

OST

Ostschweizer
Fachhochschule

Webcam basiertes Gaze-Tracking

Gaze-Tracking für die Interaktion mit dem Computerbildschirm

Dr. Lucas Falch

14. Juni 2024

EMS Institut für Entwicklung Mechatronische Systeme

Motivation

Warum Webcam basiertes Gaze-Tracking?

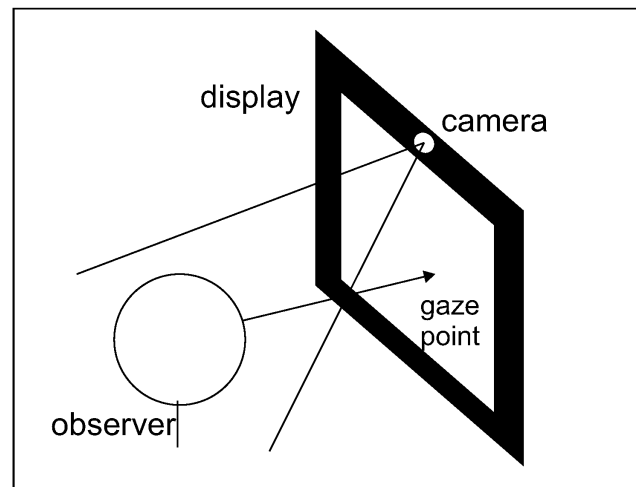
- Warum Webcam?
 - Fast jeder Arbeitsplatz verfügt über eine Webcam
 - Konventionelle Methoden benötigen spezifischen Eye Tracker (Infrarot und Stereo Kameras)
 - Webcam ist ein kostengünstiger Sensor
- Warum Gaze-Tracking?
 - Untersuchung von Benutzerfreundlichkeit
 - Alternative Methode zur Interaktion
 - Effizienzsteigerung



Forschungsfragen und Problembeschreibung

Wie kann mittels 2D Kamera der Blick auf dem Bildschirm geschätzt werden?

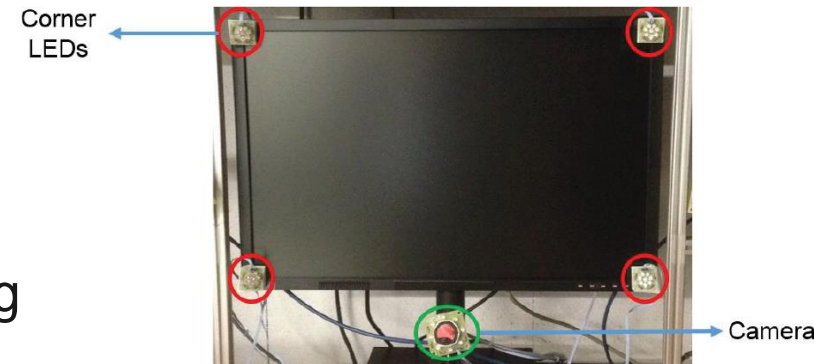
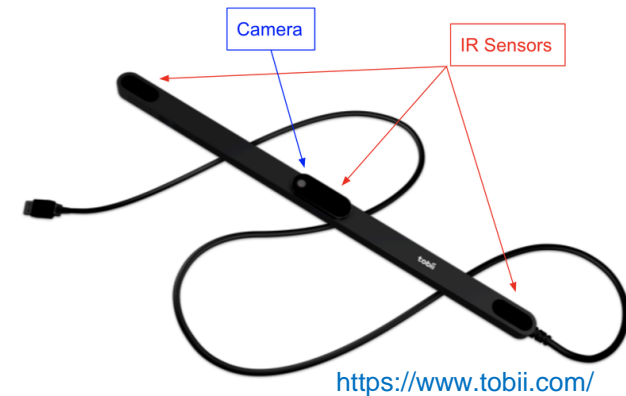
- Augen bzw. Pupille muss getrackt werden
- Blickrichtung muss auf den Bildschirm übertragen werden
- Challenges:
 - Abstand zwischen Benutzer und Bildschirm unbekannt
 - Kopfbewegungen und Rotationen erschweren Gaze-Tracking



Stand der Technik

Webcam Eye Tracking

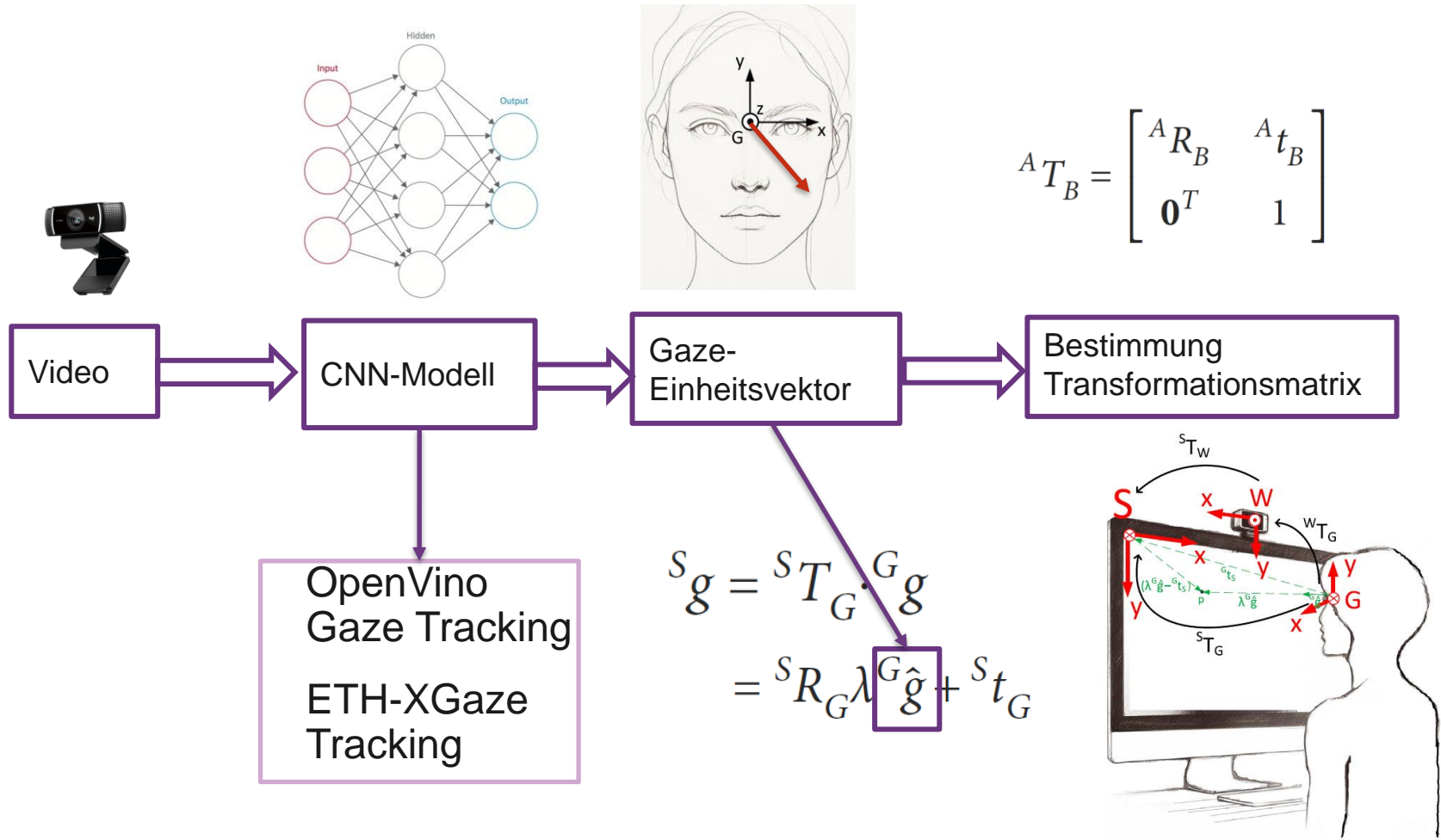
- Konventionelle Eye Tracker (wie Tobii Eye Tracker) verwenden Infrarot Sensoren
- Infrarot Licht Reflexionen, um Auge zu tracken und Blick auf Bildschirm zu projizieren
- Trainierte CNN-Modelle, um Blickrichtung zu schätzen
 - Blickrichtung aus Sicht des Users
 - Abstand von Bildschirm zu Benutzer muss bekannt sein



- Yoo and Chung, 2005; Zhu and Ji, 2005; Hennessey et al., 2006;
- Zhang et al., 2017; Zhang et al., 2020; NVIDIA, 2023; OpenVino. 2023

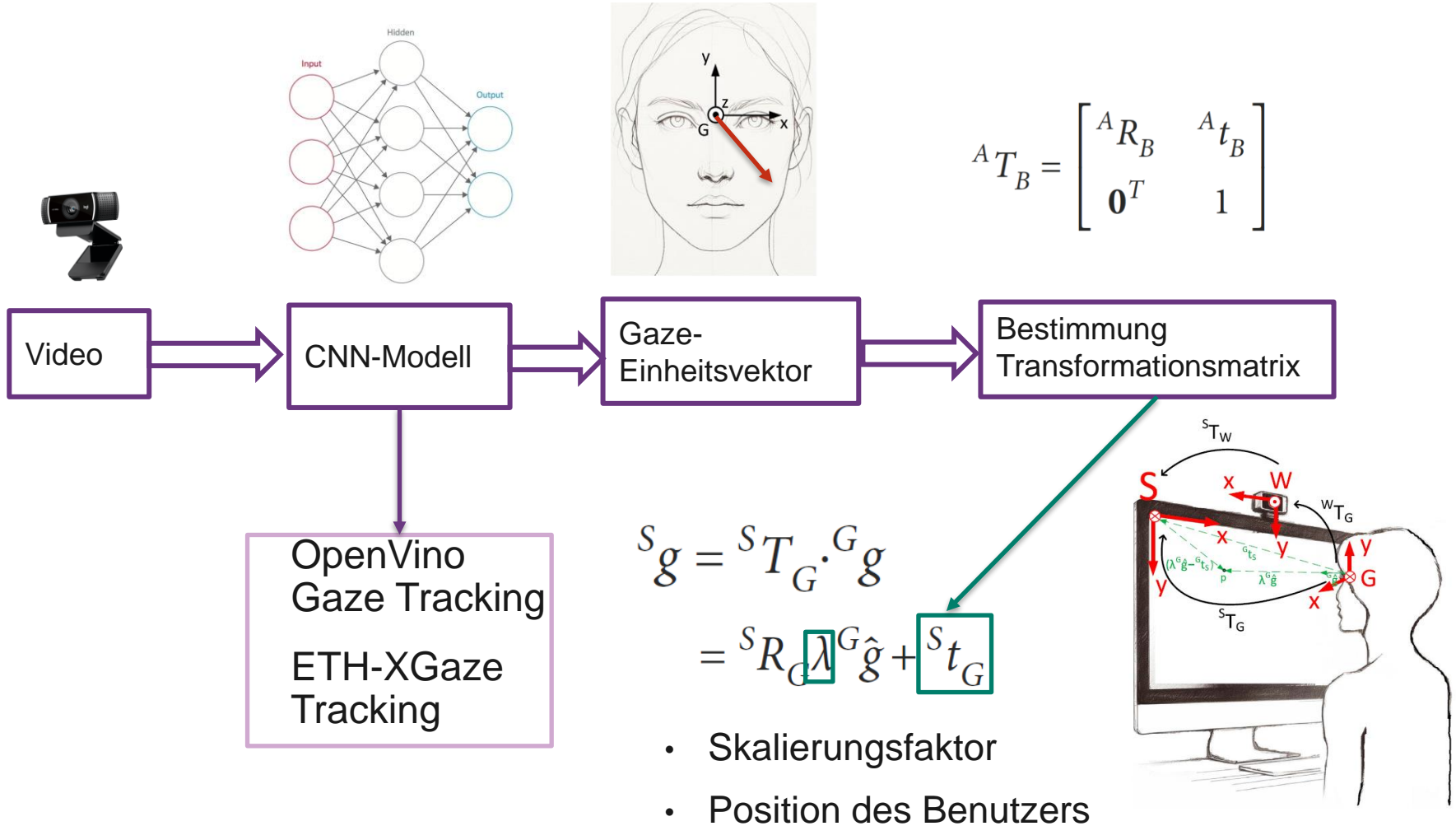
Methoden und Umsetzung

Erfassung des Blickrichtungsvektors



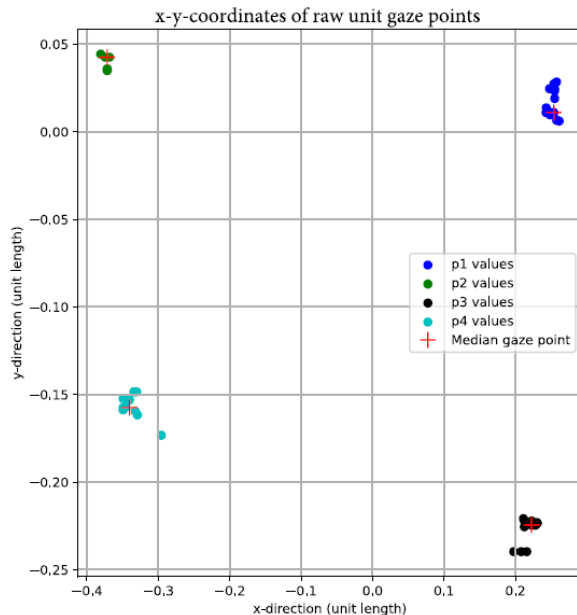
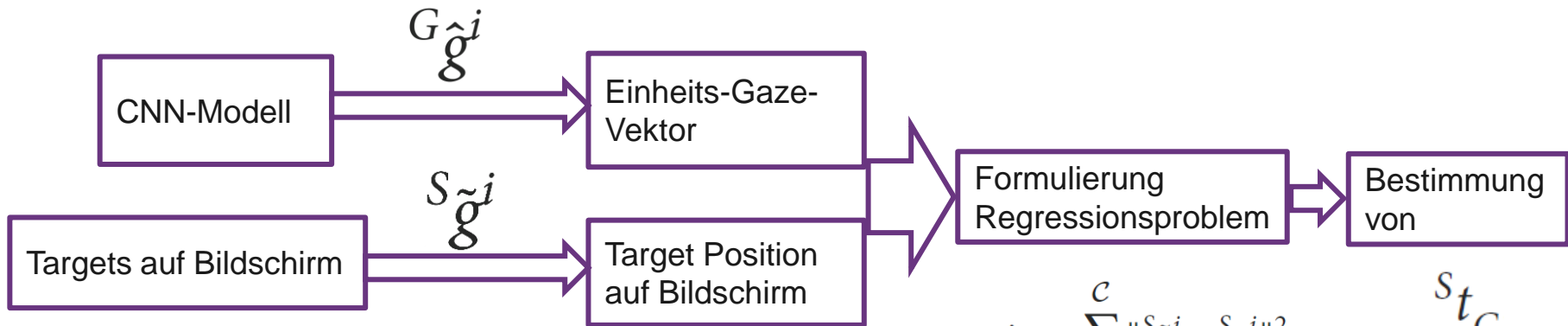
Methoden und Umsetzung

Erfassung des Blickrichtungsvektors



Methoden und Umsetzung

Bestimmung der Position mit Hilfe von Kalibrierung



$$\min_x \sum_i^c \|S_{\tilde{g}}^i - S_{g^i}\|_2^2$$

$$\text{with } S_{g^i} = {}^S R_G \lambda^{iG} \hat{g}^i + S_{t_G}$$

$$\lambda^i = \frac{{}^G Z^T \cdot (-{}^S R_G^T S_{t_G})}{{}^G Z^T G_{\hat{g}}^i}$$

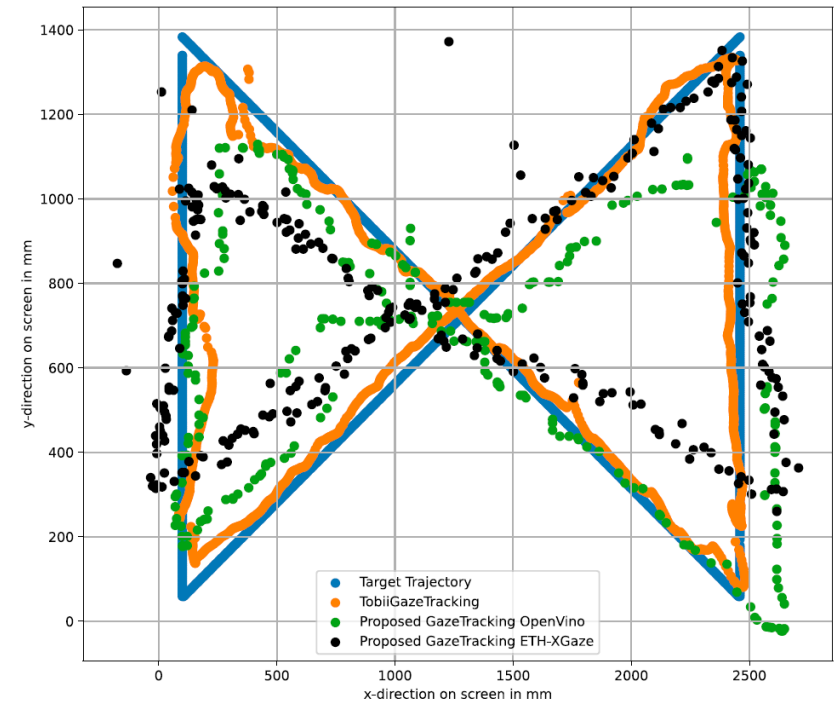
Experimente und Ergebnisse

Vergleich zu anderen Blickrichtungstools

- RMSE

	Tobii eye tracker 5	Proposed method	Proposed method
		OpenVino	ETX-Gaze
RMSE	15mm/0.9°	53mm/3.3°	50mm/3.2°

- Physikalische Messung und Schätzung durch Methode stimmt überein (Abstand: 800mm)
- Methode ist abhängig von der Genauigkeit des Einheitsvektors das vom CNN-Modell geliefert wird



Zusätzliche Informationen

Kopfbewegungen

- Über Structure from Motion (SfM) können 3D Informationen von zwei aufeinanderfolgenden 2D Bildern extrahiert werden
- Es kann eine Transformation bestimmt werden, wie sich eine Kamera im Raum bewegt (Visual Odometry)
- In dieser Anwendung bewegt sich das Bild (Benutzer) und es kann getrackt werden, wie sich der Benutzer bewegt
- Schwierigkeit:
 - Unterscheidung zwischen Rotation des Kopfes oder lateralen Bewegung des Kopfes von links nach recht.
 - Training eines neuen CNN wo Bewegungen mit berücksichtigt werden

Schlussfolgerung und mögliche Erweiterung

Webcam basiertes Gaze-Tracking

- Methode, um über eine 2D Webcam die physikalische Distanz zwischen Computerbildschirm und Benutzer zu bestimmen
- Günstige Alternative gegenüber High-End Eye Trackern
- Webcam basiertes Gaze-Tracking geeignet für grosse Benutzergruppen Studien
- Erweiterungen:
 - Spezialisiertes Training eines CNNs wo Informationen von Bewegungen des Benutzers (SfM) mit einbezieht
 - Anwendung der Lösung auf zum Beispiel Webseiten Benutzerfreundlichkeit

Lucas Falch and Katrin Solveig Lohan, "Webcam-based gaze estimation for computer screen interaction",
Frontiers in Robotics and AI, Volume 11 - 2024 | <https://doi.org/10.3389/frobt.2024.1369566>