Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences

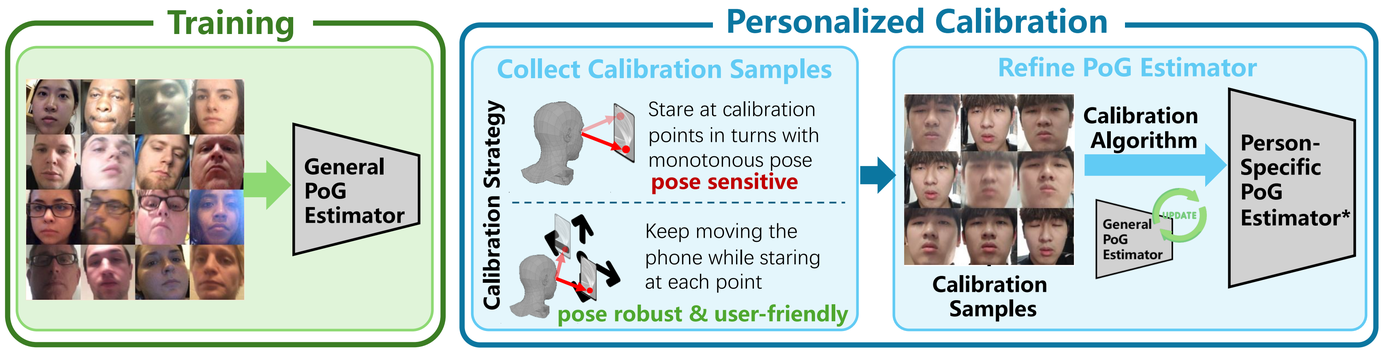
BMVC 2025

Pose-Robust Calibration Strategy for Point-of-Gaze Estimation on Mobile Phones

Yujie Zhao, Jiabei Zeng, Shiguang Shan

https://mobile-pog.github.io

Motivation



Despite recent progress in appearance-based PoG estimation, existing models still struggle to generalize across users due to individual eye differences, motivating the need for calibration.

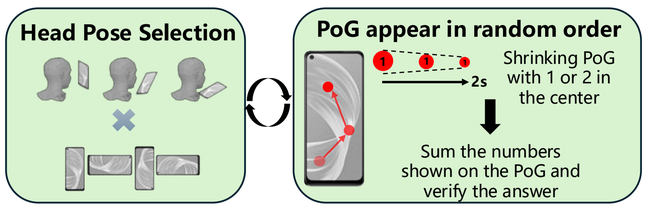
Personalized calibration works well only when the user keeps a fixed head pose, once the pose changes, accuracy drops, and current benchmarks are difficult to analyze or fix this issue.

How to ensure the model robust to pose after calibration?

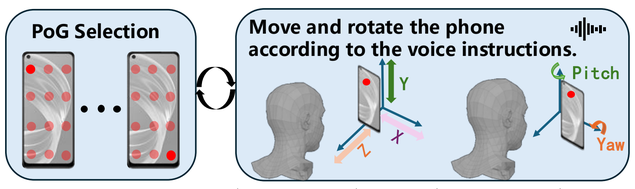
Benchmark

Since existing PoG datasets were not designed to investigate calibration factors and lack the necessary conditions for such analysis, we construct a new benchmark, MobilePoG, which contains facial images of individuals when they are fixating on designated points under either fixed or continuously changing head poses.

• *Dataset Collection*



(a) Static-MobilePoG



(b) Dynamic-MobilePoG

Each participant is instructed to adopt one of three user-to-phone relative positions and to hold the phone in one of four orientations, yielding **12 head pose configurations**.

For each pose, **55 predefined points** appeared in random order. Participants are asked to maintaining the same pose.

For each of the predefined points, participants are instructed to fixate on the target point for a period of time.

During this fixation period, they are guided by **voice instructions** to smoothly move and rotate the phone. These instructions induce head pose variations across **five degrees of freedom**: left-right (x), up-down (y), forward-backward (z), pitch, and yaw.

• *Dataset Characteristics*

Dataset	# Subject	# PoG	# Image	# Pose	Pose Controllable
TabletGaze [30]	51	35	1,785	4 Static	✓
GazeCapture [20]	1468	Random	2,445,504	Random	✗
Static-MobilePoG	32	55	606,301	12 Static	✓
Dynamic-MobilePoG	32	12	1,707,910	Continuous	✓

• *Dataset Examples*

Experiment

• *Factor Analysis of Personalized Calibration*

Model	Calibration Algorithm	1 calibration point Same	Diff.	Avg.	5 calibration points Same	Diff.	Avg.	9 calibration points Same	Diff.	Avg.	13 calibration points Same	Diff.	Avg.
iTracker	w/o calibration	2.82											
	SVR	4.17	11.17	10.60	2.30	8.93	8.42	2.21	8.47	8.02	2.27	8.31	7.92
	Linear Probe	3.82	7.08	6.81	2.40	4.13	4.00	2.17	5.30	5.07	2.29	4.58	4.43
	Finetune MLP	2.35	2.74	2.71	2.10	2.88	2.82	2.00	2.83	2.77	2.02	3.05	2.98
	Full Finetune	2.10	2.71	2.66	2.07	2.72	2.67	1.97	2.63	2.58	1.40	2.65	2.57
AFFNet	w/o calibration	2.54											
	SVR	4.18	11.17	10.60	2.18	8.40	7.93	1.69	8.91	8.40	2.03	8.60	8.17
	Linear Probe	4.18	6.84	6.62	2.02	3.29	3.19	1.97	3.69	3.57	2.05	3.55	3.45
	Finetune MLP	2.06	2.53	2.49	2.04	2.41	2.38	1.52	2.45	2.38	1.48	2.57	2.50
	Full Finetune	1.94	2.37	2.33	1.99	2.39	2.36	1.42	2.45	2.38	1.40	2.37	2.30

The table above reports the performance of calibrated estimators based on samples collected under a single head pose. **The results clearly indicate that calibration based on monotonous head pose samples fails to generalize to varying head poses.** Thus, incorporating pose diversity in the calibration data is essential for achieving robust and reliable PoG estimation in real world.

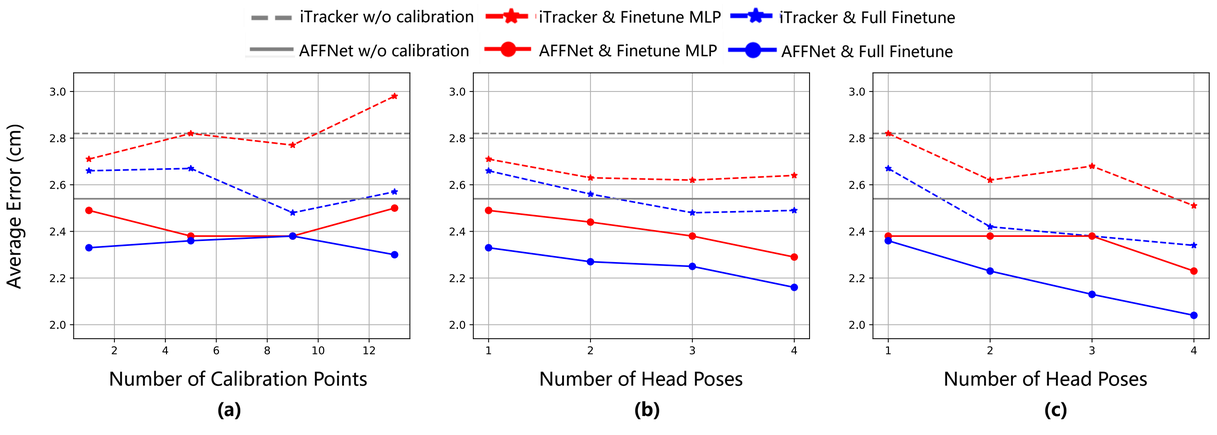


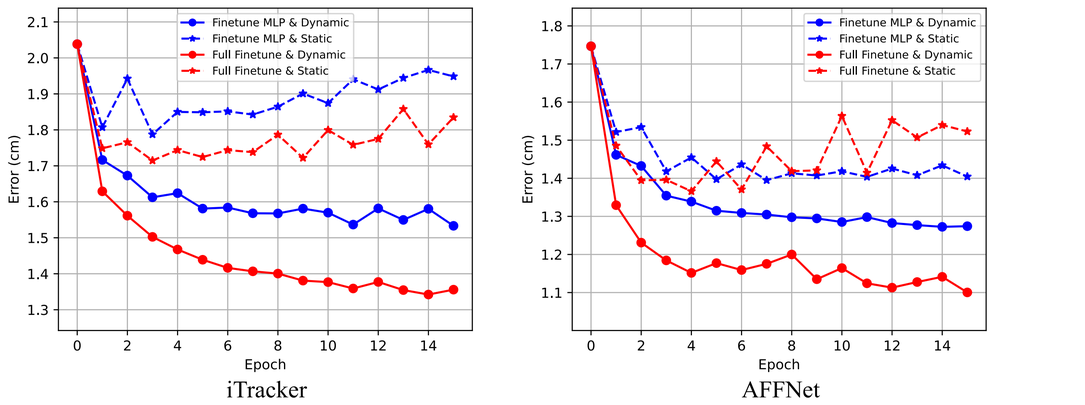
Figure 4 illustrates how the average error of the calibrated estimator changes with increasing numbers of calibration points and head poses in the calibration samples. **These findings strongly suggest that head pose diversity is a critical factor in personalized calibration.** Incorporating varied head poses into calibration samples substantially enhances the generalization ability of calibrated estimators and should be prioritized when designing practical calibration strategies.

• *Dynamic Strategy VS. Static Strategy*

As shown in the collection procedure of Dynamic-MobilePoG, we propose a new dynamic calibration strategy. **For each calibration PoG, the user just needs to rotate and translate the mobile phone within a range while keeping gaze at the PoG**, which not only allows for the collection of calibration samples with sufficiently diverse head poses, but is also user-friendly and easy to operate on mobile devices.

Model	Calibration Algorithm	1 calib. point static	dynamic	2 calib. points static	dynamic	4 calib. points static	dynamic	6 calib. points static	dynamic
iTracker	w/o calibration	2.09							
	SVR	4.18	**4.18**	3.35	**3.11**	2.26	**1.89**	2.17	**1.76**
	Linear Probe	2.86	**2.83**	2.62	**2.58**	2.03	**1.89**	1.97	**1.77**
	Finetune MLP	2.55	**1.81**	2.24	**1.89**	2.02	**1.63**	1.95	**1.53**
	Full Finetune	1.86	**1.85**	1.87	**1.74**	1.95	**1.48**	1.83	**1.36**
AFFNet	w/o calibration	1.75							
	SVR	4.18	**4.18**	3.19	**2.9**	1.60	**1.39**	1.60	**1.35**
	Linear Probe	2.77	**2.67**	2.35	2.41	1.69	**1.66**	1.58	**1.51**
	Finetune MLP	1.79	**1.57**	1.83	**1.53**	1.52	**1.33**	1.40	**1.27**
	Full Finetune	1.49	**1.43**	1.56	**1.42**	1.83	**1.27**	1.52	**1.10**

☐ Calibration Performance: Enhance robustness to pose variation.



☐ Calibration Stability: Converge to the optimal and remain stable.