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

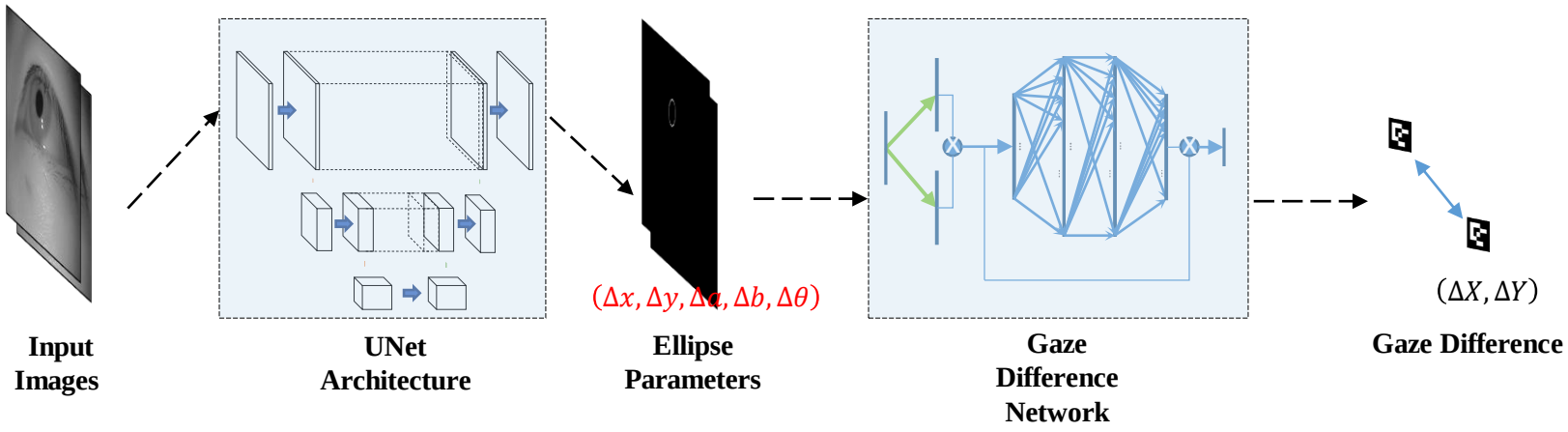
Sang Yoon Han, Nam Ik Cho

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

Overall Algorithm

❖Proposed Method

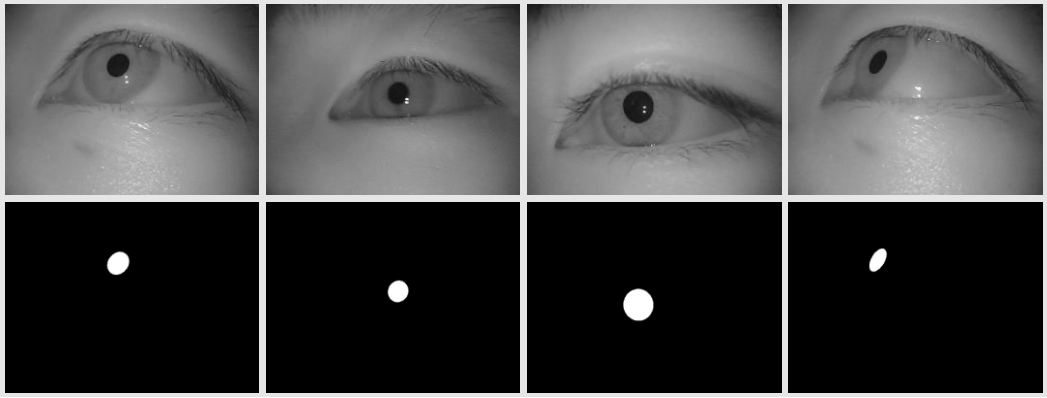
- Apply network(UNet) to frames for segment pupil candidate area.
- Confidence analysis to get frame with clear pupil area.
- Uses ellipse parameter acquired from pupil area to gaze difference network.
- Get gaze with using current frame and reference frame.



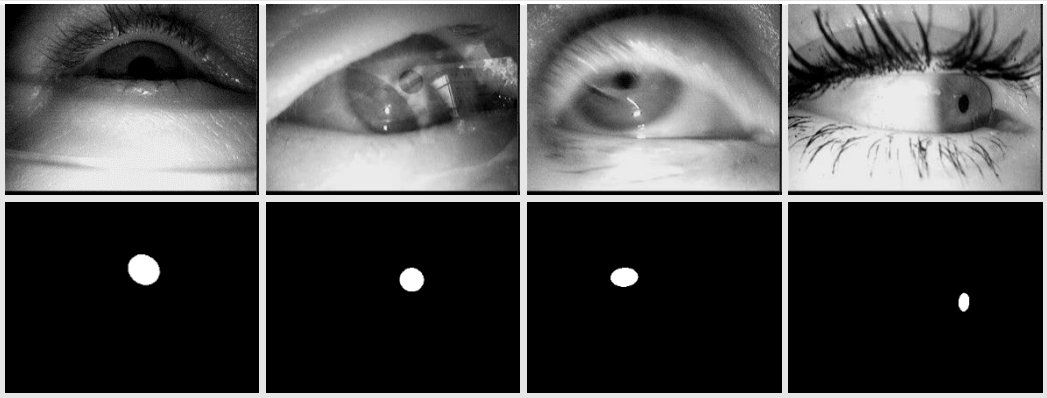
Pupil segmentation Dataset

- Unlike the existing CNN-based pupil detection methods based on the regression, we use the CNN for the **pupil segmentation.**
  - Pupil boundary edges have more feature than pupil center.
  - Traditional methods finds pupil center from fitting ellipse using pupil edges.

Own dataset (111,581 frames)



Open dataset (Else & ExCuSe, 94,113 frames)

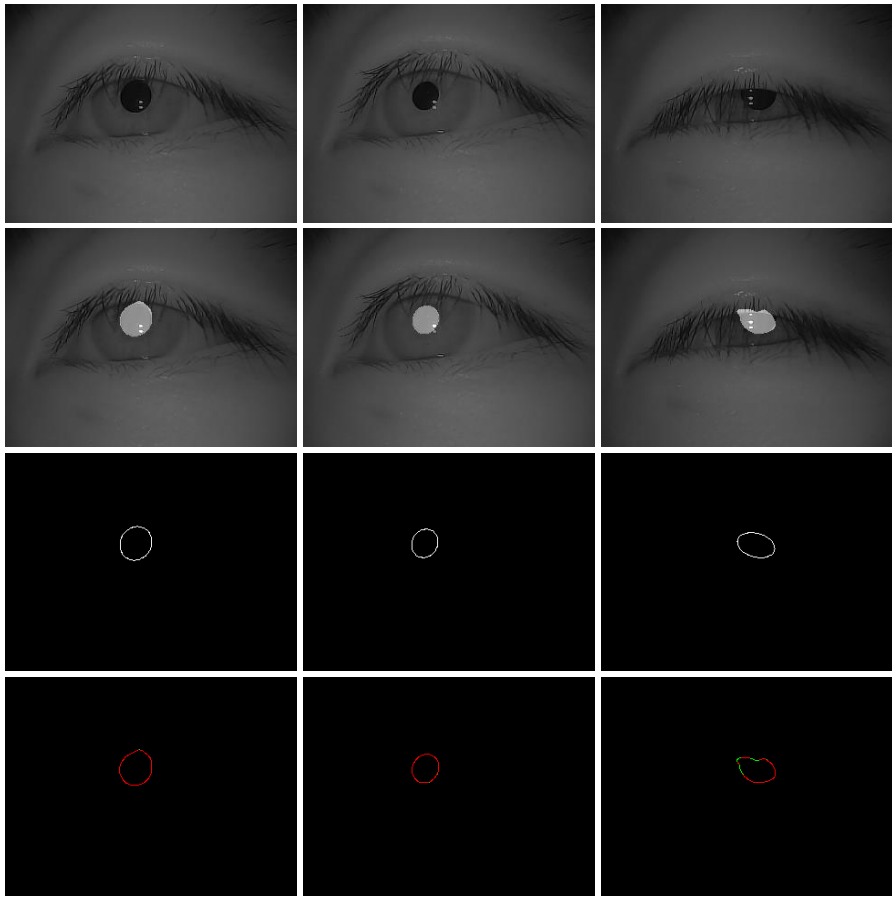


← Only have center point

Blink detection

❖Proposed Method

- Confidence analysis to get frame with clear pupil area.
  - The red pixel is the pixel belonging to the fitted ellipse
  - The green pixel is the pixel not belonging to the fitted ellipse



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

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

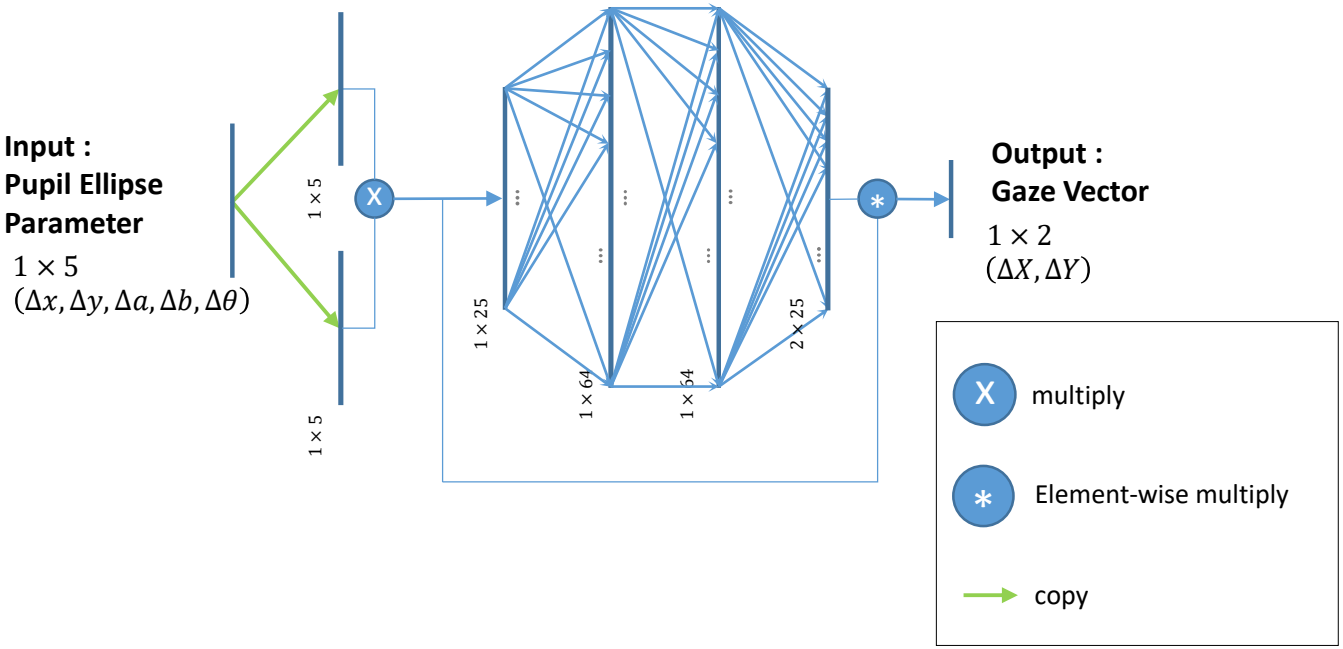
Confidence : $\varphi = \frac{\rho + \phi}{2} > 0.95$

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

Gaze Estimation Network

❖Proposed Gaze Estimation Network

- Uses ellipse parameter difference acquired from reference frame and current frame
- Multiply input parameter twice in order to use the ellipse parameter of the pupil as the second polynomial variable.
- Result of passing the fully connected layer is to do the dot product with Equ.2 again before getting to the output.

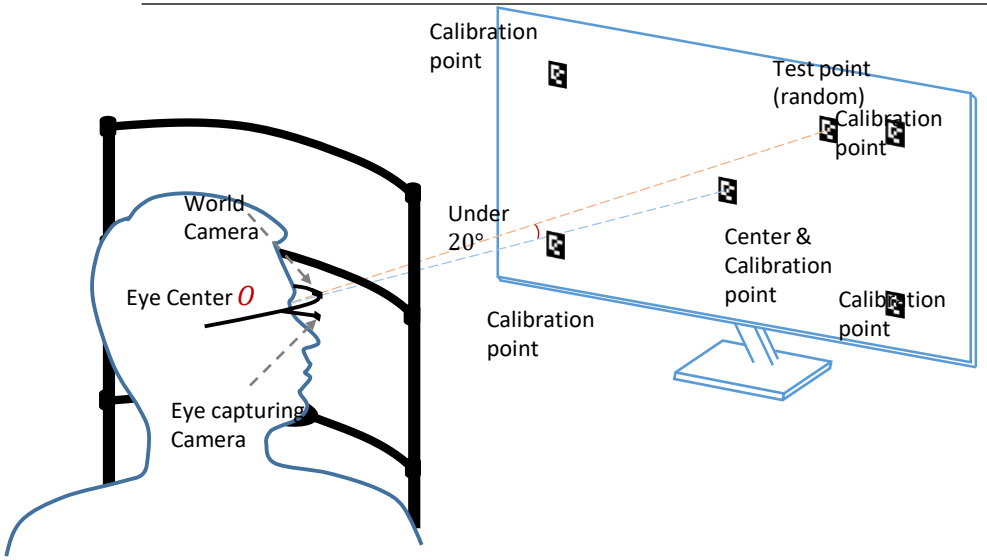


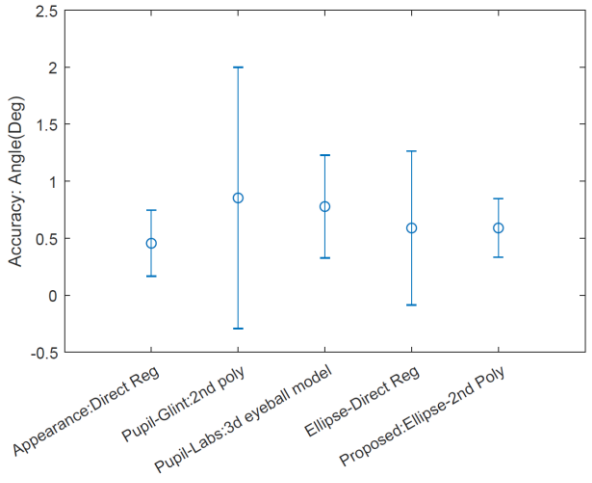
Gaze Estimation Results

❖Experiment

- The angular error in gaze detection on each users using using proposed algorithm. (unit: degree)
- Compute the error of the estimated gaze point and marker point position on the world camera while looking at the marker at each position of the monitor.
- Gazing points are composed of under 20 degree from center.

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



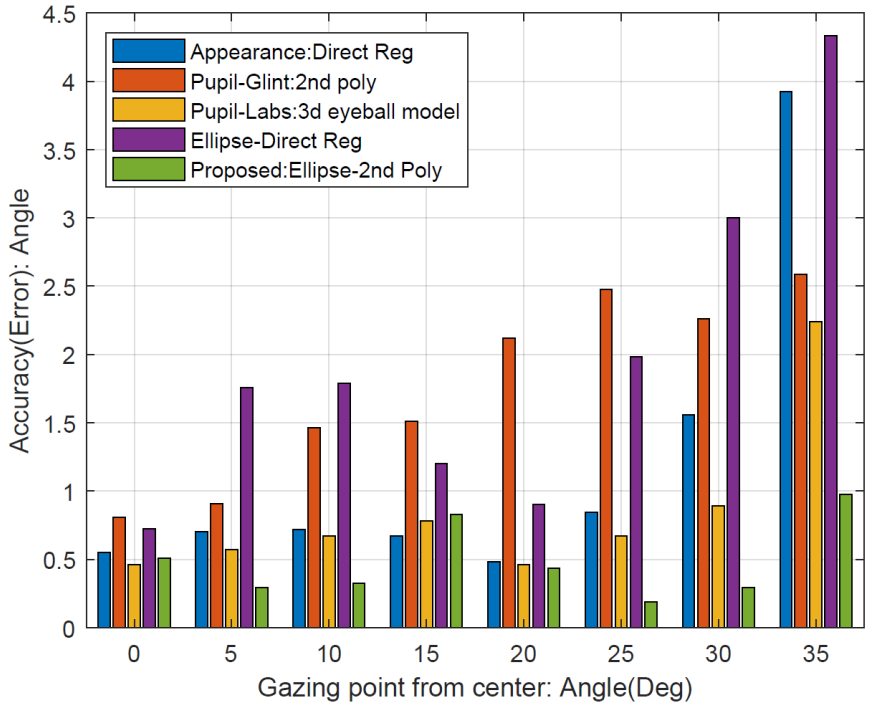


Comparison

❖Experiment

- Results outside the calibrated range show quite different accuracy.
- Accuracy of networks that directly estimate X,Y coordinate of gaze over 20deg shows bad results

	Learning based	Use polynomial	Use ellipse param	Use gaze difference	etc.
Appearance	√				Image as input
Pupil-Glint	√	√			Relative pupil position
Pupil-labs			√		3D eyeball model
Ellipse-direct	√		√	√	-
Ellipse-2nd poly	√	√	√	√	-



Comparison : Re-calibration

❖ **Experiment**

- Calibration Drift
- Recalibrating each time a calibration drift occurs makes user cumbersome
- Proposed method of using difference to minimize the calibration process when re-calibration.