Seminar/Project: Computer Vision and Deep Learning

P3: Eye-Tracker Extrinsic Calibration

Ayush Chamoli
chamoli@rptu.de
Supervisor: Mahdi Chamseddine

RPTU | augmented VISION

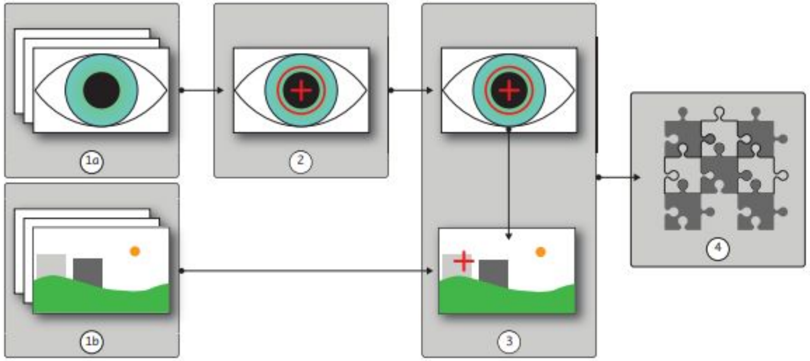
A look into the near future…

- Looking is to AR/VR, what fingers were to smartphones, or mouse was to PC.
- The “gaze” is an important source of information.
- Gaze can also be important outside the scene.
- How can one extract this information?

https://www.apple.com/de/newsroom/2023/06/introducing-apple-vision-pro/

Pupil Core

- Has one Scene Camera and two Infrared Eye Cameras.
- Detects Gaze and projects it onto the Scene Camera.
- Has an API and a GUI Application (Pupil Capture).
- API uses “dark pupil” detection method in order to detect pupil and project gaze.



https://arxiv.org/pdf/1405.0006.pdf

World Camera

- IDS 3080CP is being used as the World Camera.
- Comes with an API for accessing the functionality of Camera (Feed Buffer etc.) and changing parameters.
- Also comes with a GUI Application for tweaking settings.

Intrinsic Calibration with Chessboard

- Because of the radial distortion of both the camera, Internal Calibration is required.
- Need to calculate the Camera Calibration Matrix.

$$camera\ matrix = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

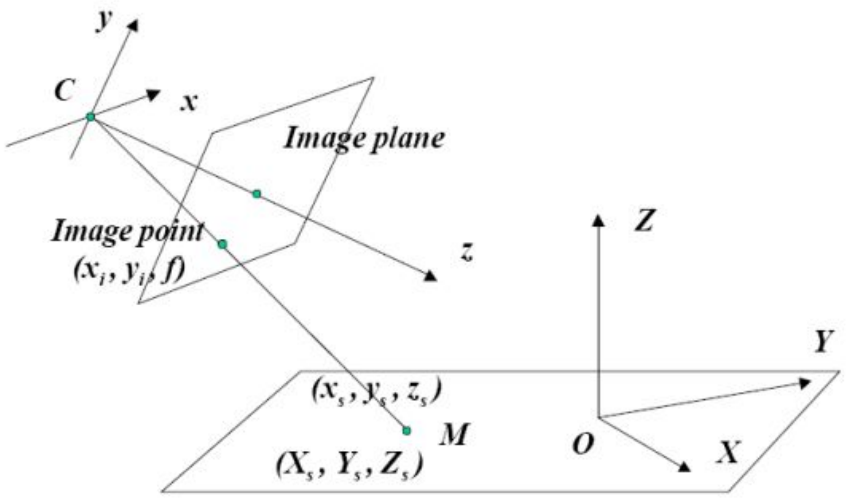
Extrinsic Calibration

- Determining the Extrinsic Parameters with respect to the control field.
- Transformation from world coordinate system to camera coordinate system:

$$\begin{pmatrix} R & t \\ 0_3^\top & 1 \end{pmatrix}$$

R - Rotation Matrix

t - Translation

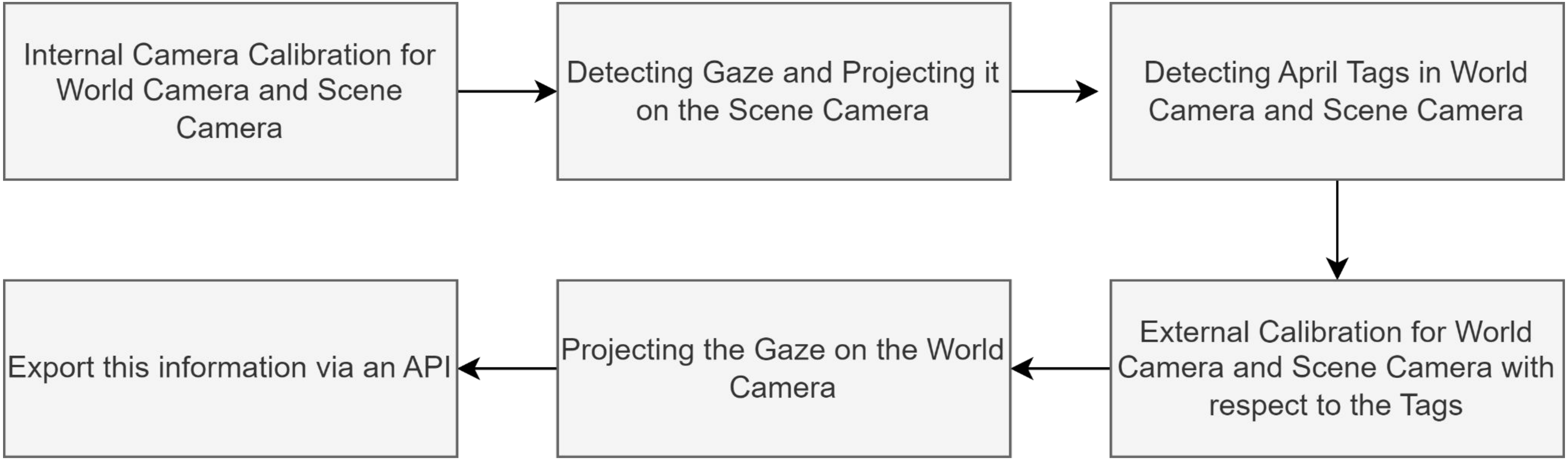


April Tags

- AprilTag is a visual fiducial system.
- It's involved in a wide variety of augmented reality, robotics and camera calibration.

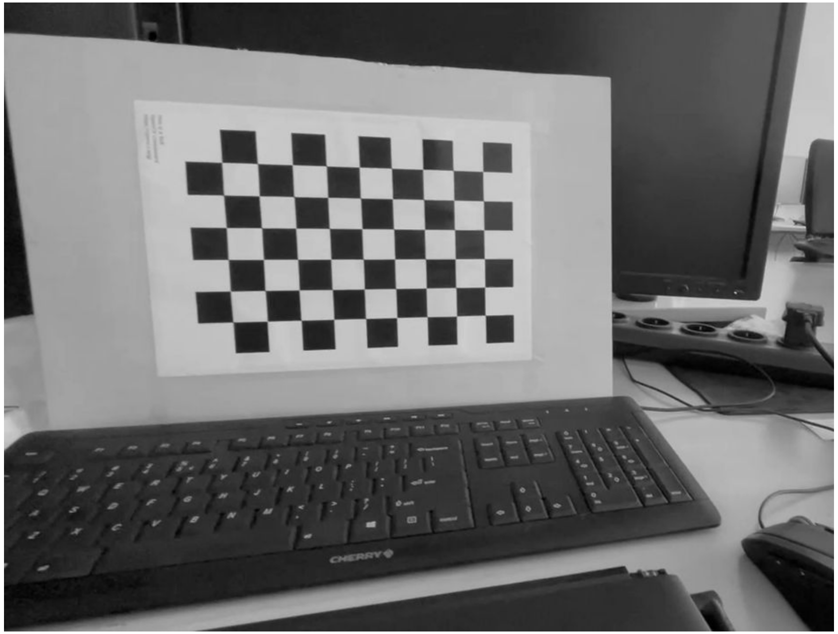
General Framework

- Proposed Framework for PupilCalib



Intrinsic Calibration of Camera

- Burst of 30 images with Chessboard pattern is taken.
- Chessboard corners are used to calculate the Instrinsic Parameters for Camera.
- Parameters are calculated for both World and Scene Cameras.

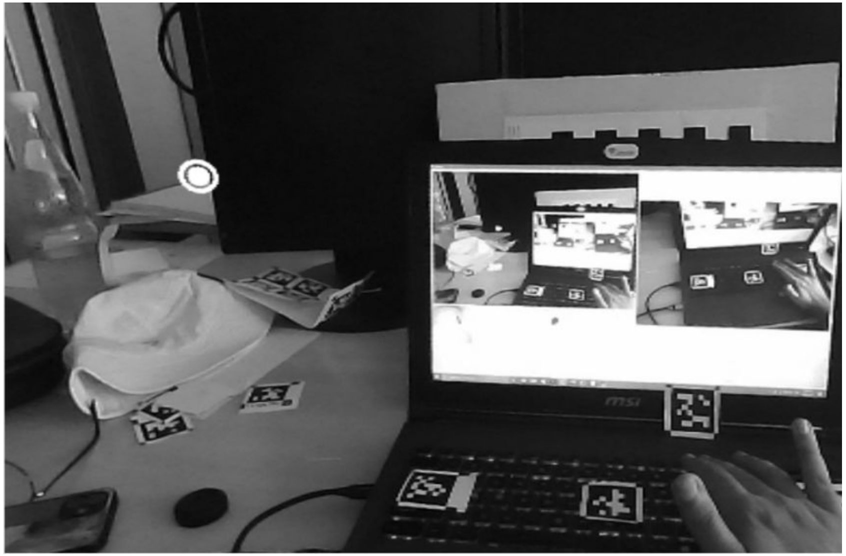


$$\begin{bmatrix} \alpha_x & s & c_x & 0 \\ 0 & \alpha_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Detecting Gaze and it’s projection.

- Pupil’s API is accessed to get the Gaze Point in 3D.
- Gaze Point is projected in Scene Camera.

3D Gaze Point = $\begin{bmatrix} x_{gaze} \\ y_{gaze} \\ z_{gaze} \\ 1 \end{bmatrix}$



Detecting Gaze and it's projection.

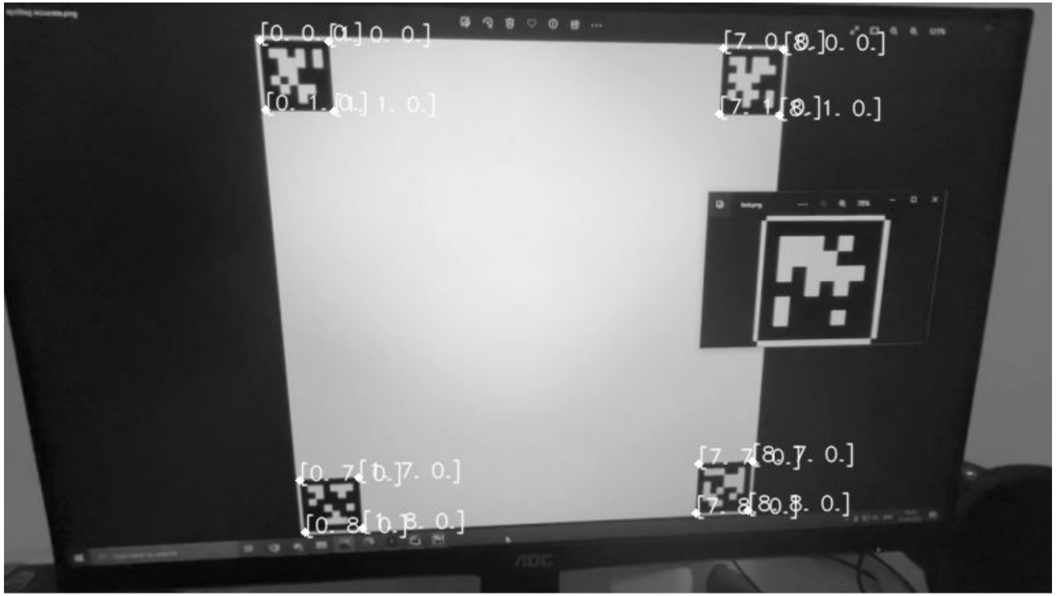
$$\begin{bmatrix} u_{gaze} \\ v_{gaze} \\ w_{gaze} \end{bmatrix} = \begin{bmatrix} \alpha_x & s & c_x & 0 \\ 0 & \alpha_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{scene} \begin{bmatrix} x_{gaze} \\ y_{gaze} \\ z_{gaze} \\ 1 \end{bmatrix}$$

$$x_{gaze}^{pix} = \frac{u_{gaze}}{w_{gaze}}$$

$$y_{gaze}^{pix} = \frac{v_{gaze}}{w_{gaze}}$$

Detecting AprilTags

- For both of the cameras, the AprilTags are detected using a detection algorithm.
- AprilTags are used as markers for global coordinates.
- By default, coplanar points are chosen by the API.



https://april.eecs.umich.edu/papers/details.php?name=wang2016iros

Extrinsic calibration for Camera

- Global coordinates are used to solve for rotation and translation vectors.
- This is done for both the World and the Scene Camera.
- Perspective-n-Point pose computation is performed.

$$\begin{bmatrix} \mathbf{R}_{world} & \mathbf{t}_{world} \\ 0 & 1 \end{bmatrix}_{4\times 4}$$

Projecting Gaze on World Camera

- Extrinsic and Intrinsic Camera parameters are used to project gaze on the world camera.
- Global gaze point is obtained in the global coordinate system.

$$\begin{bmatrix} X_{gaze} \\ Y_{gaze} \\ Z_{gaze} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{scene} & \mathbf{t}_{scene} \\ 0 & 1 \end{bmatrix}_{4\times 4}^{-1} \begin{bmatrix} x_{gaze} \\ y_{gaze} \\ z_{gaze} \\ 1 \end{bmatrix}$$

Projecting Gaze on World Camera

- Global Gaze point is used with the world camera extrinsic parameters to obtain Gaze point in world camera coordinate system.

$$\begin{bmatrix} \bar{x}_{gaze} \\ \bar{y}_{gaze} \\ \bar{z}_{gaze} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{world} & \mathbf{t}_{world} \\ 0 & 1 \end{bmatrix}_{4\times 4} \begin{bmatrix} X_{gaze} \\ Y_{gaze} \\ Z_{gaze} \\ 1 \end{bmatrix}$$

Exporting the results

- Converted to pixel coordinates.

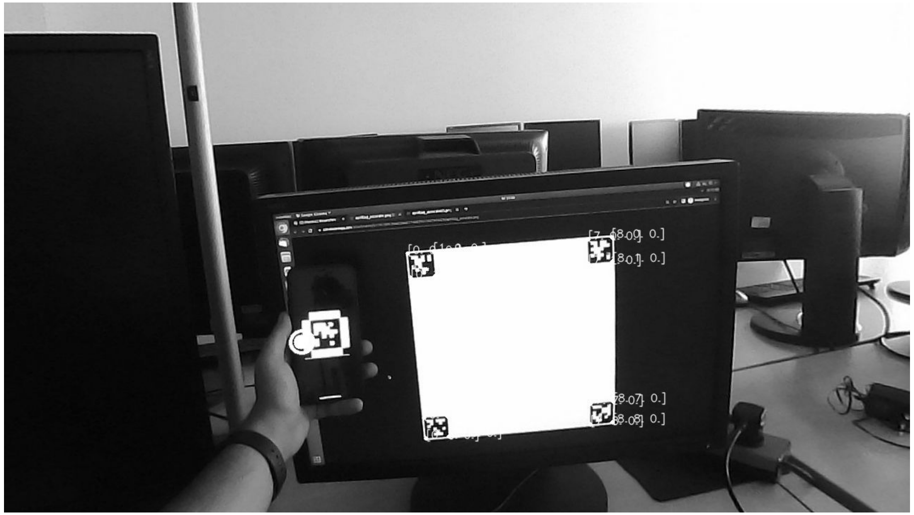
$$\begin{bmatrix} \bar{u}_{gaze} \\ \bar{v}_{gaze} \\ \bar{w}_{gaze} \end{bmatrix} = \begin{bmatrix} \alpha_x & s & c_x & 0 \\ 0 & \alpha_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{world} \begin{bmatrix} \bar{x}_{gaze} \\ \bar{y}_{gaze} \\ \bar{z}_{gaze} \\ 1 \end{bmatrix}$$

$$\bar{x}^{pix}_{gaze} = \frac{\bar{u}_{gaze}}{\bar{w}_{gaze}}$$

$$\bar{y}^{pix}_{gaze} = \frac{\bar{v}_{gaze}}{\bar{w}_{gaze}}$$

Exporting the result

- The projected 2D/3D point can be access via the API.



Metrics and Evaluation

- A target AprilTag is set and the user looks at the target AprilTag.
- The error is taken as the pixel distance between the target and the projected gaze in the world camera.
- The API provides a response at 15Hz for all the information.

Target Distance	Light Settings	Obervations	Mean	Standard Deviation
$<= 1m$	Well Lit Room	30	74.38	12.57
$> 1m$	Well Lit Room	20	102.82	15.90
$<= 1m$	Dim Lit Room	45	93.34	28.53
$> 1m$	Dim Lit Room	30	135.78	41.45

Table 1: PupilCalib performance with IDS 3080CP and Pupil Core in different settings

Limitations

- The API is limited by the response rate of the hardware chosen and can sometimes use old buffer data.
- Increasing the number of AprilTags does not improve the projection.
- Low light negatively affect the performance.
- Further gaze distance negatively affect the performance.

Target Distance	Light Settings	Obervations	Mean	Standard Deviation
$<= 1m$	Well Lit Room	30	74.38	12.57
$> 1m$	Well Lit Room	20	102.82	15.90
$<= 1m$	Dim Lit Room	45	93.34	28.53
$> 1m$	Dim Lit Room	30	135.78	41.45

Table 1: PupilCalib performance with IDS 3080CP and Pupil Core in different settings

Expanding the API

- Provides a modular class based structure with PyQt backbone for easy to use UI.
- Can be expanded to any eyetracker hardware.
- Needs updated implementation for accessing the buffer and gaze data for respective hardware.

Related Works

- Kassner, Moritz, William Patera, and Andreas Bulling. "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction." Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication. 2014.

Thank you for your attention!

DFKI GmbH
Department Augmented Vision
Trippstadterstr. 122
D-67663 Kaiserslautern

Ayush Chamoli

chamoli@rptu.de

https://av.dfki.de/