A New 3D Line of Gaze Estimation Method with Simple Marked Targets and Glasses

Samil Karahan[1,2], Yakup Genc[2], Yusuf Sinan Akgul[2]

[1]TUBITAK BILGEM, Anibal Street, 41470, Gebze, Kocaeli, Turkey
[2]GIT Vision Lab, http://vision.gyte.edu.tr, Department of Computer Engineering, Gebze Institute of Technology,41400, Kocaeli, Turkey

samil.karahan@tubitak.gov.tr
{ygenc, akgul}@bilmuh.gyte.edu.tr

PETMEI 2013

Content

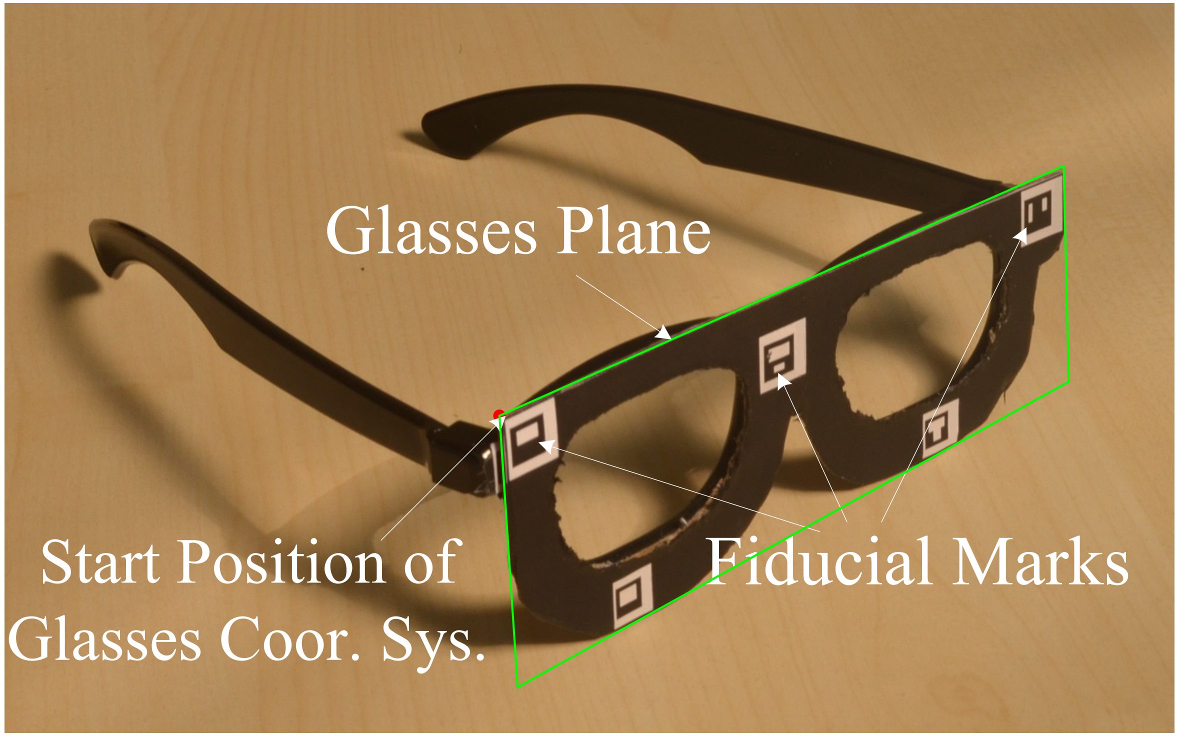
- Introduction
- The Simple Glasses and The Target Paper
- The 3D Geometry of System
- Training Stage
- Estimation of LoG
- LM Optimization
- Results
- Advantages and Drawbacks

Introduction

- We present a new Line of Gaze (LoG) method that uses a paper target with a hole for training and simple glasses for the head tracking.
- The fiducial marks are used for 3D localization via 3D camera geometry.
- The system doesn't need any extra camera or IR light sources.
- The system uses the 3D position of the cornea center and the radii of it to estimate LoG.

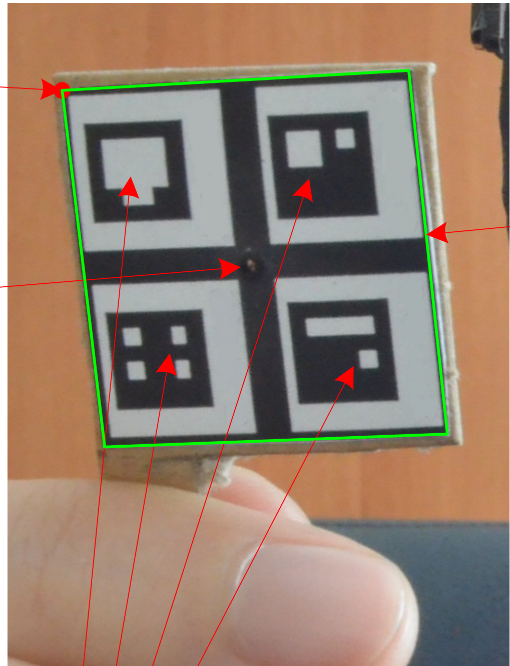
Simple Glasses

- Removed color filters on basic movie glasses
- The main purposes of glasses:
  - 3D head pose
  - Reference coordinate system on world

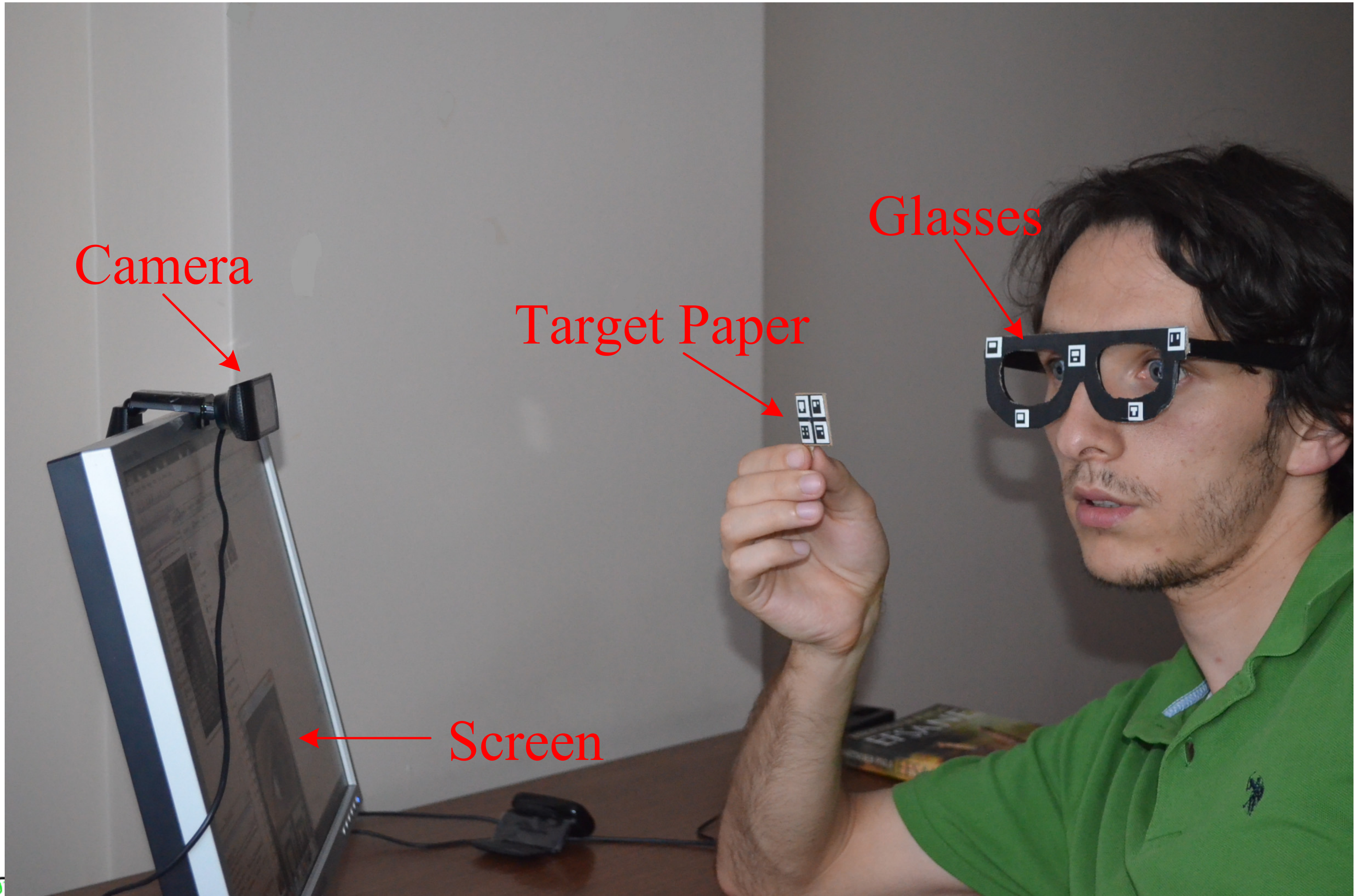


Target Paper

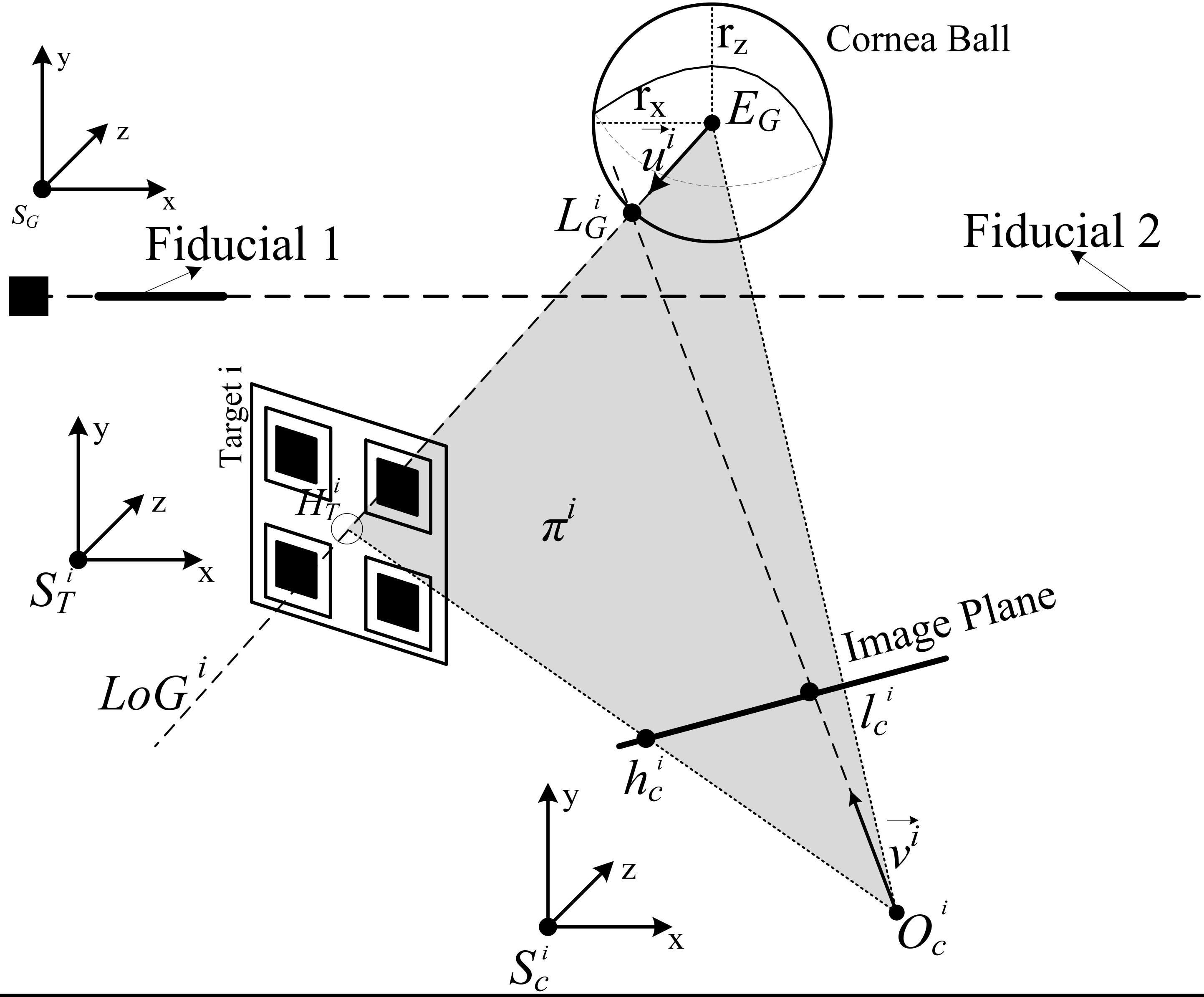
- The target paper makes our calibration more robust.
- It is used during training stage to find localization of cornea center and radii of it.



General Image of System

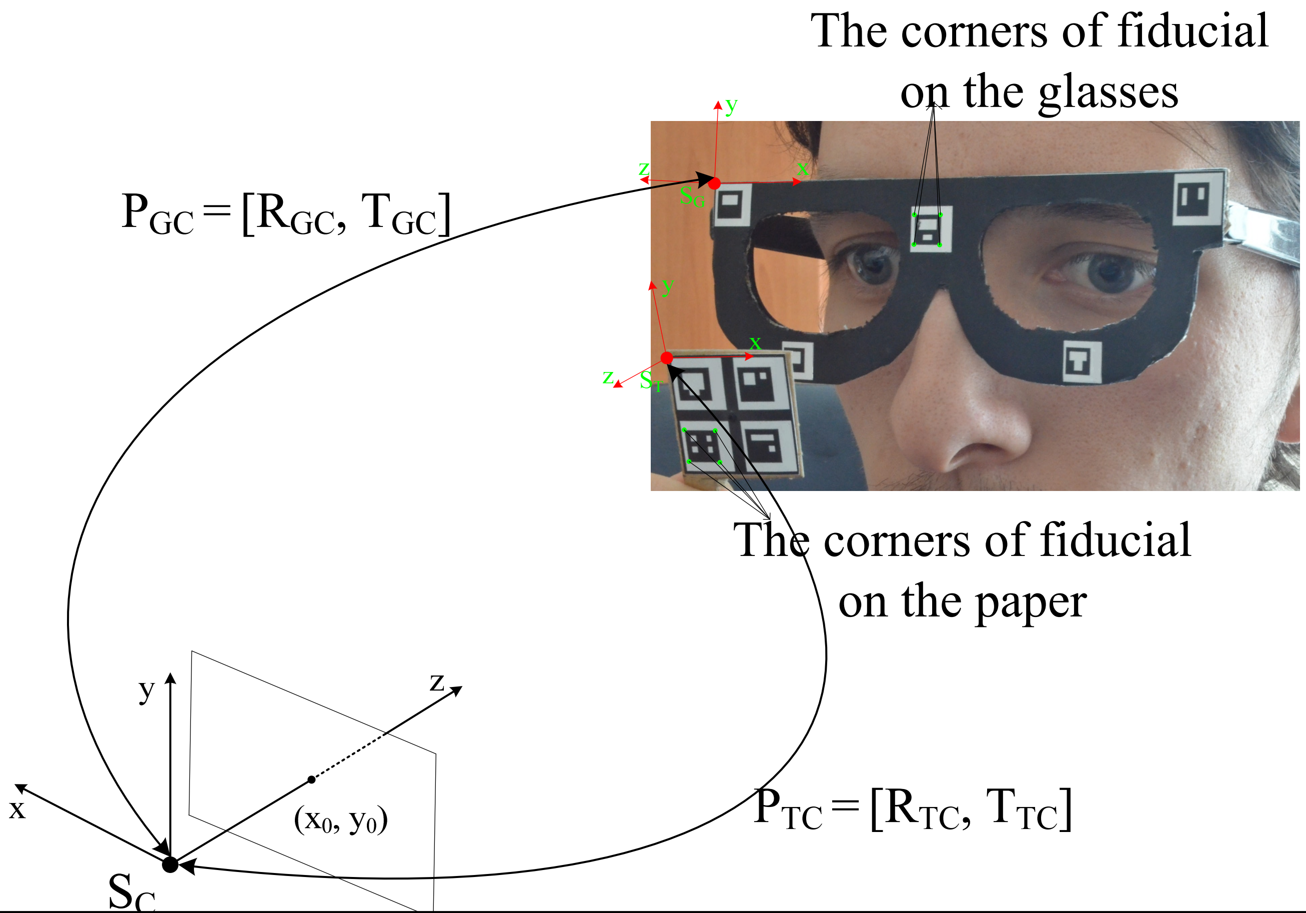


The 3D Geometry of System



Training Stage

- Detection of Fiducial Marker
- Calculation of Transformation Matrix P_{GC}^{i} and P_{TC}^{i}



Define A Plane Using Three Points

- The system defines a plane on which the 3D position of cornea center is located.
- Three points define a plane in homegenous geometry.

$$\begin{bmatrix} X_1^T \\ X_2^T \\ X_3^T \end{bmatrix} \pi = 0$$

The First Point : The 3D position of the camera center on the glasses coordinate system.

$$M_{GC}^i O_G^i = O_c^i \longrightarrow O_c^i = [0 \quad 0 \quad 0 \quad 1]^T$$

Define A Plane Using Three Points

- The Second Point : Pupil On Image
  - We detect the center of the pupil on an image by using the gradient field method of (Timm and Bart).

$$l_C^i = \left[\left(l_x^i - x_0\right)/d_{px} \quad \left(l_y^i - y_0\right)/d_{py} \quad f \quad 1 \right]$$

- The Third Point :The Target Position

$$h_C^i = \left[\left(h_x^i - x_0\right)/d_{px} \quad \left(h_y^i - y_0\right)/d_{py} \quad f \quad 1 \right]$$

Estimation of 3D Position of The Cornea

- All planes are used to estimate the 3D position of cornea.
- Common property of all defined planes is that the 3D position of cornea center is located on them.

$$\begin{bmatrix} \pi_1^T \\ \pi_2^T \\ \vdots \\ \vdots \\ \pi_n^T \end{bmatrix} \begin{bmatrix} E_{Gx}/E_{Gt} \\ E_{Gy}/E_{Gt} \\ E_{Gz}/E_{Gt} \\ 1 \end{bmatrix} = 0$$

↓

- SVD solution of this lineer system

The Radii of Cornea

- The target 3D position is transformed to glasses coordinate system via transformation matrix.

$$E_G + r.\overrightarrow{u_i} = O_G^i + a_i.\overrightarrow{v_i}$$

$\overrightarrow{u_i}$: From E_G towards H_G^i

$\overrightarrow{v_i}$: From O_G^i towards l_G^i

- r and a_i are unknown variables.

- All linear equations, which are derived from all defined planes, are used to estimate the radii.

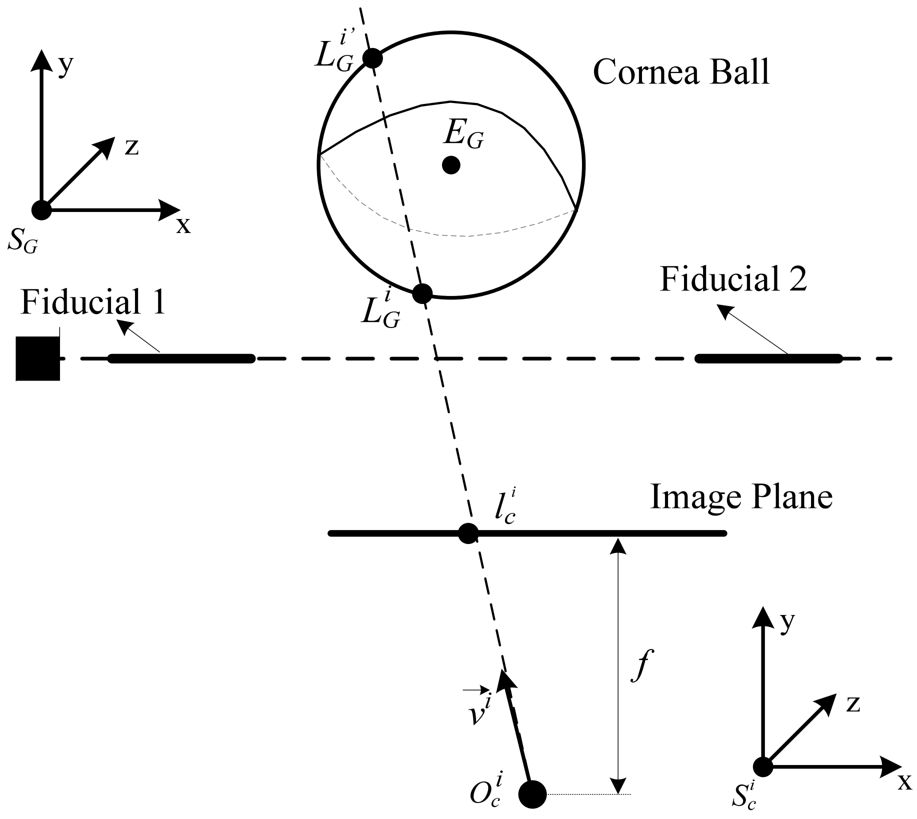
Test Stage : Estimation of Gaze

- r is known, but $\overrightarrow{u^i}$ is not because of not using the paper.

$$E_G + r.\overrightarrow{u_i} = O_G^i + a_i.\overrightarrow{v_i}$$

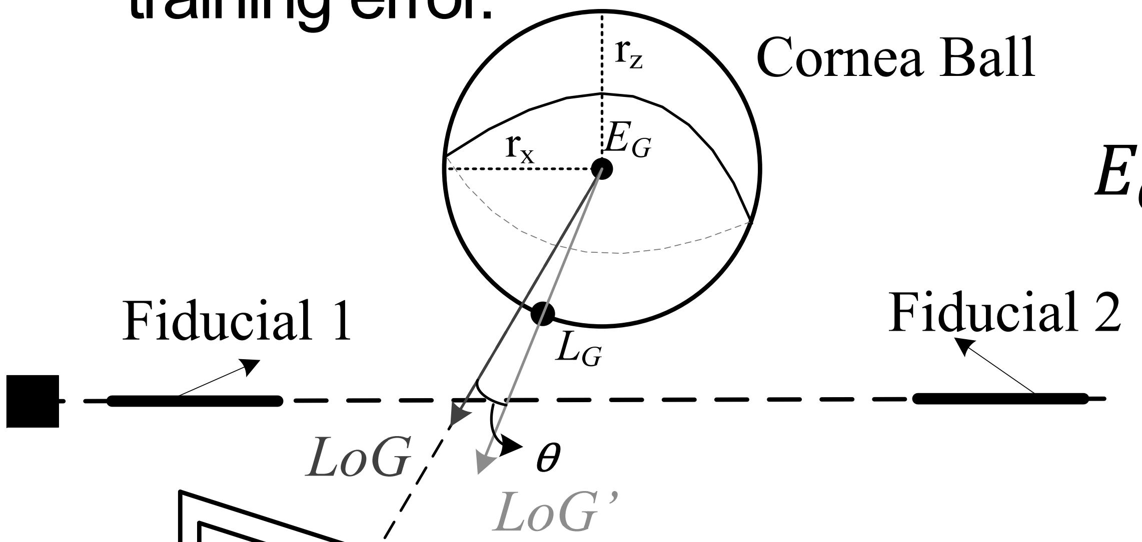
- $\overrightarrow{u^i}$ is unit vector;

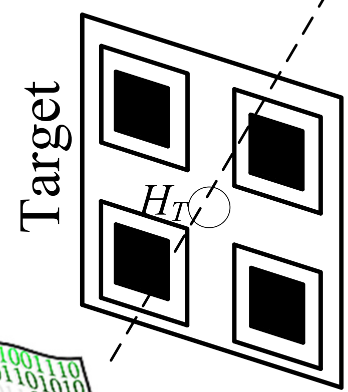
$$\left(O_G^i - E_G + a_i.\overrightarrow{v_i}\right)/r = \overrightarrow{u_i}$$



Levenberg-Marquardt Optimization

- SVD does not give exact solution because of model and calibration errors.
- The main goal of this optimization is minimized the system training error.





$$E_G + r.\overrightarrow{u_i} = O_G^i + a_i.\overrightarrow{v_i}$$

- The method updates E_G and r vectors to diffirence between estimated LoG and real LoG.

Results

- Properties of first user dataset.

Head Pose (Degree)	Cornea Center (x,y,z) cm	Radii Of Cornea (r_x, r_y, r_z) cm	Image Count
-10° < Head Pose <+10°	(3.32, 1.8, 4.22)	(1.38, 1.27, 1.31)	9
-20° < Head Pose <+20°	(3.31, 1.8, 4.25)	(1.40, 1.29, 1.33)	13
-∞° < Head Pose <+∞°	(3.32, 1.8, 4.24)	(1.38, 1.28, 1.33)	13

- The gaze estimation result of first user not applied LM

Experiments	Min.	Max.	Mean	Median
Pupil Detection (Timm and Bart)	0.31°	6.88°	3.72°	3.37°
-10° < Head Pose < +10°	2.7°	6.63°	4.88°	4.98°
-20° < Head Pose < +20°	1.05°	6.88°	3.47°	3.03°
-∞° < Head Pose <+∞°	0.31°	5.23°	3.17°	3.37°
Hand Detected Pupil Centers	0.39°	6.83°	3.58°	3.55°

Results

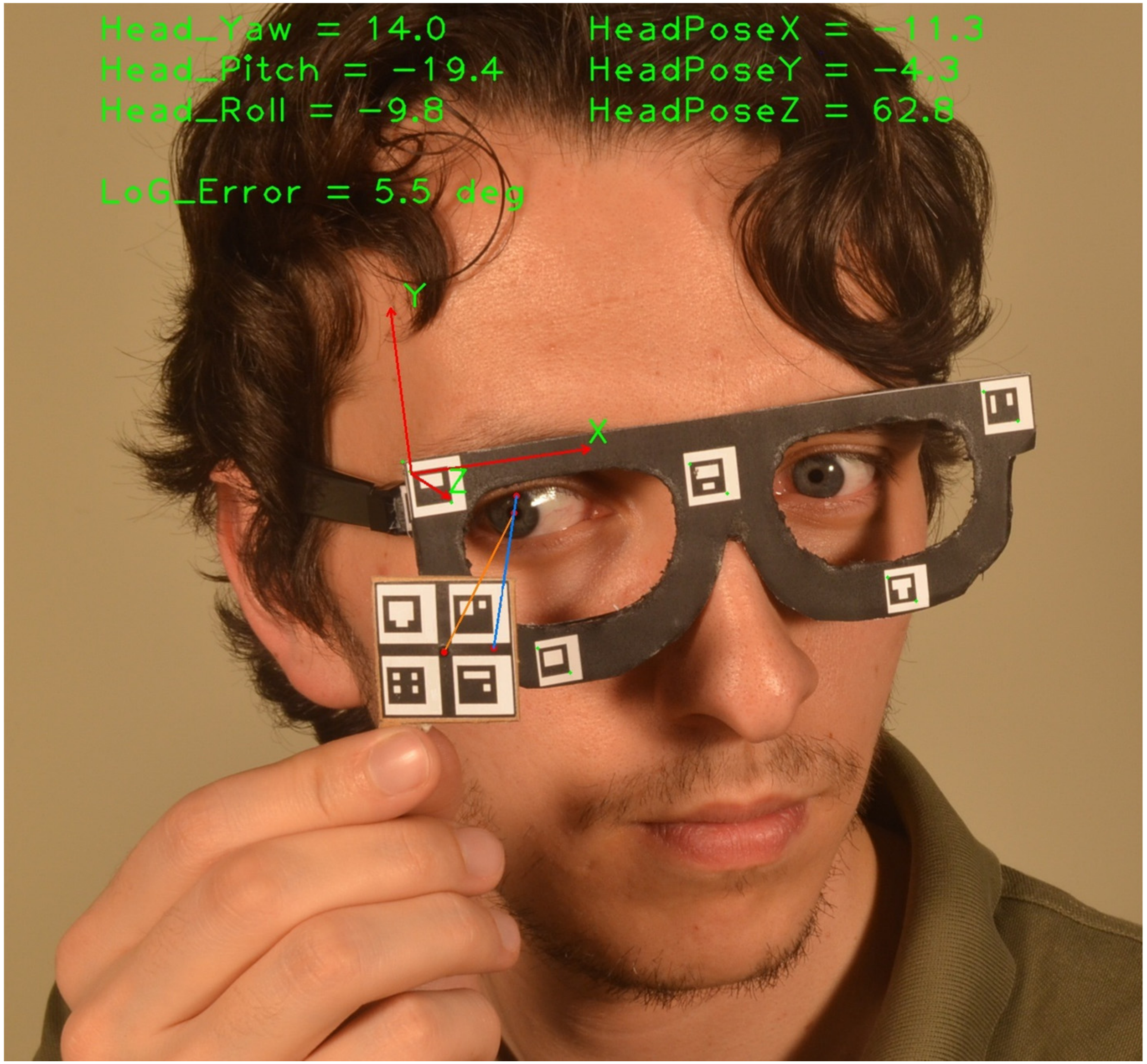
- The gaze estimation result of first user applied LM

Experiments	Min.	Max.	Mean	Median
Pupil Detection (Timm and Bart)	0.34°	6.85°	3.52°	3.43°
-10° < Head Pose < +10°	2.0°	6.65°	4.81°	5.07°
-20° < Head Pose < +20°	0.73°	6.85°	3.3°	3.01°
-∞° < Head Pose <+∞°	0.34°	4.96°	2.85°	3.32°
Hand Detected Pupil Centers	0.62°	6.41°	3.27°	3.22°

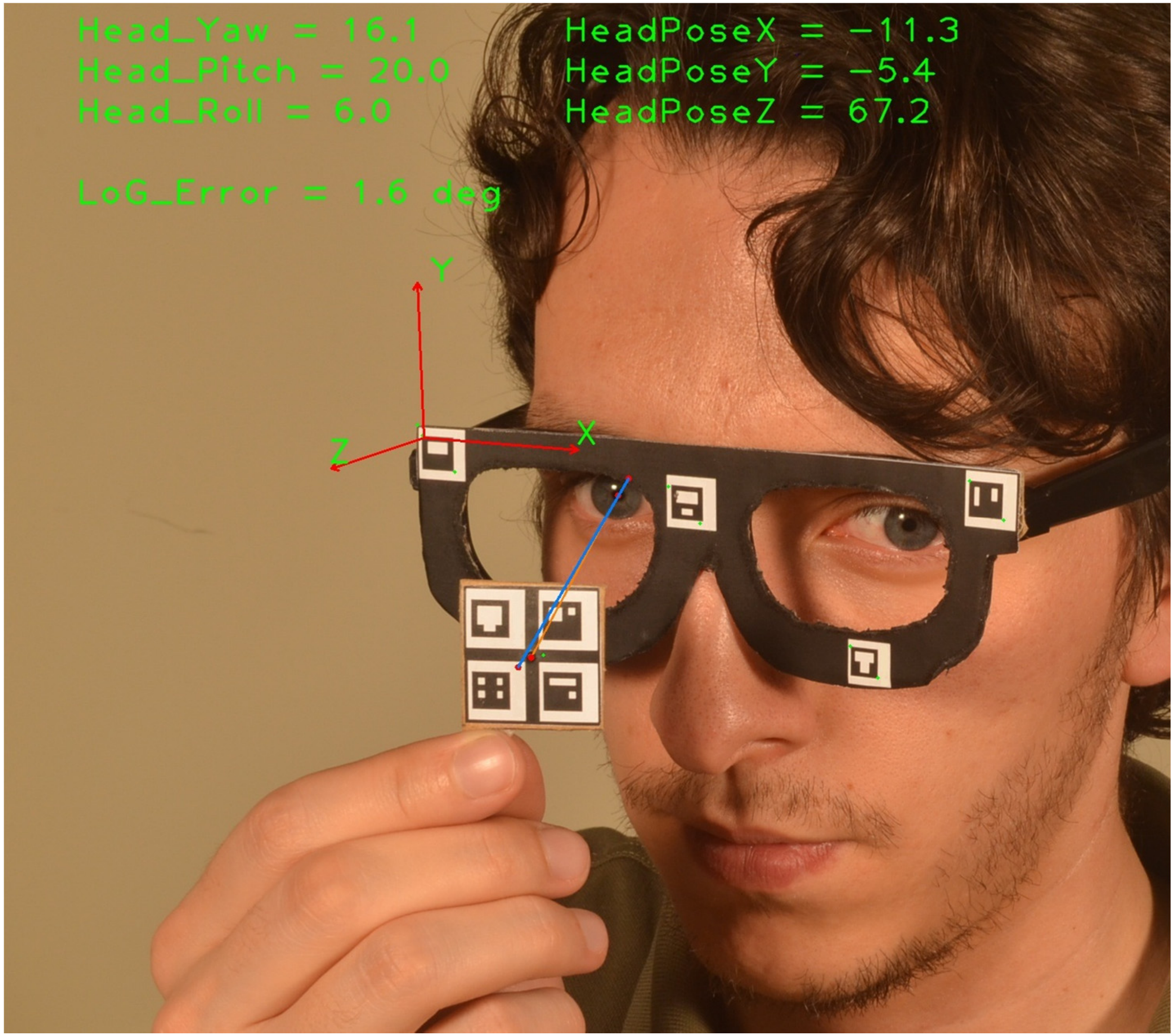
- The gaze estimation result of second user applied LM

Experiments	Min.	Max.	Mean	Median
-5° < Head Pose < +5°	2.40°	8.52°	4.81°	3.84°
-10° < Head Pose < +10°	0.50°	6.81°	3.21°	3.11°
-∞° < Head Pose <+∞°	1.31°	8.97°	4.22°	2.99°
Hand Detected Pupil Centers	0.50°	8.97°	4.12°	3.71°

Some Result Images of Users



Some Result Images of Users



Some Result Images of Users

Some Result Images of Users

Advantages and Drawbacks

- Some Advantages of System
  - Head movements are not restricted.
  - More robust calibration with the target paper.
  - Calculation of head position and orientation.
  - Calculation of the 3D position and radii of cornea.
- Drawbacks of System
  - The glasses is assumed fixed
  - The glasses occludes the eye when head position is extreme.

Thank You!!

Any Questions??