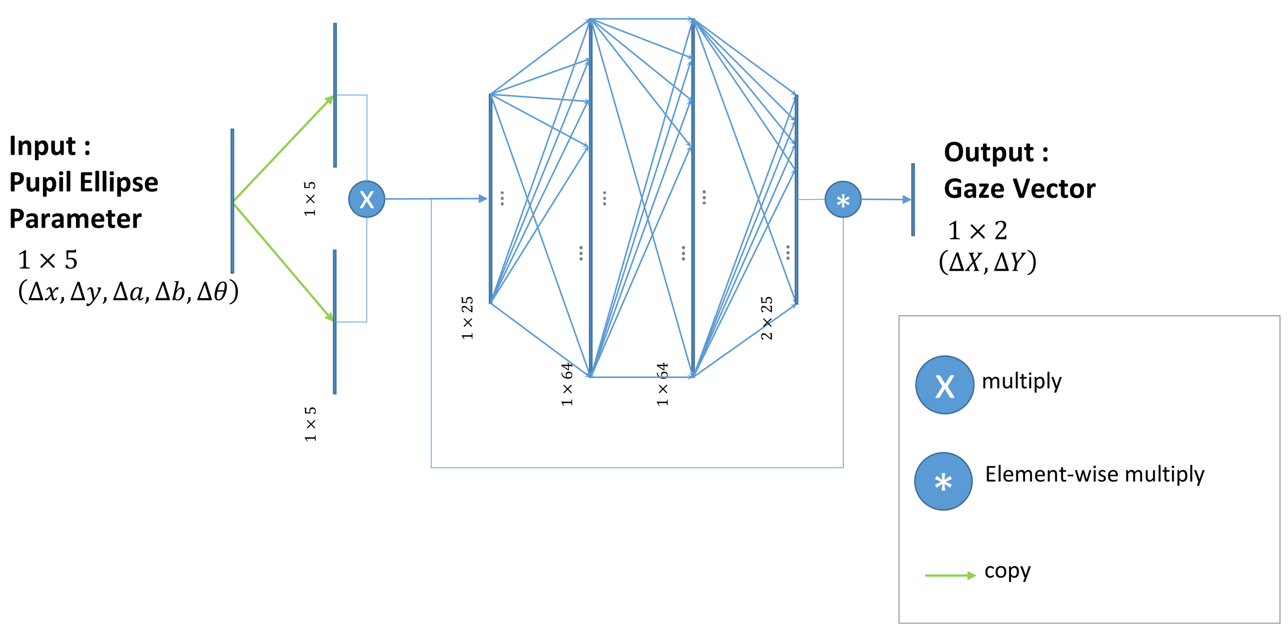


Figure 3. Proposed gaze estimation network

EVALUATION

Table 1. The angular error in gaze detection on each users using proposed algorithm (unit: °).

Gaze Point \ User	User1	User2	User3	User4	User5	User6	User7	User8	User9	User10	User11	User12	User13	User14	User15	Average
1	0.33	0.50	0.23	0.52	1.88	0.54	0.31	0.23	0.40	0.88	0.23	0.79	2.34	0.60	1.15	0.73
2	0.32	0.13	0.12	0.13	0.40	0.21	0.25	0.32	0.31	0.92	0.13	0.78	2.38	0.66	1.22	0.55
3	0.24	0.08	0.23	0.13	0.60	0.14	0.10	0.20	0.33	0.92	0.20	0.82	2.50	0.60	1.11	0.55
4	0.22	0.19	0.33	0.14	0.75	0.31	0.29	0.25	0.34	0.87	0.12	0.78	2.32	0.65	1.23	0.59
5	0.19	0.11	0.21	0.22	0.59	0.24	0.36	0.22	0.36	0.90	0.12	0.77	2.29	0.59	1.20	0.56
Average	0.26	0.20	0.23	0.23	0.85	0.29	0.26	0.24	0.35	0.90	0.16	0.79	2.37	0.62	1.18	0.59

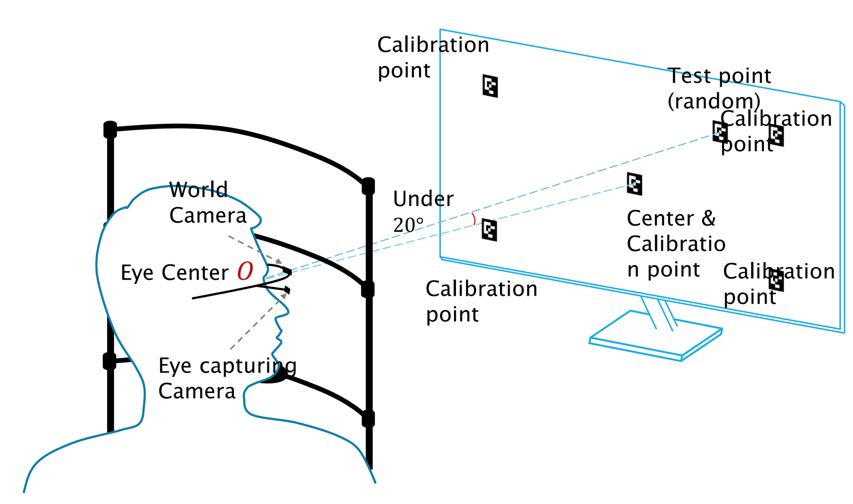


Figure 4. Illustration of capturing training data for gaze estimation network.

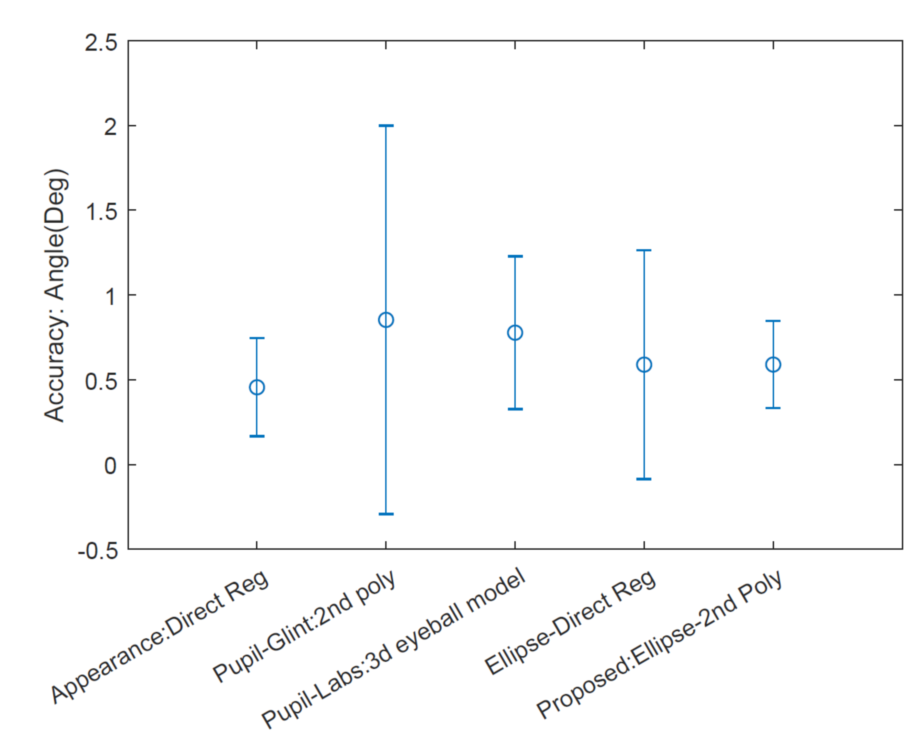


Figure 5. Accuracy comparison in terms of mean and variance of angle estimation error.

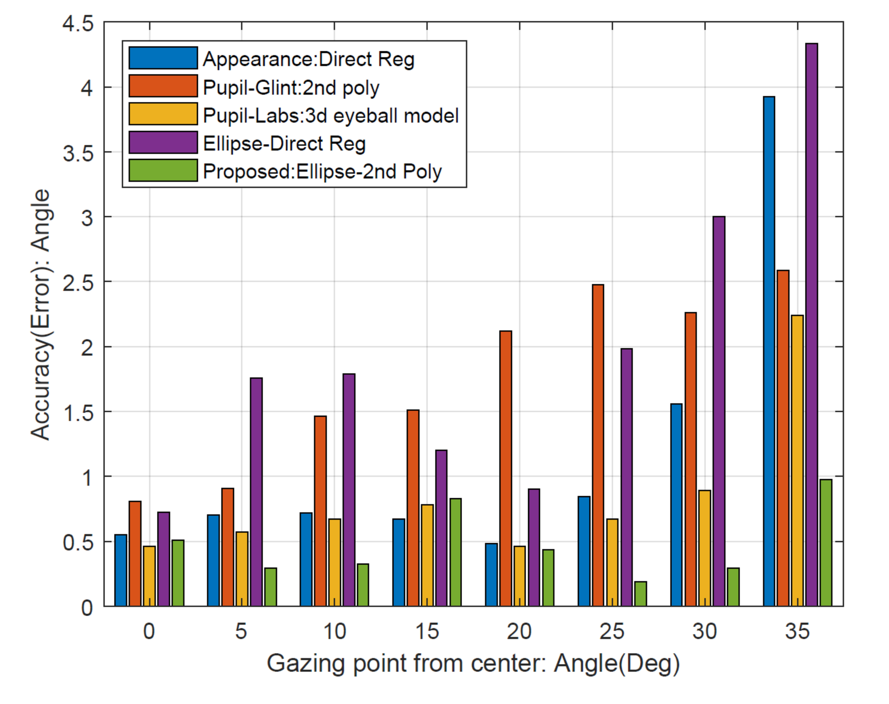


Figure 6. Accuracy comparison according to the gaze angle.

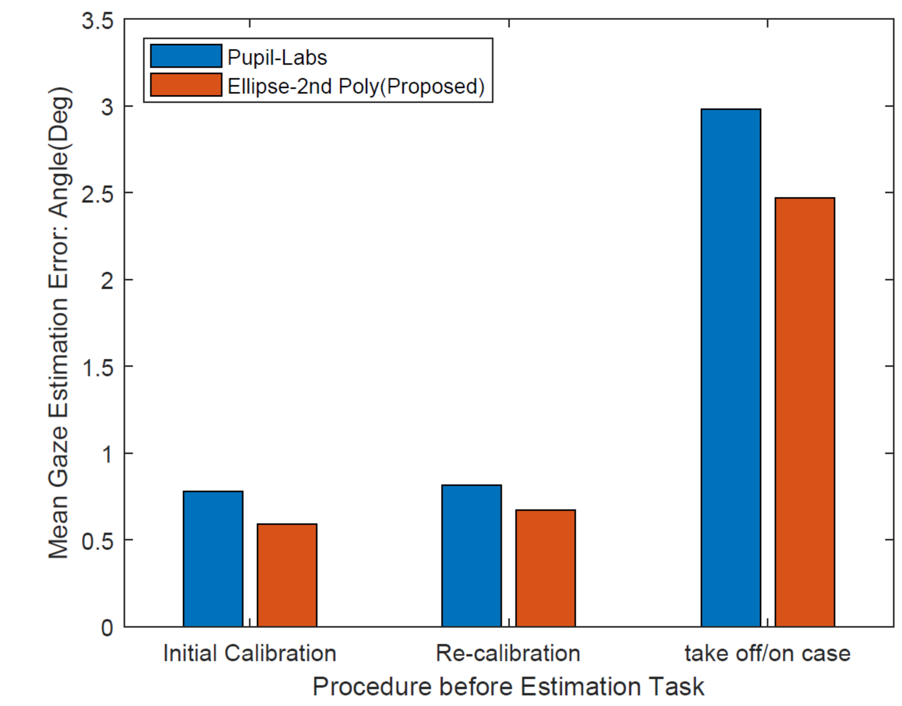


Figure 7. Mean gaze estimation error for each mode for each of three gaze pointing tasks.

CONCLUSION

- ✓ We propose a user-independent gaze estimation method by splitting the problem into two steps: extracting common features among people (elliptic parameters of pupils), and then mapping the features to the gaze angle, each of which is performed by a lightweight neural network.
- ✓ The proposed method requires very few calibration points (≤5), which is important for practicability.
- ✓ Our gaze estimation method also works robustly to uncalibrated points while other deep learning methods suffer from overfitting to the calibrated points.
- ✓ We make the network learn gaze difference so that it could easily respond to re-wearing of head-mount/glass devices or calibration drift situations.