ETRA 2025

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

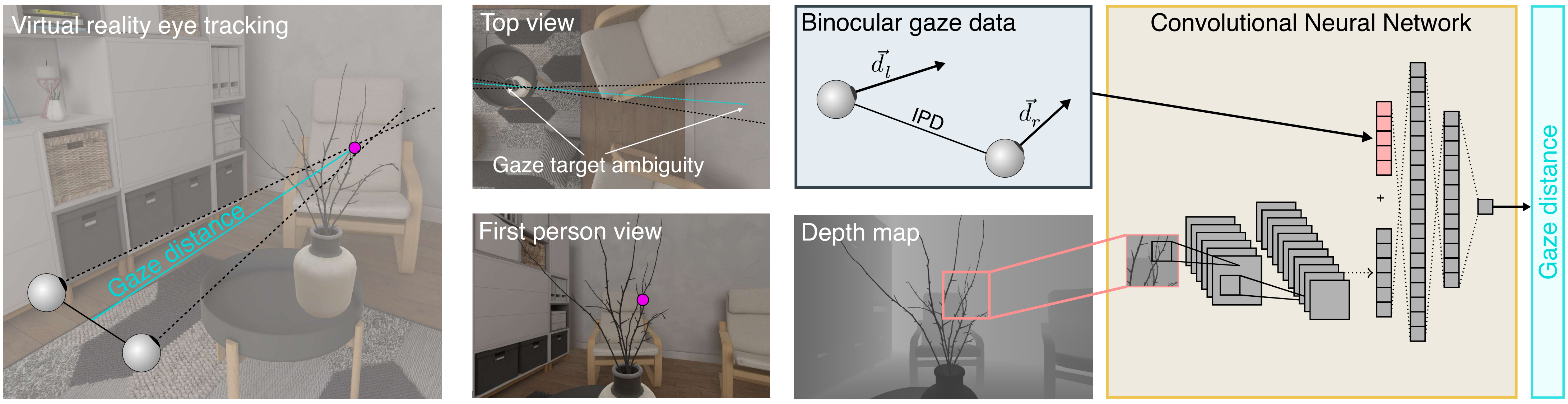
Universitätsklinikum Tübingen

ZEISS

CNN-based estimation of gaze distance in virtual reality using eye tracking and depth data

Anna-Lena von Behren[1], Yannick Sauer[1,2], Björn R. Severitt[1], Siegfried Wahl[1,2]

[1] Institute for Ophthalmic Research, Eberhard-Karls-University of Tübingen, Elfriede-Aulhorn-Strasse 7, Tübingen, Germany | [2] Carl Zeiss Vision International GmbH, Turnstrasse 27, Aalen, Germany



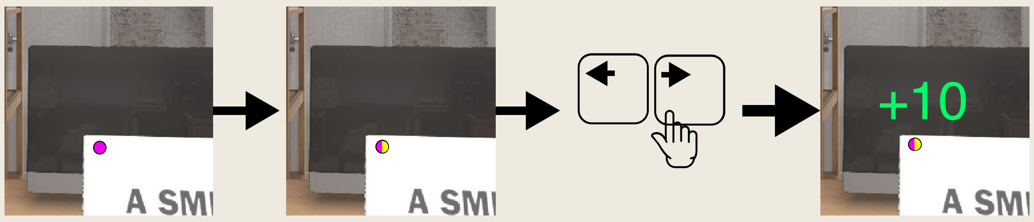
Background

- Depth of field (blur) simulation in VR [1] needs *gaze distance (d)*, not just gaze direction
- **Vergence** (eye angle) gives a rough estimate, but is often inaccurate [2]
- Using **depth at center** of gaze point works, but fails with small targets or close to object edges
- Combining vergence and depth stats (mean, std) in a **SVR** model improves estimation [3]
- Our **CNN** approach uses the full depth map around gaze point for more robust results
- Depth of field blur scales with repciprocal distance (1/*d* in diopters), we train a second model using 1/*d* as target

Dataset

VR fixation game

For collecting ground truth data: participants fixated targets placed on the surface of objects and respond to color changes for points



Recorded data during 1 s of fixation:
- Binocular gaze data (gaze direction and interpupillary distance)
- Depth map cropped around gaze location
- Ground truth distance of fixation target
- 110,996 samples from 41 participants and 280 target locations

Example scenes with example target locations

References

[1] Kramida, G. 2015. Resolving the vergence-accommodation conflict in head-mounted displays. IEEE transactions on visualization and computer graphics, 22(7), 1912-1931.1931.
[2] Wang, R. I., Pelfrey, B., Duchowski, A. T., & House, D. H. 2012. Online gaze disparity via bioncular eye tracking on stereoscopic displays. In 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission (pp. 184-191). IEEE.
[3] Weier, M., Roth, T., Hinkenjann, A., & Slusallek, P. 2018. Predicting the gaze depth in head-mounted displays using multiple feature regression. In Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications (pp. 1-9).

Results and Discussion

Comparison of CNN with baseline models

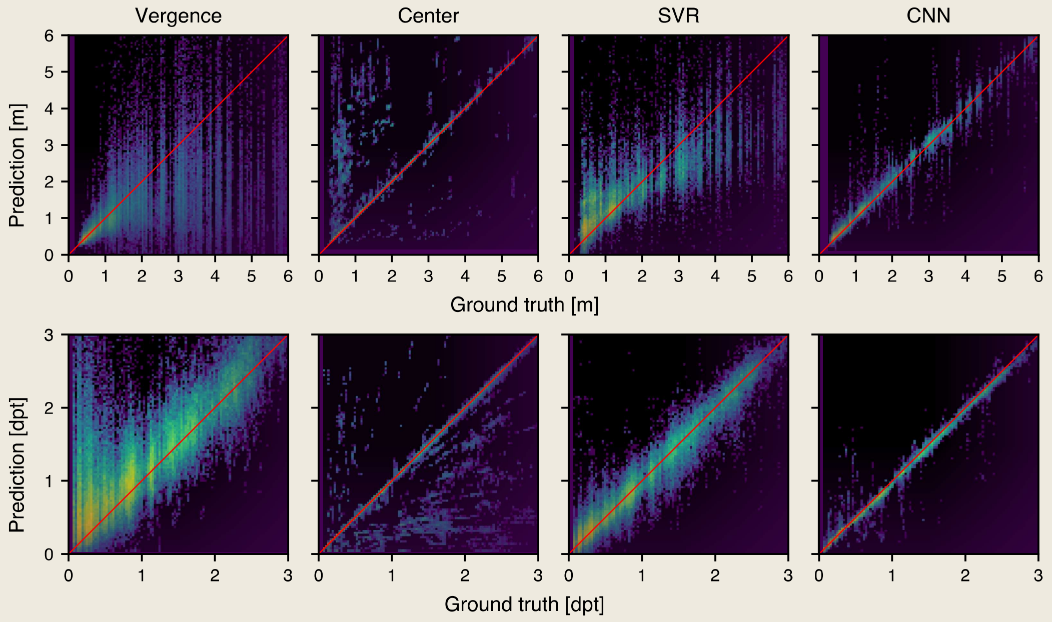
Distance models (meters)

Model	MSE	RMSE	MAE	ME	R^2
Vergence	7.62	2.76	1.35	-1.01	-1.85
Center	7.09	2.66	0.80	0.69	0.47
SVR	0.92	0.96	0.62	**0.01**	0.24
CNN	**0.18**	**0.42**	**0.17**	-0.02	**0.97**

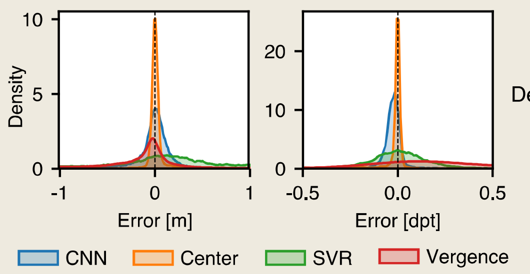
Reciprocal dist. models (diopters)

Model	MSE	RMSE	MAE	ME	R^2
Vergence	27.89	5.28	0.67	0.57	-0.02
Center	0.22	0.47	0.19	-0.15	0.58
SVR	0.04	0.21	0.13	**0.00**	0.92
CNN	**0.01**	**0.09**	**0.04**	-0.02	**0.99**

- **Vergence**-based estimates are inaccurate, especially for far distances
- **Center**-based method estimates mostly correct distance, but many outliers for small targets
- **SVR** model reduces outliers but has increased variance
- **CNN** model has best performance, wider error distribution than center but substantially fewer outliers



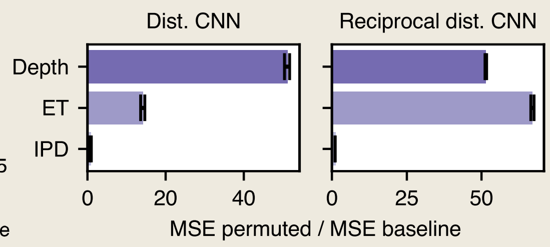
Error distribution



Our CNN has a narrower error distribution compared to vergence and SVR [2]. Compared to the center method, our approach reduces outliers, proving that binocular gaze data provides relevant information to help interpret the depth data.

Feature importance analysis

Random permutation of input features



Both depth and eye tracking data contribute to estimation: successful data fusion by the model. For reciprocal distance estimation, influence of gaze data increases. Interpupillary distance (IPD) appears to have minimal relevance.

yannick.sauer@uni-tuebingen.de

www.eye-tuebingen.de/zeiss-vision-science-lab

This research is supported by the German Research Foundation (DFG): SFB 1233, Robust Vision: Inference Principles and Neural Mechanisms, TP TRA, project No. 276693517 and by the European Union's Horizon Europe research and innovation programme under the Marie Skodowska–Curie grant PLACES: PLAsticity of perception in real and virtual spaCES, agreement No. 101086206.