

# Remote, Non-Contact Gaze Estimation with Minimal Subject Cooperation

by

Elias Daniel Guestrin

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy

Graduate Departments of the Edward S. Rogers Sr. Department of Electrical and  
Computer Engineering and the Institute of Biomaterials and Biomedical Engineering,  
University of Toronto

© Copyright by Elias Daniel Guestrin 2010

# Remote, Non-Contact Gaze Estimation with Minimal Subject Cooperation

Elias Daniel Guestrin

Doctor of Philosophy

Edward S. Rogers Sr. Department of Electrical and Computer Engineering and  
Institute of Biomaterials and Biomedical Engineering  
University of Toronto

2010

## Abstract

This thesis presents a novel system that estimates the point-of-gaze (where a person is looking at) remotely while allowing for free head movements and minimizing personal calibration requirements. The point-of-gaze is estimated from the pupil and corneal reflections (virtual images of infrared light sources that are formed by reflection on the front corneal surface, which acts as a convex mirror) extracted from eye images captured by video cameras. Based on the laws of geometrical optics, a detailed general mathematical model for point-of-gaze estimation using the pupil and corneal reflections is developed. Using this model, the full range of possible system configurations (from one camera and one light source to multiple cameras and light sources) is analyzed. This analysis shows that two cameras and two light sources is the simplest system configuration that can be used to reconstruct the optic axis of the eye in 3-D space, and therefore measure eye movements, without the need for personal calibration. To estimate the point-of-gaze, a simple single-point personal calibration procedure is needed. The performance of the point-of-gaze estimation depends on the geometrical arrangement of the cameras and light sources and the method used to reconstruct the optic axis of the eye. Using a comprehensive simulation framework developed from the mathematical model, the performance of several gaze

## Abstract

estimation methods of varied complexity is investigated for different geometrical system setups in the presence of noise in the extracted eye features, deviation of the corneal shape from the ideal spherical shape and errors in system parameters. The results of this investigation indicate the method(s) and geometrical setup(s) that are optimal for different sets of conditions, thereby providing guidelines for system implementation. Experimental results with adults, obtained with a system that follows those guidelines, exhibit RMS point-of-gaze estimation errors of 0.4-0.6° of visual angle (comparable to the best commercially available systems, which require multiple-point personal calibration procedures). Preliminary results with infants demonstrate the ability of the proposed system to record infants' visual scanning patterns, enabling applications that are very difficult or impossible to carry out with previously existing technologies (e.g., study of infants' visual and oculomotor systems).

## Acknowledgments

I would like to express my gratitude to the following persons:

- My research supervisor, Prof. Moshe Eizenman, for his invaluable support and for giving me an almost unlimited freedom in my work.
- The members of my Ph.D. Advisory Committee, Prof. Berj Bardakjian and Prof. Manfredi Maggiore, for their valuable feedback and for serving in the Departmental and Final Oral Examination Committees.
- The Chair of my Departmental Oral Examination Committee, Prof. Shahrokh Valaei.
- Prof. Milos Popovic for his support and advice in the transition to the next stage of my academic career and for serving in my Final Oral Examination Committee.
- The External Examiner, Prof. Qiang Ji from the Rensselaer Polytechnic Institute, for his appraisal of this thesis and for also serving in my Final Oral Examination Committee.
- Prof. Carlos H. Morimoto from the Universidade de São Paulo for kindly providing the eye images shown in Fig. 1.3.
- Mr. Tony Jares and Mr. Franz Schuh for their invaluable assistance in building the prototype point-of-gaze estimation system.
- The staff of the Graduate Office of the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, Ms. Sarah Cherian (former Graduate Studies Administrator), Ms. Darlene Gorzo (present Graduate Studies Administrator) and Ms. Judith Levene (Graduate Programs Assistant) for their friendly and prompt assistance over the years. I especially want to thank Sarah for being so caring and for her invaluable support in difficult and stressful times. Her kindness will stay in my memory forever.
- The staff of the Institute of Biomaterials and Biomedical Engineering, Ms. Esther Oostdyk (Business Officer), Ms. Yvonne MacNeil and Ms. Sheema Inayatulla (former Managers of Operations) and Ms. Tammy Chan (former Student Affairs Coordinator), for their assistance in diverse matters.

## Acknowledgments

- The parents of the infants that participated in the experiments for their kind cooperation and patience. In particular, I want to thank the parents of one of the infants for granting us permission to use images of their child, such as the one shown in Fig. 1.2, in academic publications.
- My former and present fellow graduate students, Bryon Braymore, Andrew Rae, Ruth Sapir-Pichhadze, Jeff Kang, Jerry Lam, Dmitri Model and Daniel Cassel for enriching my graduate experience and for many interesting discussions. In particular, I want to thank Jerry for his work on the image processing algorithms for binocular eye-tracking and Jeff and Dmitri for their assistance in C++ coding and helping supervise summer students.
- The numerous summer students that have contributed to the implementation, testing and debugging of the software for the gaze estimation systems developed in our lab.
- My uncle, Héctor, and my aunt, Delfina, for their encouragement.
- My friend Karen, with whom I shared my life during this long journey.
- Most importantly, my parents, Ester and Oscar, my sister, Analía, my brothers, Abraham and David, and my almost-97-year old grandmother, Mecha, for their unflagging support and encouragement over the years.

I would also like to thank the Institute of Electrical and Electronics Engineers, Inc. (IEEE) for granting me permission to reproduce the diagrams shown in Fig. 1.5, Fig. 1.6 and Fig. 1.7.

Lastly, I would like to acknowledge the financial support from the University of Toronto, the Edward S. Rogers Sr. Scholarship program, and the Ontario Graduate Scholarship (OGS) program, as well as the grants to Prof. Moshe Eizenman from the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC) that made the development of the project possible.

## Table of Contents

<b>Acknowledgments .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>vi</b>
<b>List of Tables .....</b>	<b>xii</b>
<b>List of Figures.....</b>	<b>xix</b>
<b>List of Most Frequently Used Acronyms.....</b>	<b>xxii</b>
<b>Chapter 1 - Introduction .....</b>	<b>1</b>
1.1 Review of remote, non-contact gaze estimation methods .....	2
1.1.1 Appearance-based / view-based.....	3
1.1.2 Pupil/iris centers combined with 3-D head pose estimation from facial features...	4
1.1.3 Limbus image back-projection .....	5
1.1.4 Pupil and corneal reflections.....	6
1.1.4.1 “Cross-ratios” methodologies.....	10
1.1.4.2 Corneal reflection-pupil center vector.....	11
1.1.4.2.1 Pioneering work - Principles of operation.....	12
1.1.4.2.2 Behavior of the corneal reflection-pupil center vector in the presence of head movements.....	16
1.1.4.2.3 Subsequent works.....	17
1.1.4.3 Direct model-based 3-D gaze reconstruction .....	20
1.2 Goals .....	22
1.3 Organization of this thesis .....	23
<b>Chapter 2 - Mathematical Model .....</b>	<b>27</b>
2.1 Introduction.....	27
2.2 Eye model .....	27
2.2.1 Eye coordinate system (ECS) .....	29

2.2.2	Visual axis.....	30
2.2.3	Corneal model.....	31
2.2.3.1	Spherical corneal model .....	31
2.2.3.2	Ellipsoidal corneal model .....	32
2.2.4	Pupil model .....	35
2.3	Camera model .....	36
2.3.1	Camera coordinate system (CCS).....	36
2.3.2	Image coordinate system (ICS).....	36
2.3.3	Perspective projection .....	36
2.3.4	Lens distortion .....	38
2.3.5	Transformation to the ICS .....	38
2.4	Light source model .....	39
2.5	Integration of the eye, camera and light source models .....	39
2.5.1	World coordinate system (WCS) .....	39
2.5.2	Relation between the ECS, CCS and WCS .....	40
2.5.3	Eye model with respect to the WCS .....	41
2.5.4	Integrated model .....	43
2.5.4.1	Integrated model for a spherical cornea .....	43
2.5.4.2	Generalization of the integrated model to non-spherical corneas .....	48
2.6	Generation of synthetic eye feature data.....	50
2.6.1	Orientation of the ECS with respect to the WCS.....	50
2.6.2	Points of reflection and refraction.....	51
2.6.2.1	Points of reflection and refraction on a spherical cornea .....	51
2.6.2.1.1	Points of reflection on a spherical cornea.....	51
2.6.2.1.2	Points of refraction on a spherical cornea .....	53
2.6.2.2	Points of reflection and refraction on an ellipsoidal cornea .....	56

2.6.2.2.1	Points of reflection on an ellipsoidal cornea .....	57
2.6.2.2.2	Points of refraction on an ellipsoidal cornea .....	58
2.6.3	Validation of the points of reflection and refraction.....	59
2.6.4	Projection onto the camera image plane .....	60
2.7	Summary of model assumptions and simplifications .....	60
<b>Chapter 3 - Gaze Estimation with a Spherical Corneal Model.....</b>	<b>62</b>	
3.1	Introduction.....	62
3.2	Determination of the 3-D coordinates of the center of curvature of the cornea .....	63
3.2.1	One camera and one light source .....	65
3.2.2	One camera and two light sources .....	66
3.2.3	One camera and more than two light sources .....	69
3.2.4	Two cameras and one light source.....	69
3.2.5	Two cameras and two light sources .....	72
3.2.6	Two cameras and more than two light sources .....	76
3.2.7	Beyond two cameras (with multiple light sources).....	76
3.3	Reconstruction of the optic axis of the eye .....	76
3.3.1	Based on the pupil center with explicit refraction model – Method I (two cameras only) .....	78
3.3.2	Based on the pupil center with explicit refraction model – Method II .....	79
3.3.2.1	One camera .....	81
3.3.2.2	Two cameras.....	81
3.3.3	Based on the pupil center with implicit refraction model .....	83
3.3.3.1	One camera .....	83
3.3.3.2	Two cameras.....	84
3.3.4	Based on the pupil boundary with explicit refraction model .....	85
3.3.4.1	One camera .....	87

3.3.4.2 Two cameras.....	88
3.3.5 Brief summary of the different methods .....	89
3.4 Reconstruction of the visual axis and estimation of the point-of-gaze.....	89
3.5 Summary of the analysis of the different system configurations and discussion of their personal calibration requirements .....	90
<b>Chapter 4 - Simulation Study .....</b>	<b>95</b>
4.1 Introduction.....	95
4.2 Point-of-gaze estimation methods studied .....	96
4.2.1 Calculation of the 3-D coordinates of the center of curvature of the cornea .....	96
4.2.2 Subsequent reconstruction of the optic axis of the eye in 3-D space.....	98
4.3 General simulation methodology.....	99
4.4 Model parameters for the simulations .....	100
4.4.1 Eye parameters.....	100
4.4.2 System parameters .....	102
4.4.2.1 Cameras .....	102
4.4.2.1.1 Extrinsic parameters.....	102
4.4.2.1.2 Intrinsic parameters .....	103
4.4.2.2 Light sources .....	104
4.4.2.3 Scene.....	105
4.5 Evaluation of the inherent point-of-gaze estimation bias .....	105
4.6 Sensitivities of the point-of-gaze estimation with two cameras and two light sources.....	108
4.6.1 Sensitivity to noise in the image coordinates of the eye features .....	110
4.6.2 Effect of corneal asphericity .....	120
4.6.3 Sensitivity to error in a selected system parameter.....	130
4.6.4 Summary of simulation results .....	135

4.7 Sensitivities of the point-of-gaze estimation with two cameras and four light sources.....	136
4.7.1 Effects of noise in the image coordinates of the eye features and corneal asphericity as a function of system setup for a fixed position of the center of the eyeball .....	139
4.7.2 Effects of noise in the image coordinates of the eye features and corneal asphericity for the selected system setup under head movements .....	147
<b>Chapter 5 - Experimental Results .....</b>	<b>153</b>
5.1 Introduction.....	153
5.2 System setups.....	153
5.2.1 Initial system setup .....	154
5.2.2 New system setup .....	154
5.3 System calibration.....	155
5.4 Experimental results with adults .....	156
5.5 Experimental results with infants.....	177
5.6 Summary of experimental results .....	179
<b>Chapter 6 - Summary of Contributions and Future Work .....</b>	<b>180</b>
6.1 Summary of contributions.....	180
6.2 Future work .....	183
6.2.1 Point-of-gaze estimation methodologies.....	183
6.2.2 Applications .....	184
<b>References .....</b>	<b>186</b>
<b>Appendix A - Relation Between the ECS and the WCS .....</b>	<b>195</b>
A.1 Notes on the torsion of the eye .....	198
A.1.1 Analytical expressions of Listing's law .....	200
A.1.1.1 Listing's law and the generation of synthetic eye feature data.....	203
A.1.1.2 Listing's law and the estimation of the point-of-gaze .....	205

A.1.1.3 Listing's law and the single-point personal calibration procedure .....	206
A.1.2 Quantification of the torsion predicted by Listing's law .....	208
<b>Appendix B - Outline of the Derivation of (2.4)-(2.6).....</b>	<b>209</b>
<b>Appendix C - Relation Between the CCS and the WCS .....</b>	<b>210</b>
<b>Appendix D - Appendices to Chapter 3 .....</b>	<b>213</b>
D.1 Appendix to Section 3.2.2 .....	213
D.2 Appendix to Section 3.2.4 .....	217
<b>Appendix E - Sensitivity to Errors in System Parameters .....</b>	<b>221</b>
E.1 Sensitivity to errors in the coordinates of the light sources .....	221
E.2 Sensitivity to errors in the camera parameters .....	222

## List of Tables

Table 4.1: Eye parameters for the simulations.....	101
Table 4.2: Intrinsic camera parameters for the simulations .....	104
Table 4.3: Inherent point-of-gaze estimation bias when $o_X = 50$ mm and the pupil diameter is 8 mm.....	107
Table 4.4: Inherent point-of-gaze estimation bias when $o_X = 150$ mm and the pupil diameter is 8 mm.....	107
Table 4.5: System setups.....	110
Table 4.6: Noise sensitivity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm].....	112
Table 4.7: Noise sensitivity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Dispersion [mm].....	112
Table 4.8: Noise sensitivity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm].....	113
Table 4.9: Noise sensitivity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Dispersion [mm].....	113
Table 4.10: Noise sensitivity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm].....	114
Table 4.11: Noise sensitivity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Dispersion [mm].....	114
Table 4.12: Noise sensitivity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm].....	115
Table 4.13: Noise sensitivity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Dispersion [mm].....	115
Table 4.14: Effect of corneal asphericity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm].....	122
Table 4.15: Effect of corneal asphericity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm].....	122
Table 4.16: Effect of corneal asphericity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm].....	123
Table 4.17: Effect of corneal asphericity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm].....	123

Table 4.18: Effect of corneal asphericity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm] .....	124
Table 4.19: Effect of corneal asphericity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm] .....	124
Table 4.20: Effect of corneal asphericity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm] .....	125
Table 4.21: Effect of corneal asphericity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm] .....	125
Table 4.22: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm] .....	131
Table 4.23: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm] .....	131
Table 4.24: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm] .....	132
Table 4.25: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm] .....	132
Table 4.26: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm] .....	133
Table 4.27: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm] .....	133
Table 4.28: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm] .....	134
Table 4.29: Effect of a $0.1^\circ$ error in the tilt angle of camera 1 – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm] .....	134
Table 4.30: Point-of-gaze estimation methods subject to further consideration .....	136
Table 4.31: Effect of corneal asphericity – System with 4 light sources with $l_{1-2,y} = 0$ mm and $o_X = 150$ mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) – MAX Bias [mm] .....	138
Table 4.32: Effect of corneal asphericity – System with 4 light sources with $l_{1-2,y} = 0$ mm and $o_X = 150$ mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) – RMS Bias [mm] .....	138
Table 4.33: Angle of convergence of the cameras, $\theta_{\text{opt.ax.cam's}}$ , for each tested value of $o_X$ .....	140
Table 4.34: Positions of the center of the eyeball ( <b>d</b> ) for the simulations with head movements [mm] .....	148

Table 4.35: Noise sensitivity – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Dispersion</i> [mm] .....	150
Table 4.36: Noise sensitivity – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Dispersion</i> [ $^{\circ}$ ] .....	150
Table 4.37: Effect of corneal asphericity – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Bias</i> [mm] .....	150
Table 4.38: Effect of corneal asphericity – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Bias</i> [ $^{\circ}$ ] .....	150
Table 4.39: Total point-of-gaze estimation error – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Error</i> [mm] .....	151
Table 4.40: Total point-of-gaze estimation error – System with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm – <i>RMS Error</i> [ $^{\circ}$ ] .....	151
Table 5.1: Experimental point-of-gaze estimation error for Subject 1 – <i>Left eye – RMS Bias</i> [mm] .....	162
Table 5.2: Experimental point-of-gaze estimation error for Subject 1 – <i>Right eye – RMS Bias</i> [mm] .....	162
Table 5.3: Experimental point-of-gaze estimation error for Subject 1 – <i>Left eye – RMS Dispersion</i> [mm] .....	162
Table 5.4: Experimental point-of-gaze estimation error for Subject 1 – <i>Right eye – RMS Dispersion</i> [mm] .....	162
Table 5.5: Experimental point-of-gaze estimation error for Subject 1 – <i>Left eye – RMS Error</i> [mm] .....	163
Table 5.6: Experimental point-of-gaze estimation error for Subject 1 – <i>Right eye – RMS Error</i> [mm] .....	163
Table 5.7: Experimental point-of-gaze estimation error for Subject 1 – <i>Left eye – RMS Error</i> [ $^{\circ}$ ] .....	163
Table 5.8: Experimental point-of-gaze estimation error for Subject 1 – <i>Right eye – RMS Error</i> [ $^{\circ}$ ] .....	163
Table 5.9: Experimental point-of-gaze estimation error for Subject 2 – <i>Left eye – RMS Bias</i> [mm] .....	167
Table 5.10: Experimental point-of-gaze estimation error for Subject 2 – <i>Right eye – RMS Bias</i> [mm] .....	167
Table 5.11: Experimental point-of-gaze estimation error for Subject 2 – <i>Left eye – RMS Dispersion</i> [mm] .....	167

Table 5.12: Experimental point-of-gaze estimation error for Subject 2 – <i>Right eye – RMS Dispersion [mm]</i> .....	167
Table 5.13: Experimental point-of-gaze estimation error for Subject 2 – <i>Left eye – RMS Error [mm]</i> .....	168
Table 5.14: Experimental point-of-gaze estimation error for Subject 2 – <i>Right eye – RMS Error [mm]</i> .....	168
Table 5.15: Experimental point-of-gaze estimation error for Subject 2 – <i>Left eye – RMS Error [°]</i> .....	168
Table 5.16: Experimental point-of-gaze estimation error for Subject 2 – <i>Right eye – RMS Error [°]</i> .....	168
Table 5.17: Experimental point-of-gaze estimation error for Subject 3 – <i>Left eye – RMS Bias [mm]</i> .....	172
Table 5.18: Experimental point-of-gaze estimation error for Subject 3 – <i>Right eye – RMS Bias [mm]</i> .....	172
Table 5.19: Experimental point-of-gaze estimation error for Subject 3 – <i>Left eye – RMS Dispersion [mm]</i> .....	172
Table 5.20: Experimental point-of-gaze estimation error for Subject 3 – <i>Right eye – RMS Dispersion [mm]</i> .....	172
Table 5.21: Experimental point-of-gaze estimation error for Subject 3 – <i>Left eye – RMS Error [mm]</i> .....	173
Table 5.22: Experimental point-of-gaze estimation error for Subject 3 – <i>Right eye – RMS Error [mm]</i> .....	173
Table 5.23: Experimental point-of-gaze estimation error for Subject 3 – <i>Left eye – RMS Error [°]</i> .....	173
Table 5.24: Experimental point-of-gaze estimation error for Subject 3 – <i>Right eye – RMS Error [°]</i> .....	173
Table A.1: Quantification of the torsion predicted by Listing’s law (all angles are in degrees) .....	208
Table E.1: Inherent bias (no error in system parameters) – <i>MAX Bias [mm]</i> .....	226
Table E.2: Inherent bias (no error in system parameters) – <i>RMS Bias [mm]</i> .....	226
Table E.3: Effect of error in the coordinates of the light sources – Error in $l_{1,X} = 10 \text{ mm}$ – <i>MAX Bias [mm]</i> .....	226

Table E.4: Effect of error in the coordinates of the light sources – Error in $l_{1,X} = 10$ mm – <i>RMS Bias [mm]</i> .....	226
Table E.5: Effect of error in the coordinates of the light sources – Error in $l_{1,Y} = 10$ mm – <i>MAX Bias [mm]</i> .....	227
Table E.6: Effect of error in the coordinates of the light sources – Error in $l_{1,Y} = 10$ mm – <i>RMS Bias [mm]</i> .....	227
Table E.7: Effect of error in the coordinates of the light sources – Error in $l_{1,Z} = 10$ mm – <i>MAX Bias [mm]</i> .....	227
Table E.8: Effect of error in the coordinates of the light sources – Error in $l_{1,Z} = 10$ mm – <i>RMS Bias [mm]</i> .....	227
Table E.9: Effect of error in the coordinates of the light sources – Error in $l_{2,X} = 10$ mm – <i>MAX Bias [mm]</i> .....	228
Table E.10: Effect of error in the coordinates of the light sources – Error in $l_{2,X} = 10$ mm – <i>RMS Bias [mm]</i> .....	228
Table E.11: Effect of error in the coordinates of the light sources – Error in $l_{2,Y} = 10$ mm – <i>MAX Bias [mm]</i> .....	228
Table E.12: Effect of error in the coordinates of the light sources – Error in $l_{2,Y} = 10$ mm – <i>RMS Bias [mm]</i> .....	228
Table E.13: Effect of error in the coordinates of the light sources – Error in $l_{2,Z} = 10$ mm – <i>MAX Bias [mm]</i> .....	229
Table E.14: Effect of error in the coordinates of the light sources – Error in $l_{2,Z} = 10$ mm – <i>RMS Bias [mm]</i> .....	229
Table E.15: Effect of error in the coordinates of the light sources – Error in $l_{3,X} = 10$ mm – <i>MAX Bias [mm]</i> .....	229
Table E.16: Effect of error in the coordinates of the light sources – Error in $l_{3,X} = 10$ mm – <i>RMS Bias [mm]</i> .....	229
Table E.17: Effect of error in the coordinates of the light sources – Error in $l_{3,Y} = 10$ mm – <i>MAX Bias [mm]</i> .....	230
Table E.18: Effect of error in the coordinates of the light sources – Error in $l_{3,Y} = 10$ mm – <i>RMS Bias [mm]</i> .....	230
Table E.19: Effect of error in the coordinates of the light sources – Error in $l_{3,Z} = 10$ mm – <i>MAX Bias [mm]</i> .....	230
Table E.20: Effect of error in the coordinates of the light sources – Error in $l_{3,Z} = 10$ mm – <i>RMS Bias [mm]</i> .....	230

Table E.21: Effect of error in the coordinates of the light sources – Error in $l_{4,X} = 10 \text{ mm} - MAX Bias [\text{mm}]$ .....	231
Table E.22: Effect of error in the coordinates of the light sources – Error in $l_{4,X} = 10 \text{ mm} - RMS Bias [\text{mm}]$ .....	231
Table E.23: Effect of error in the coordinates of the light sources – Error in $l_{4,Y} = 10 \text{ mm} - MAX Bias [\text{mm}]$ .....	231
Table E.24: Effect of error in the coordinates of the light sources – Error in $l_{4,Y} = 10 \text{ mm} - RMS Bias [\text{mm}]$ .....	231
Table E.25: Effect of error in the coordinates of the light sources – Error in $l_{4,Z} = 10 \text{ mm} - MAX Bias [\text{mm}]$ .....	232
Table E.26: Effect of error in the coordinates of the light sources – Error in $l_{4,Z} = 10 \text{ mm} - RMS Bias [\text{mm}]$ .....	232
Table E.27: Effect of error in the camera parameters – Error in $o_{1,X} = 1 \text{ mm} - MAX Bias [\text{mm}]$ .....	232
Table E.28: Effect of error in the camera parameters – Error in $o_{1,X} = 1 \text{ mm} - RMS Bias [\text{mm}]$ .....	232
Table E.29: Effect of error in the camera parameters – Error in $o_{1,Y} = 1 \text{ mm} - MAX Bias [\text{mm}]$ .....	233
Table E.30: Effect of error in the camera parameters – Error in $o_{1,Y} = 1 \text{ mm} - RMS Bias [\text{mm}]$ .....	233
Table E.31: Effect of error in the camera parameters – Error in $o_{1,Z} = 1 \text{ mm} - MAX Bias [\text{mm}]$ .....	233
Table E.32: Effect of error in the camera parameters – Error in $o_{1,Z} = 1 \text{ mm} - RMS Bias [\text{mm}]$ .....	233
Table E.33: Effect of error in the camera parameters – Error in $\theta_{\text{cam},1} = 0.1^\circ - MAX Bias [\text{mm}]$ .....	234
Table E.34: Effect of error in the camera parameters – Error in $\theta_{\text{cam},1} = 0.1^\circ - RMS Bias [\text{mm}]$ .....	234
Table E.35: Effect of error in the camera parameters – Error in $\theta_{\text{cam},1} = 1^\circ - MAX Bias [\text{mm}]$ .....	234
Table E.36: Effect of error in the camera parameters – Error in $\theta_{\text{cam},1} = 1^\circ - RMS Bias [\text{mm}]$ .....	234
Table E.37: Effect of error in the camera parameters – Error in $\phi_{\text{cam},1} = 0.1^\circ - MAX Bias [\text{mm}]$ .....	235

Table E.38: Effect of error in the camera parameters – Error in $\varphi_{\text{cam},1} = 0.1^\circ$ – <i>RMS Bias</i> [mm].....	235
Table E.39: Effect of error in the camera parameters – Error in $\varphi_{\text{cam},1} = 1^\circ$ – <i>MAX Bias</i> [mm].....	235
Table E.40: Effect of error in the camera parameters – Error in $\varphi_{\text{cam},1} = 1^\circ$ – <i>RMS Bias</i> [mm].....	235
Table E.41: Effect of error in the camera parameters – Error in $\kappa_{\text{cam},1} = 0.1^\circ$ – <i>MAX Bias</i> [mm].....	236
Table E.42: Effect of error in the camera parameters – Error in $\kappa_{\text{cam},1} = 0.1^\circ$ – <i>RMS Bias</i> [mm].....	236
Table E.43: Effect of error in the camera parameters – Error in $\kappa_{\text{cam},1} = 1^\circ$ – <i>MAX Bias</i> [mm].....	236
Table E.44: Effect of error in the camera parameters – Error in $\kappa_{\text{cam},1} = 1^\circ$ – <i>RMS Bias</i> [mm].....	236
Table E.45: Effect of error in the camera parameters – Error in <i>focusing distance</i> <sub>1</sub> = 50 mm – <i>MAX Bias</i> [mm] .....	237
Table E.46: Effect of error in the camera parameters – Error in <i>focusing distance</i> <sub>1</sub> = 50 mm – <i>RMS Bias</i> [mm] .....	237
Table E.47: Effect of error in the camera parameters – Error in $o_Y$ for both cameras = 10 mm – <i>MAX Bias</i> [mm] .....	237
Table E.48: Effect of error in the camera parameters – Error in $o_Y$ for both cameras = 10 mm – <i>RMS Bias</i> [mm] .....	237
Table E.49: Effect of error in the camera parameters – Error in $\varphi_{\text{cam. pair}} = 1^\circ$ – <i>MAX Bias</i> [mm].....	238
Table E.50: Effect of error in the camera parameters – Error in $\varphi_{\text{cam. pair}} = 1^\circ$ – <i>RMS Bias</i> [mm].....	238

## List of Figures

Fig. 1.1: Schematic diagram of the right eye as viewed from above.....	3
Fig. 1.2: Image of the face of a 6-month old infant showing 4 corneal reflections in each eye.....	7
Fig. 1.3: (a) Bright pupil image. (b) Dark pupil image.....	8
Fig. 1.4: Cross-ratios method. (a) System setup with four light sources positioned around a computer screen, and a camera under the screen. (b) Image of the eye as captured by the camera, showing the four corneal reflections and the pupil center.....	11
Fig. 1.5: Optical system with illumination collinear with the optic axis of the camera used in the system from [49].....	12
Fig. 1.6: Schematic diagram of the cubic-foot version of the system from [49]. .....	13
Fig. 1.7: Principle of operation of the corneal reflection-pupil center vector methodology. ....	14
Fig. 2.1: Optic axis, visual axis and nodal points of the eye (right eye as viewed from above). ....	28
Fig. 2.2: Visual axis in the ECS.....	30
Fig. 2.3: Rotation of the $x_o$ - and $y_o$ - axes relative to the $x_{\text{eye}}$ - and $y_{\text{eye}}$ - axes by an angle $\delta$ .....	34
Fig. 2.4: Pupil model in the ECS. ....	35
Fig. 2.5: Perspective projection and relation between the CCS and the ICS.....	37
Fig. 2.6: Orientation of the optic axis of the eye with respect to the WCS. ....	41
Fig. 2.7: Ray-tracing diagram, showing schematic representations of the eye, a camera and a light source. ....	44
Fig. 2.8: Representation of the point of reflection $\mathbf{q}_j$ in parametric form. ....	53
Fig. 2.9: Representation of the point of refraction $\mathbf{r}_j$ in parametric form. ....	55
Fig. 4.1: Inherent point-of-gaze estimation bias of the <i>PCTriang</i> method when $o_X = 150$ mm and the pupil diameter is 8 mm.....	107
Fig. 4.2: Inherent point-of-gaze estimation bias as a function of the pupil diameter when $o_X = 150$ mm. ....	108
Fig. 4.3: Noise sensitivity of the <i>CPlIntTriang-4CR + PB</i> method .....	116
Fig. 4.4: Noise sensitivity of the <i>CNoPInt-4CR + PB</i> method.....	117

Fig. 4.5: Noise sensitivity of the different methods for the subsequent reconstruction of the optic axis of the eye with system setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) when the 3-D coordinates of the center of curvature of the cornea are calculated with the <i>CNoPlInt-4CR</i> method.....	118
Fig. 4.6: Noise sensitivity of the different methods for the subsequent reconstruction of the optic axis of the eye with system setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ , $\theta_{\text{l.s.}} = 90^\circ$ ) when the 3-D coordinates of the center of curvature of the cornea are calculated with the <i>CNoPlInt-4CR</i> method.....	119
Fig. 4.7: Effect of corneal asphericity – Bias of the <i>CPIIntTriang-4CR + PCTriang</i> method.....	126
Fig. 4.8: Effect of corneal asphericity – Bias of the <i>CPIIntInv-3CR + PCPIInt</i> method .....	127
Fig. 4.9: Effect of corneal asphericity – Bias of the <i>CNoPlInt-4CR + PB</i> method .....	128
Fig. 4.10: Effect of corneal asphericity – Bias of the <i>CNoPlInt-3CR + PB</i> method .....	129
Fig. 4.11: Effect of corneal asphericity – System with 4 light sources with $l_{1-2,Y} = 0$ mm and $o_X = 150$ mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ).....	139
Fig. 4.12: Noise sensitivity as a function of $o_X$ for the system with 4 light sources with $l_{1-2,Y} = 0$ mm .....	143
Fig. 4.13: Effect of corneal asphericity – point-of-gaze estimation bias as a function of $o_X$ for the system with 4 light sources with $l_{1-2,Y} = 0$ mm .....	143
Fig. 4.14: Total point-of-gaze estimation error as a function of $o_X$ for the system with 4 light sources with $l_{1-2,Y} = 0$ mm.....	144
Fig. 4.15: Noise sensitivity of the system with 4 light sources as a function of $l_{1-2,Y}$ when $o_X = 150$ mm .....	145
Fig. 4.16: Effect of corneal asphericity – point-of-gaze estimation bias of the system with 4 light sources as a function of $l_{1-2,Y}$ when $o_X = 150$ mm.....	146
Fig. 4.17: Total point-of-gaze estimation error of the system with 4 light sources as a function of $l_{1-2,Y}$ when $o_X = 150$ mm .....	146
Fig. 4.18: Point-of-gaze estimation results under head movements for the system with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm.....	149
Fig. 4.19: Point-of-gaze estimation bias under head movements for the system with 4 light sources with $l_{1-2,Y} = 50$ mm and $o_X = 150$ mm .....	152
Fig. 5.1: Initial system setup .....	155
Fig. 5.2: New system setup .....	155

Fig. 5.3: Top-view schematic representation of the camera calibration setup.....	156
Fig. 5.4: Experimental mean point-of-gaze estimates for Subject 1.....	164
Fig. 5.5: Experimental mean point-of-gaze estimates for Subject 1.....	165
Fig. 5.6: Experimental point-of-gaze estimates for Subject 1. ....	166
Fig. 5.7: Experimental mean point-of-gaze estimates for Subject 2.....	169
Fig. 5.8: Experimental mean point-of-gaze estimates for Subject 2.....	170
Fig. 5.9: Experimental point-of-gaze estimates for Subject 2. ....	171
Fig. 5.10: Experimental mean point-of-gaze estimates for Subject 3.....	174
Fig. 5.11: Experimental mean point-of-gaze estimates for Subject 3.....	175
Fig. 5.12: Experimental point-of-gaze estimates for Subject 3. ....	176
Fig. 5.13: Recording of the stimulus position (solid line) and the estimates of the infant's point-of-gaze (dotted line: right eye, dashed line: left eye) as a function of time for the fixation experiment. ....	177
Fig. 5.14: Optokinetic nystagmus in a 7-month old infant. ....	178
Fig. A.1: Rotation of the $X''Y''Z''$ coordinate system about the $Y''$ -axis by the pan angle $\theta_{\text{eye}}$ ..	197
Fig. A.2: Rotation of the $X'''Y'''Z'''$ coordinate system about the $X'''$ -axis by the tilt angle $\varphi_{\text{eye}}$ ..	197
Fig. A.3: Rotation of the $X^VY^VZ^V$ coordinate system about the $Z^V$ -axis by the torsion angle $\kappa_{\text{eye}}$ ..	198
Fig. C.1: Rotation of the $X'Y'Z'$ coordinate system about the $Y'$ -axis by the pan angle $\theta_{\text{cam}}$ ..	211
Fig. C.2: Rotation of the $X''Y''Z''$ coordinate system about the $X''$ -axis by the tilt angle $\varphi_{\text{cam}}$ ...	212
Fig. C.3: Rotation of the $X'''Y'''Z'''$ coordinate system about the $Z'''$ -axis by the torsion angle $\kappa_{\text{cam}}$ .....	212
Fig. E.1: Orientation of the optic axis of a camera with respect to the WCS.....	223

## List of Most Frequently Used Acronyms

<i>Acronym</i>	<i>Meaning</i>
2-D	Two-dimensional
3-D	Three-dimensional
CCD	Charge-Coupled Device
CCS	Camera Coordinate System
<i>CNoPlInt-3CR</i>	See Section 4.2.1, 5) (p. 97)
<i>CNoPlInt-4CR</i>	See Section 4.2.1, 4) (p. 97)
<i>CPlIntInv-3CR</i>	See Section 4.2.1, 2) (p. 96)
<i>CPlIntPsInv-4CR</i>	See Section 4.2.1, 1) (p. 96)
<i>CPlIntTriang-4CR</i>	See Section 4.2.1, 3) (p. 97)
ECS	Eye Coordinate System
ICS	Image Coordinate System
IR	Infrared
LCD	Liquid Crystal Display
LED	Light-Emitting Diode
MAX	Maximum
<i>PB</i>	See Section 4.2.2, 4) (p. 98)
<i>PB-EF</i>	See Section 4.2.2, 5) (p. 98)
<i>PCPlInt</i>	See Section 4.2.2, 1) (p. 98)
<i>PCTriang</i>	See Section 4.2.2, 3) (p. 98)
<i>PCTriangRefr</i>	See Section 4.2.2, 2) (p. 98)
RMS	Root-Mean-Square
WCS	World Coordinate System

# Chapter 1

## Introduction

The analysis of eye movements and visual scanning patterns is used in a large variety of applications [1] such as the study and assessment of normal and abnormal functions of the visual and oculomotor systems [2]-[9], the study and assessment of mood, perception, attention and learning disorders [10]-[12], the analysis of reading behavior [13], driving research/safety [14]-[18], pilot training [19], ergonomics [20], and marketing/advertising research [21]. The voluntary control of the eyes can serve as an input modality in multimodal human-computer interfaces [22] and assistive devices that allow severely motor-impaired persons (especially those that can neither move nor speak such as victims of severe spinal cord injuries or advanced amyotrophic lateral sclerosis) to communicate and control their immediate home environment [23]-[26], and even access the Internet and control a wheelchair.

When a person with normal vision examines an object, his/her eyes are oriented in such a way that the image of the object falls on the region of highest acuity of the retina (fovea) of each eye. The point-of-gaze (i.e., where a person is looking at) can thus be formally defined as the point within the visual field that is imaged on the center of the fovea.

Gaze estimation systems use calibration procedures to estimate a set of subject-specific parameters that are needed for the accurate calculation of the gaze direction or the point-of-gaze. These personal calibration procedures require active subject participation since subjects are required to fixate one or more specific points at specific times. Since the personal calibration requirements limit the usability of gaze estimation systems, one of the ultimate goals in the gaze estimation field is to be able to eliminate the need for personal calibration, or, at least, minimize the requirement of active subject participation. While eliminating or minimizing the personal calibration requirements can be beneficial in general, it is critical for applications in which the subjects cannot follow instructions (e.g., infants, mentally challenged individuals, non-human subjects), when subjects cannot maintain an accurate and stable fixation (e.g., Age-related Macular Degeneration), when subjects cannot fixate (e.g., oscillopsia) or when subjects are not

willing to cooperate (i.e., covert operation).

The next section presents an extensive critical review of relevant literature on remote, non-contact gaze estimation, covering the principles of operation, accuracy and personal calibration requirements.

## 1.1 Review of remote, non-contact gaze estimation methods

A good general survey of methods for measuring eye movements and estimating the point-of-gaze<sup>1</sup> is provided in [27]. More recent surveys with emphasis on remote point-of-gaze estimation methods are given in [28], [29]. A shorter, recent survey, specific to remote point-of-gaze estimation methods, can be found in [30].

The point-of-gaze can be formally defined as the point within the visual field that is imaged on the highest acuity region of the retina known as the fovea (Fig. 1.1). Alternatively, the point-of-gaze can be defined as the intersection of the visual axis with the scene. The visual axis is the line defined by the center of the fovea and the nodal point of the eye and it exhibits a subject-specific deviation from the optic axis (Fig. 1.1 –the nodal point of the eye is at the intersection of the visual axis and the optic axis, located approximately at the tip of the arrow for the crystalline lens). Ideally, the optic axis of the eye is the axis of symmetry of the eyeball and contains the centers of curvature of all the optical surfaces (front and back surfaces of the cornea, front and back surfaces of the lens). Relevant aspects of the eye anatomy are treated in more detail in Section 2.2.

Remote, non-contact gaze estimation systems are based on the analysis of eye features and, sometimes, face features extracted from video images. Video-based gaze estimation involves two major steps. The first step consists of using image processing algorithms to detect and extract the eye features / face features. The second step consists of using those features to calculate the gaze direction / point-of-gaze. This thesis is concerned with the second step.

---

<sup>1</sup> The term “eye-tracking” is often used to refer collectively to the measurement of eye movements and the determination of the gaze direction / point-of-gaze in space. In many cases, however, this term is simply used to refer to the detection and tracking of eyes in image sequences, without necessarily determining gaze.

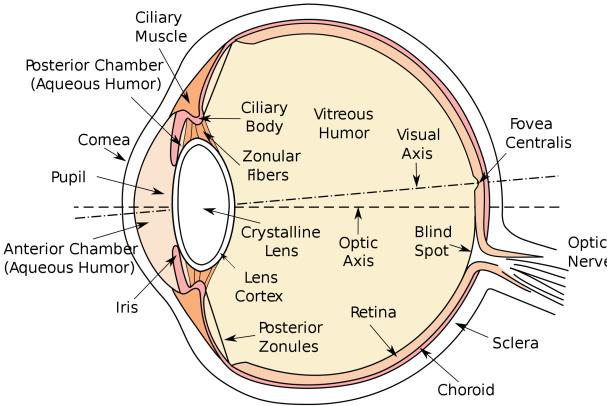


Fig. 1.1: Schematic diagram of the right eye as viewed from above.  
(Reproduced from [31] with minor changes.)

There are several different approaches to remote video-based gaze estimation. Based on the features or set of features used, video-based gaze estimation methodologies can be divided into the four major classes described in Sections 1.1.1-1.1.4.

### 1.1.1 Appearance-based / view-based

In appearance-based / view-based methods, the features used to estimate the point-of-gaze are the intensity levels of all the pixels in a rectangular image region containing the eye. Essentially, the eye image regions are treated as points in a high-dimensional space (e.g., a 20 pixel by 20 pixel intensity image region is considered as a 400-component vector or a point in a 400-dimensional space). Two examples are described briefly. In [32], the intensity levels of all the pixels in the eye image region are used as inputs to an artificial neural network (ANN). The ANN is trained using 2000 sample eye images with their corresponding coordinates of the point-of-gaze on a computer screen. The average accuracy of the point-of-gaze estimates generated by the fully trained ANN is around  $2^\circ$  of visual angle (a better accuracy of about  $1.5^\circ$  is obtained when an “offset-table” that is manually created is used to correct for the larger errors in the point-of-gaze estimates). In [33], about 250 sample images (represented as points in a 4000-dimensional space) with their corresponding coordinates of the point-of-gaze on a computer screen are used to form an “appearance manifold” for each subject. Then, given a test image, the point-of-gaze is estimated by nearest-neighbor linear interpolation. An average accuracy of about  $0.4^\circ$  of visual angle is reported for 3 subjects.

A practical disadvantage of appearance-based / view-based methods is that they need to be trained for each subject (note that the training is taking into account subject-specific eye parameters such as the deviation of the visual axis from the optic axis of the eye). The preceding examples show that substantial subject cooperation is required for the collection of the sample data (eye images with their corresponding coordinates of the point-of-gaze) during the training phase. Furthermore, to properly tolerate head movements, sample data should be collected not only for a large number of points-of-gaze but also for many different head positions. This would require a significantly larger amount of sample data and even greater subject cooperation.

### 1.1.2 Pupil/iris centers combined with 3-D head pose estimation from facial features

A good example of this class of gaze estimation methods is provided in [15], [34], [35]. Facial features are tracked in images captured by two remote cameras in order to estimate the 3-D head pose (position and orientation) using stereo matching and 3-D reconstruction techniques. From the 3-D head pose, the 3-D position of the centers of the eyeballs can be inferred and, by combining this information with the estimated positions of the pupil centers or iris centers, the position and orientation of the eyes in 3-D space can be determined. This method requires the knowledge of a set of subject-specific eye parameters (radius of the eyeball, radius of the iris), which are obtained through a personal calibration procedure [35]. Ideally, this methodology would reconstruct the optic axes of the eyes in 3-D space. The subject-specific deviation of the visual axis from the optic axis of the eye is not explicitly modeled. The point-of-gaze estimates are approximated either as the intersection of the estimated optic axes of both eyes or as the intersection of a “gaze line” with the scene. This approximate “gaze line” goes through the midpoint between the eyes (cyclopean eye) and its direction is the average of the estimated directions of the optic axes of both eyes. A point-of-gaze estimation accuracy of about  $3^\circ$  of visual angle is reported in [35]. This limited accuracy can be attributed to two main reasons: not modeling the deviation of the visual axis from the optic axis of the eye and the limited image resolution of the pupil or iris since wide field-of-view cameras are used to capture full-face images.

Methodologies that use a single camera to track facial features to estimate the 3-D head pose, and determine the position and orientation of the eyes in 3-D space using the pupil/iris centers, have also been proposed [36], [37]. However, the 3-D head pose determination using a single camera

tends to be less accurate and so is the determination of the point-of-gaze (comments in [15], [34], [35]; results in [37]).

The work described in [15], [34], [35] led to a commercial system [38]. This commercial system can provide a better accuracy of about  $1^\circ$  of visual angle after each subject completes a personal calibration procedure in which (s)he is required to fixate multiple known target points in the scene. Another commercial system that follows a similar approach in one of its operating modes is [39].

This class of gaze estimation methods is, probably, the most suitable for driving research/safety applications in which the driver's focus of attention (e.g., front, rear-view mirrors, dashboard, etc.) is to be monitored.

### 1.1.3 Limbus image back-projection

Point-of-gaze estimation methods based on the back-projection of the image of the limbus (iris-sclera boundary) captured by a single camera are described in [40]-[42]. The general idea can be summarized as follows. The limbus can be modeled as a circle. The image of the limbus on the camera image plane is an ellipse (perspective projection of a circle). The shape of the ellipse (ratio between major and minor axes, orientation of the major axis) is a function of the orientation of the plane of the limbus with respect to the camera, while the size of the ellipse is a function of the limbus radius and the distance between the limbus and the camera. Therefore, if the radius of the limbus is known, the image of the limbus can be back-projected to space and the 3-D position of the limbus center and the 3-D orientation of the normal to the plane of the limbus can be reconstructed. The back-projection of the image of the limbus yields, in general, two possible solutions for the position of the limbus center and the orientation of the limbus plane. Nevertheless, by using appropriate geometrical/anatomical constraints the spurious solution can be discarded. In this way, these methods essentially reconstruct the optic axis of the eye in 3-D space.

One limitation of these methods is that the accuracy degrades significantly when the camera has a (close to) fronto-parallel view of the limbus [40]-[42]. This is because, when the angle between the optic axis of the eye and the optic axis of the camera is small, the shape of the limbus image is very close to a circle and the sensitivity of the change of the limbus image shape with respect

to eye orientation is significantly reduced. Another key issue with these methods is that, in general, the top part and/or the bottom part of the iris are occluded by the eyelids and only the lateral edges between the iris and the sclera can be used to fit an ellipse and apply the methodology described in the previous paragraph. In order to be able to fit an ellipse reliably in such conditions, a high-resolution image of the iris is required. This issue has been addressed by using two remote cameras: a “head-pose” camera that has a wide field-of-view, somewhat larger than the whole face, and a “gaze” camera mounted on a motorized pan and tilt unit that has a narrow field-of-view and takes zoom-in images of both eyes [40] or only one eye [41], [42]. The “head-pose” camera is used to track facial features and estimate the 3-D head pose. The information about the head pose is used to aim the “gaze” camera at the eye(s).

In order to reconstruct the optic axis of the eye, the radius of the limbus has to be calibrated for each subject. To estimate the point-of-gaze, the deviation between the visual and optic axes of the eye must be also calibrated for each subject. Although the need for calibration is briefly acknowledged in [40]-[42], the calibration procedure is not described. To obtain those eye parameters, a calibration procedure in which the subject has to fixate at least two known target points should be performed. Two calibration points is the theoretical minimum and more points should be used in practice to obtain a reasonably robust calibration.

Point-of-gaze estimation errors of  $1^\circ$  of visual angle or less are reported in [40]-[42] (presumably after personal calibration, and as long as (close to) fronto-parallel views of the eye are avoided).

#### 1.1.4 Pupil and corneal reflections

The pupil and corneal reflections have been used for gaze estimation for over 45 years [43]. They are, in fact, the most commonly used eye features in commercial remote gaze estimation systems (e.g., [39], [44]-[47], among many others). The corneal reflections (first Purkinje images, glints) are virtual images of near-infrared light sources that illuminate the eye. These virtual images are formed by reflection on the front surface of the cornea, which acts as a convex mirror. The corneal reflections appear as bright spots as shown in Fig. 1.2. The main reason to use near-infrared illumination is that it is invisible to the human eye.



Fig. 1.2: Image of the face of a 6-month old infant showing 4 corneal reflections in each eye. The pupil centers are indicated by crosses and the corneal reflections are enclosed by squares.

The appearance of the pupil depends on the location of the illumination source(s) relative to the video camera. When there is a near-infrared illumination source on the optic axis of the camera (“on-axis illumination”)<sup>2</sup> or very close to it, the light that enters the eye is reflected on the retina back to the camera, making the pupil appear brighter than the surrounding iris. This is called the “bright pupil effect” (Fig. 1.3(a)), which is similar to the “red eye effect” that is observed in photographs taken with a flash. The intensity of the bright pupil is highly dependent on the pupil diameter [48] since the energy of the light that reaches the retina is proportional to the square of

---

<sup>2</sup> Most commonly, this is achieved by using a near-infrared light emitting diode (NIR LED) located on the optic axis of the camera, in front of and very close to the lens of the camera, or using a ring of NIR LEDs around the lens of the camera.

the pupil diameter. For this reason, when the pupil diameter is small (less than 4 mm) the contrast between the bright pupil and the iris is poor, making the extraction of the pupil from the eye image very difficult or impossible. Consequently, systems that use the bright pupil effect have to be used in dim ambient light conditions [23], [49]. The intensity of the bright pupil also depends on the gaze direction with respect to the camera/illumination [48], [50] since the reflectivity of the retina is not uniform. In addition, the intensity of the bright pupil varies with the ethnicity of the subjects [48]. Experience has shown that in about 5 to 10 % of the population the near-infrared reflectivity of the retina is not enough for the bright pupil effect to work properly [23], [51].

When the near-infrared illumination is away from the optic axis of the camera (“off-axis illumination”), the pupil appears as a dark circle against the surrounding iris since the light reflected by the retina is not collected back by the camera. This is called the “dark pupil effect” (Fig. 1.3(b) –see also Fig. 1.2), which is not affected by the pupil diameter and therefore it can be used in a larger variety of applications and environments.



Fig. 1.3: (a) Bright pupil image. (b) Dark pupil image.  
The brightest spots in the images are corneal reflections.  
(Courtesy of Prof. Carlos H. Morimoto.)

The combination of the bright pupil and dark pupil effects has been proposed to improve the detection of the pupil. In the method described in [52]-[55], on-axis illumination (bright pupil) is used for every other video field/frame and off-axis illumination (dark pupil) is used for the rest of the video fields/frames. The difference between consecutive bright and dark pupil images (“bright-dark pupil difference method”) removes (or minimizes) the effect of ambient infrared illumination and emphasizes the pupil against the background, thereby improving the detection of the pupil and the extraction of the pupil-iris boundary. A drawback of this technique is that it

can result in pupil extraction errors when the eye moves between consecutive video fields/frames. Methods that, at the cost of increased system complexity, overcome this problem by obtaining bright and dark pupil images simultaneously are described in [51] and [56].

The imaging system described in [51] uses a dual camera system with an illumination source consisting of a ring of LEDs around the common collection aperture. In the ring of LEDs, the inner LEDs have a nominal wavelength of 850 nm and the outer LEDs have a nominal wavelength of 950 nm. The dual camera system consists of two synchronized charge-coupled device (CCD) cameras with 850 nm and 950 nm optical filters and a di-chromic beam splitter. The 850 nm image has bright pupils while the 950 nm images have dark pupils. The difference between these two simultaneously captured images emphasizes the pupil against the background.

The system described in [56] is even more complex as it is meant to improve the detection of not only the pupil but also the corneal reflection. That system uses simultaneous polarized off-axis illumination of nominal wavelength of 850 nm ( $\lambda_1$ ) and on-axis illumination of nominal wavelength of 950 nm ( $\lambda_2$ ), together with a specially designed camera with 3 CCD image sensors. Using an appropriate optical construct, light at wavelength  $\lambda_2$  is directed to  $CCD_3$  while light at wavelength  $\lambda_1$  is split between  $CCD_1$  and  $CCD_2$ . Part of the light at wavelength  $\lambda_1$  enters  $CCD_1$  through a polarizer (perpendicular to the polarizer used for the illumination) while the other part enters  $CCD_2$  directly. As a result,  $CCD_1$  captures a dark pupil image without corneal reflections,  $CCD_2$  captures a dark pupil image with the corneal reflection due to the off-axis illumination, and  $CCD_3$  captures a bright pupil image with the corneal reflection due to the on-axis illumination. The difference between the images from  $CCD_3$  and  $CCD_1$  emphasizes the pupil against the background, while the difference between the images from  $CCD_2$  and  $CCD_1$  isolates the corneal reflection produced by the off-axis illumination.

Beyond the appearance of the pupil and the extraction of the pupil and corneal reflection(s), there is a large variety of gaze estimation methodologies that use these eye features. The different methodologies can be grouped into three major classes according to the way in which the pupil and corneal reflections are used to estimate gaze. These major methodology classes are described in Sections 1.1.4.1-1.1.4.3.

### 1.1.4.1 “Cross-ratios” methodologies

The “cross-ratios” method for point-of-gaze estimation was originally presented in [57] as tolerating natural head movements and not requiring a personal calibration procedure, among other desired attributes. The work in [57] proposes a mapping of the pupil center in an eye image to a corresponding point-of-gaze estimate in a plane by exploiting the invariance of cross-ratios in projective geometry. To estimate the point-of-gaze in a plane, four light sources are placed in that plane and the subject’s eye is imaged using a video camera. Fig. 1.4(a) shows a typical system setup in which the four light sources ( $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4$ ) are placed around a computer screen (the scene plane) upon which the point-of-gaze is estimated ( $\hat{\mathbf{g}}$ ), and the camera is placed under the screen. Fig. 1.4(b) shows an image of the eye as captured by the camera, in which the four corneal reflections ( $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ ) and the pupil center,  $\mathbf{u}_p$ , are identified. The work in [57] suggests that the cross-ratios formed by features in the scene plane ( $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4, \hat{\mathbf{g}}$ ) are equal to the cross-ratios formed by the corresponding features in the camera’s imaging plane ( $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_p$ ). In other words, it suggests that the coordinates of the four corneal reflections ( $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ ) are related to the coordinates of the four light sources ( $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4$ ) by a projective transformation, and that the coordinates of the pupil center,  $\mathbf{u}_p$ , are related to the coordinates of the point-of-gaze by the same projective transformation. However, this is incorrect, resulting in very large point-of-gaze estimation bias for some subjects [58]-[62]<sup>3</sup>. A detailed study of the main sources of the point-of-gaze estimation bias of the cross-ratios method is presented in [61], [62].

In order to reduce the estimation bias, error compensation techniques using a personal calibration procedure in which the subject has to fixate multiple known target points have been proposed [58]-[60]. With the error compensation techniques presented in [59] and [60], point-of-gaze estimation accuracies of about  $1^\circ$  of visual angle can be obtained under small head movements (an analytical justification of the empirically derived error compensation technique presented in [59] is provided in [62]). Under larger head movements the performance degrades gradually, mainly as the distance between the eye and the system changes with respect to the eye-system distance at which the personal calibration procedure was carried out.

<sup>3</sup> References [60]-[62] present work done in collaboration with other students in our research lab.

The systems described in [57]-[59] have a fifth, on-axis light source in order to use the bright-dark pupil difference method described in [52]-[55] (see beginning of Section 1.1.4). The system described in [60]-[62] uses the dark pupil effect.

Another limitation of the cross-ratios method is that it requires the simultaneous formation of four corneal reflections for the estimation of the point-of-gaze. This is not always possible as one or more of the reflections of the light sources can fall on the sclera rather than on the cornea when subjects gaze at the periphery of the screen, unless the distance between the eye and the system is large relative to the computer screen size (e.g., for a 19" screen, the distance between the eye and the system has to be about 1 m [60]-[62]). This physical constraint limits the utility of the cross-ratios method.

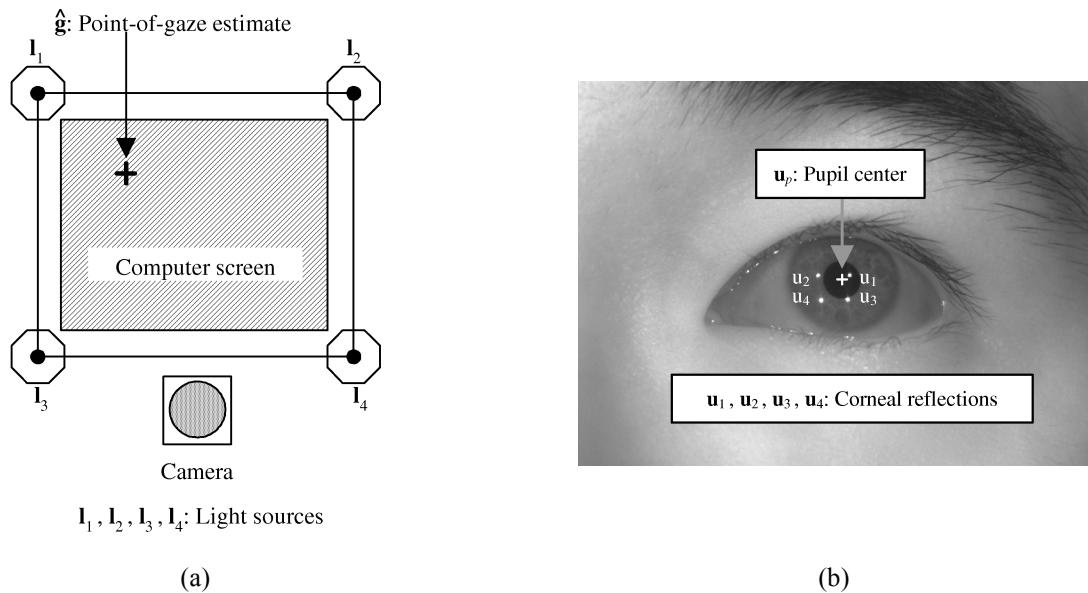


Fig. 1.4: Cross-ratios method. (a) System setup with four light sources positioned around a computer screen, and a camera under the screen. (b) Image of the eye as captured by the camera, showing the four corneal reflections and the pupil center.

#### 1.1.4.2 Corneal reflection-pupil center vector

This methodology class was predominant from the early 70's to the late 90's. Because of its importance in the field, it is described here in detail. First, pioneering work using the corneal reflection-pupil center vector is presented in Section 1.1.4.2.1 and used to explain the principles of operation of this methodology class. Next, the behavior of the corneal reflection-pupil center vector in the presence of head movements is discussed in Section 1.1.4.2.2. Finally, several subsequent works in this methodology class are described in Section 1.1.4.2.3.

### 1.1.4.2.1 Pioneering work - Principles of operation

Pioneering work on video-based remote gaze estimation using the pupil center and one corneal reflection that was done in the early 70's is described in [49]. This was one of the first fully operational remote, non-contact gaze estimation systems and had an accuracy of about  $1^\circ$  of visual angle. In that system, a vidicon TV camera with a telephoto lens and a near-infrared filter was used to obtain images of one eye. The near-infrared illumination was designed so that the illumination axis was coincident with the optic axis of the camera (Fig. 1.5), resulting in a bright pupil. The field of view of the camera was just slightly larger than an eye in order to be able to estimate the centers of the pupil and the corneal reflection as accurately as possible. Two versions of the system were described: one that allowed for eye translation in a volume of one cubic inch and one that allowed for eye translation in a volume of one cubic foot. The cubic-foot version incorporated a two-axis moving mirror system to keep the eye within the field of view of the camera as the head moved, and a servo-controlled focusing lens to keep the image of the eye in focus on the camera sensor (Fig. 1.6). The autofocus lens system worked by determining the amount of blur of the pupil-iris edge.

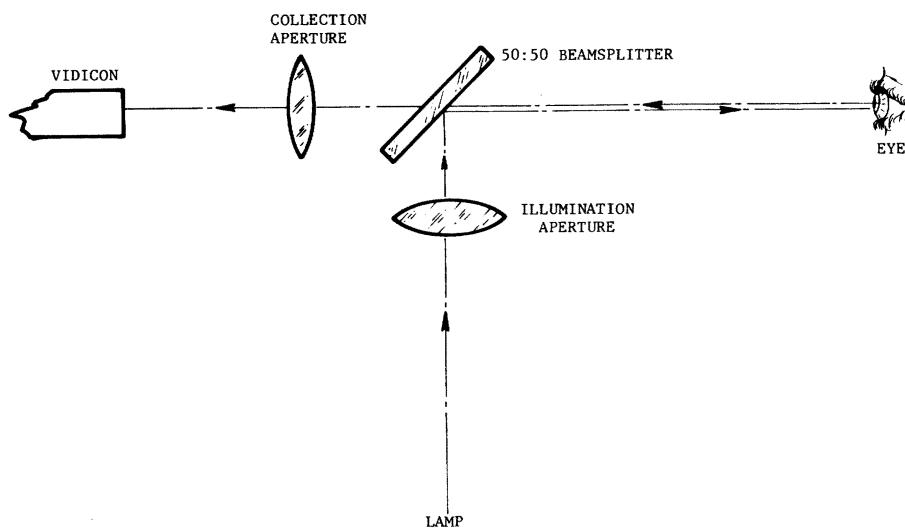


Fig. 1.5: Optical system with illumination collinear with the optic axis of the camera used in the system from [49].  
 (© 1974 IEEE. Reproduced from [49] by permission.)

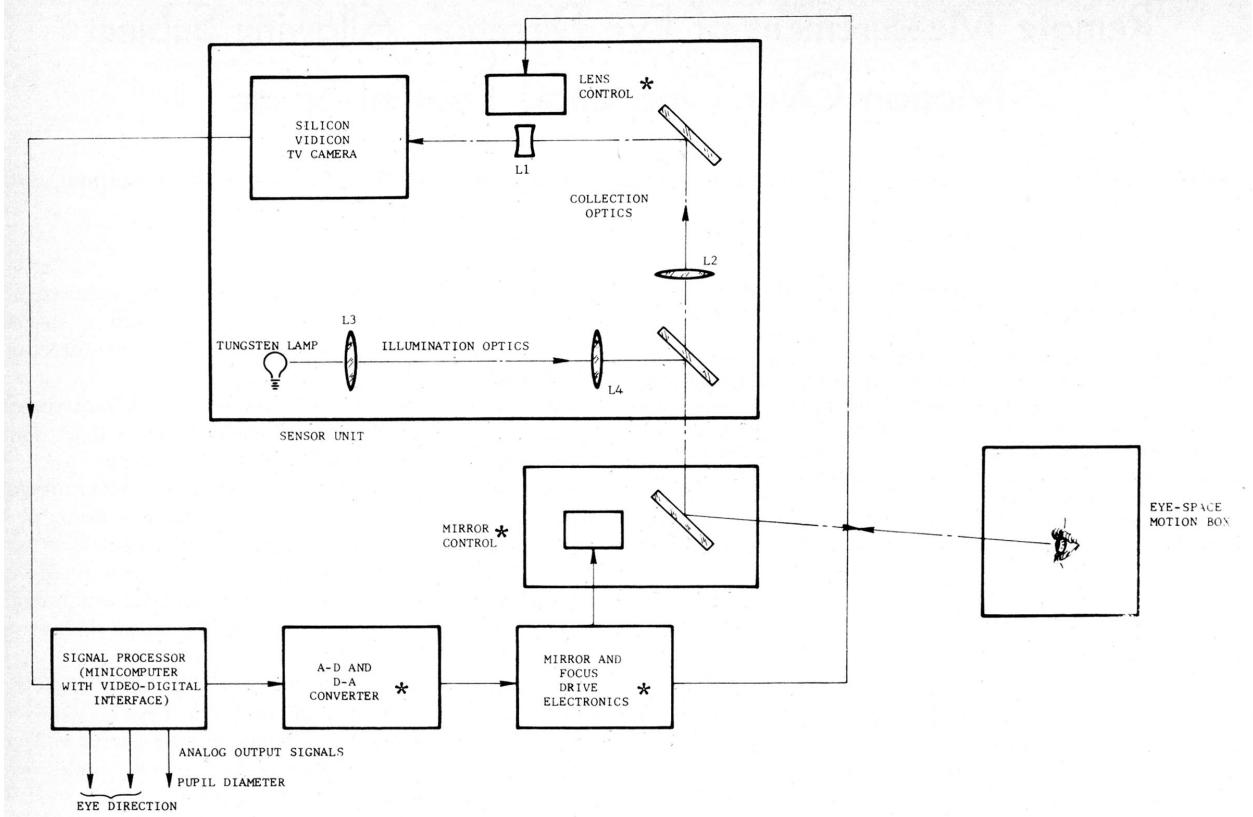


Fig. 1.6: Schematic diagram of the cubic-foot version of the system from [49].

Items indicated by '\*' are not part of the cubic-inch version.

(© 1974 IEEE. Reproduced from [49] by permission.)

The basic gaze estimation principle can be described with the aid of Fig. 1.7. As explained before, the reflection of the illumination source on the front surface of the cornea results in the formation of a virtual image of the source referred to as corneal reflection. If the front surface of the cornea is assumed to behave as a convex spherical mirror and the illumination is collimated (at optical infinity [63]), the virtual image of the illumination source is located halfway between the center of curvature and the front surface of the cornea, in the line that is parallel to the illumination axis / optic axis of the camera and passes through the center of curvature of the cornea. The distance between the center of the pupil and that line is equal to the product of the distance from the center of the pupil to the center of curvature of the cornea ( $K$ ) and the sine of the angle between the optic axis of the eye and the optic axis of the camera ( $\theta$ ), as indicated in Fig. 1.7. In reality, due to the refraction at the cornea, the camera observes a virtual image of the pupil (entrance pupil) rather than the true pupil, and therefore, it is more appropriate to refer to the center of the entrance pupil and  $K$  should represent the distance from the center of the entrance pupil to the center of curvature of the cornea [64].

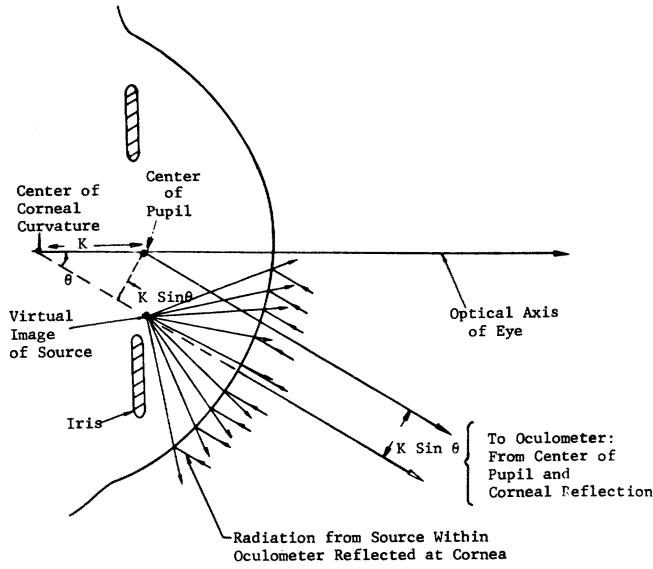


Fig. 1.7: Principle of operation of the corneal reflection-pupil center vector methodology.  
 (© 1974 IEEE. Reproduced from [49] by permission.)

The image of the corneal reflection and the image of the pupil center in the camera image plane define a vector that is hereafter referred to as the corneal reflection-pupil center vector. The direction of this vector indicates the direction of eye rotation. The magnitude of the corneal reflection-pupil center vector is

$$mK \sin \theta, \quad (1.1)$$

where  $m$  is the image magnification,  $K$  is the distance from the center of the entrance pupil to the center of curvature of the cornea, and  $\theta$  is the angle between the optic axis of the eye and the optic axis of the camera. The parameter  $m$  depends on the optical characteristics of the imaging system and the distance between the eye and the camera. To a good approximation,  $m$  is inversely proportional to the distance between the eye and the camera (as the eye gets closer to the camera, the magnitude of the corneal reflection-pupil center vector increases even when the angle  $\theta$  stays constant). The parameter  $K$  varies from subject to subject.

To estimate the gaze direction relative to the optic axis of the camera, the angular deviation of the visual axis from the optic axis of the eye must be taken into account. To do this, the following equation is used in [49] (only equation in the article):

$$\begin{aligned}\theta_x &= X_0 + J_1 X_{\text{REL}} \\ \theta_y &= Y_0 + J_2 Y_{\text{REL}} ,\end{aligned}\tag{1.2}$$

where  $X_{\text{REL}}$  and  $Y_{\text{REL}}$  are the components of the corneal reflection-pupil center vector,  $J_1$  and  $J_2$  are scale factors such that  $J_1 X_{\text{REL}}$  and  $J_2 Y_{\text{REL}}$  represent the orientation of the optic axis of the eye relative to the optic axis of the camera,  $X_0$  and  $Y_0$  account for the deviation of the visual axis from the optic axis of the eye, and  $\theta_x$  and  $\theta_y$  indicate the orientation of the visual axis of the eye (gaze direction) relative to the optic axis of the camera. This expression is only valid if the distance between the eye and the camera is constant or if the distance between the eye and the camera is measured (in this case, from the focusing distance of the autofocus lens system) and the terms  $J_1 X_{\text{REL}}$  and  $J_2 Y_{\text{REL}}$  are scaled accordingly (e.g.,  $X_{\text{REL}}$  and  $Y_{\text{REL}}$  are normalized by the image magnification). In addition, (1.2) is a reasonable approximation only for angles up to 10-20°. To maintain accuracy for an extended angular range, a polynomial equation is used to compensate for errors due to modeling assumptions and approximations. The polynomial equation compensates not only for the approximation  $\sin \theta \approx \theta$  used in (1.2), but also for the lack of modeling of the refraction at the cornea (in strict terms, the distance from the center of the entrance pupil to the center of curvature of the cornea,  $K$ , is a function of the orientation of the eye relative to the camera and the pupil diameter) and for corneal asphericity (deviation of the shape of real corneas from the ideal spherical shape –i.e., the cornea flattens towards the boundary with the sclera). The coefficients of (1.2) ( $J_1$ ,  $J_2$ ,  $X_0$  and  $Y_0$ ) and the coefficients of the polynomial equation are obtained through a personal calibration procedure in which the subject has to fixate a sequence of known target points (typically 28 points arranged in a 4 x 7 grid) for which  $\theta_x$  and  $\theta_y$  are known. The coefficients are optimized to minimize the root-mean-square (RMS) error between the estimated and the actual gaze directions (the coefficients of (1.2) would be optimized in a first stage and the coefficients of the polynomial equation would be optimized in a second stage).

To estimate the point-of-gaze in 3-D space with respect to a world coordinate system (WCS), the position of the eye and the gaze direction with respect to the WCS must be known. For the system in [49], the position of the eye with respect to the WCS can be determined as a function of the orientation of the mirrors that are used to keep the eye in the field of view of the camera when the head moves, and the distance between the eye and the camera that is obtained from the

autofocus system. The gaze direction with respect to the WCS can be determined as a function of the gaze direction relative to the optic axis of the camera and the orientation of the mirrors (the optic axis of the camera changes its direction at the mirrors). In summary, if the eye is allowed to translate, the distance between the eye and the camera and the orientation of the tracking mirrors must be taken into account to be able to estimate the point-of-gaze with respect to the WCS. Even though this is recognized in [49], no equations are provided (note that the system described in [49] was a commercial system).

#### 1.1.4.2.2 Behavior of the corneal reflection-pupil center vector in the presence of head movements

From the preceding presentation, it is clear that if the center of rotation of the eye is stationary (e.g., the head does not move) and the system components are also stationary, then there exists a one-to-one correspondence between the corneal reflection-pupil center vector and the point-of-gaze on a planar surface such as a computer or projection screen since a change in the point-of-gaze is obtained by eye rotation only. On the other hand, when the eye translates, the corneal reflection-pupil center vector by itself does not provide enough information to determine the point-of-gaze. In fact, there are infinite possible corneal reflection-pupil center vectors for a fixed point-of-gaze, and infinite points-of-gaze can be associated with a given corneal reflection-pupil center vector. To illustrate this, consider the following two examples in which the eye translates along the optic axis of the camera. For the first example, suppose that the point-of-gaze is kept constant at a certain distance from the optic axis of the camera. As the distance between the eye and the camera changes, the gaze direction relative to the camera has to change accordingly to maintain the point-of-gaze (e.g., as the distance between the eye and the camera decreases, the angle between the optic axis of the eye and the optic axis of the camera has to increase). Therefore, the magnitude of the corneal reflection-pupil center vector changes not only due to the change of the camera magnification but also due to the change of the angle between the optic axis of the eye and the optic axis of the camera (the magnitude of the corneal reflection-pupil center vector for a constant point-of-gaze is roughly inversely proportional to the square of the distance between the eye and the camera). This shows that there are infinite possible corneal reflection-pupil center vectors for a fixed point-of-gaze. For the second example, it is easy to understand that it is always possible to modify the point-of-gaze as the distance between the eye and the camera changes in such a way that the corneal reflection-pupil center vector remains

constant. This shows that infinite points-of-gaze can be associated with a given corneal reflection-pupil center vector. Similar examples can be given for lateral eye translation. Of course, for the system described in [49], given the corneal reflection-pupil center vector, the distance between the eye and the camera and the orientation of the tracking mirrors, there is a unique possible point-of-gaze on a planar surface.

#### 1.1.4.2.3 Subsequent works

It follows from the previous discussion that when the center of rotation of the eye and the system components are stationary, it is then possible to determine a function that maps the corneal reflection-pupil center vector to the point-of-gaze on a plane (2-D mapping function) [23], [28], [65]-[67]<sup>4</sup>. These 2-D mapping functions are most commonly polynomial functions of the coordinates of the corneal reflection-pupil center vector (in [23] a linear function is suggested; in [28] and [66] a second-degree polynomial is used; in [65] both linear functions and second-degree polynomials are tested; in [67] polynomials of first to fourth degree are studied). The coefficients of the 2-D mapping function are obtained through a personal calibration procedure in which the subject has to fixate multiple known target points in the scene plane. Of course, if the eye translates relative to the position adopted during the calibration, large point-of-gaze estimation errors can occur unless translation is limited to a few millimeters or the coefficients of the 2-D mapping function are recalibrated. A quantification of the point-of-gaze estimation errors due to eye translation is provided in [28]. A technique that automatically readjusts the mapping function to compensate, to some extent, for lateral head movements, but does not compensate for changes in the distance between the eye and the camera, is presented in [65].

The system described in [68] represents an attempt to estimate the point-of-gaze on a computer screen in the presence of head movements using a mapping function. It uses a relatively narrow field-of-view camera (“gaze camera”) with a mirror galvanometer system, and a stereo pair of fixed wide field-of-view cameras with near-infrared ring light sources around their lenses. The wide field-of-view cameras are used to locate the eyes and determine the pan and tilt angles of the mirror galvanometer system to keep the eye being tracked within the field-of-view of the

---

<sup>4</sup> The systems described in [23] and [65] use the bright pupil effect, the system described in [66] uses the bright-dark pupil difference method described in [52]-[55] (see beginning of Section 1.1.4), whereas [28] and [67] present simulation studies.

gaze camera as the head moves. The centers of the pupil and the two corneal reflections produced by the light sources mounted on the wide field-of-view cameras are extracted from the images captured by the gaze camera and used to estimate the point-of-gaze. The point-of-gaze is estimated using a third-degree polynomial function of the pan and tilt angles of the mirror galvanometer system, the distance between the two corneal reflections and the vector from the midpoint between the two corneal reflections to the pupil center. The coefficients of this polynomial function are obtained through a personal calibration procedure in which each subject is instructed to fixate a sequence of known target points on the computer screen ( $4 \times 4$  grid). For each target point, the subject has to maintain fixation while moving the head about slowly in all three dimensions. This requirement for the calibration can be as inconvenient as requiring the subject to maintain the head still. Although a median error of about  $0.8^\circ$  of visual angle is reported from experiments with three subjects, several point-of-gaze estimates with errors of as much as  $8-9^\circ$  of visual angle are observed in the graphs provided. It is difficult to know how much of these point-of-gaze estimation errors are due to errors in the extraction of the eye feature coordinates from the images and how much is due to modeling / methodology errors.

The recent study presented in [67] considers a variety of polynomial mapping functions using the corneal reflection-pupil center vector and variants of it. Variants include the use of the average of two corneal reflection-pupil center vectors obtained with two light sources and the normalization of this average corneal reflection-pupil center vector by the distance between the two corneal reflections. By normalizing the corneal reflection-pupil center vector by the distance between the two corneal reflections, the effect of the variation of the distance between the eye and the camera is reduced to some extent.

Systems that, similarly to [49], use some explicit information about the 3-D position of the eye to estimate the point-of-gaze have, in general, significantly better performance under head movements. Some examples are found in [18], [69]-[73]. In order to capture high resolution images of one eye while allowing for head movements, the systems described in [18], [69], [71] and [72] use a narrow field-of-view camera (gaze camera) mounted on a motorized pan and tilt unit, while the system described in [70] uses a fixed narrow field-of-view camera and a mirror mounted on a motorized pan and tilt unit.

The camera in the system described in [18] uses an autofocus lens system. A near-infrared light emitting diode (LED) is placed on the optic axis of the camera, in front of the lens, very close to the lens (bright pupil effect). The autofocus system uses an asymmetric aperture built into the camera lens, which causes the image of the corneal reflection to vary in size and orientation as the eye moves forward and backward. The magnitude of the corneal reflection blur provides information about how far out-of-focus the camera is, and the orientation of the corneal reflection image indicates whether the camera is focused too near or too far. The focusing distance of the autofocus lens system is used as an estimate of the distance between the eye and the camera. Note that equations are not provided in [18] and that these concepts are implemented in commercial systems [44].

The camera in the systems described in [69] and [70] uses a lens with motorized focus and zoom mechanisms. A near-infrared LED is placed on the optic axis of the camera, in front of the lens, very close to the lens (a second near-infrared LED is located off axis and is used for pupil detection only using the bright-dark pupil difference method described in [52]-[55] and summarized at the beginning of Section 1.1.4). An ultrasonic distance meter is used to measure the distance between the eye and the camera. The measurements from the ultrasonic distance meter are used to control the focus and zoom of the camera lens [74], and for the subsequent calculation of the point-of-gaze. By adjusting the zoom so that the image magnification (factor  $m$  in (1.1)) stays constant, the magnitude of the corneal reflection-pupil center vector varies as a function of the orientation of the eye relative to the camera only, independently of the distance between the eye and the camera. An average error of about 0.7° of visual angle is reported in [69].

The systems described in [71] and [72] use an illumination source consisting of an array of near-infrared LED's attached to the side of the gaze camera rather than on the optic axis of the camera (dark pupil effect). In [71] two additional fixed cameras with a wide field-of-view (larger than the whole face) are used to track both eyes and estimate the 3-D position of the pupil centers with respect to a WCS (attached to a computer display) using 3-D stereo matching / triangulation. In [72] only one additional fixed camera with a wide field-of-view is used to track both eyes and estimate the 3-D position of the pupil center of one of them. In this case, the calculation of the 3-D position of the pupil center assumes that the inter-ocular distance is known and that the line connecting both pupil centers is parallel to the scene plane [75]. In either case,

the estimated 3-D position of the pupil center of the eye being tracked is used to control the pan and tilt angles of the gaze camera, and for the subsequent calculation of the point-of-gaze. A point-of-gaze estimation accuracy of 0.4° of visual angle is reported in [72].

The system described in [73] uses two fixed cameras (stereo camera pair) with a relatively wide field-of-view. A ring light source is placed around the lens of each camera to mimic a light source located at the center of the lens (this system uses the bright-dark pupil difference method described in [52]-[55] and summarized at the beginning of Section 1.1.4). In one of the methodologies presented in this article, a 2-D mapping from a corneal reflection-pupil center vector extracted from one of the cameras to the point-of-gaze in a computer display is calibrated for a fixed head position. Using the pupil centers extracted from the eye images captured by both cameras, the 3-D position of the center of the entrance pupil (virtual image of the pupil) is obtained approximately through 3-D triangulation and used as input to a head motion compensation function.

The methodology from [71] and the methodology from [73] described in the previous paragraph are limited to the estimation of the point-of-gaze in a plane, and can therefore be classified as “2-D mapping methods with 3-D head motion compensation”. On the other hand, the methodologies from [69], [70] and [72] reconstruct the visual axis of the eye in 3-D space and therefore can estimate the point-of-gaze in a 3-D scene.

It is clear from the discussions in this section that the estimation of the point-of-gaze using the corneal reflection-pupil center vector in the presence of head movements is associated with a significant system complexity (i.e., one camera plus an autofocus lens system or an ultrasonic distance meter, or multiple cameras) while using a personal calibration procedure in which the subject is required to fixate a sequence of multiple known target points (typically, 9 or more points). The gaze estimation methodology class described in the next section represents a significant improvement in terms of system complexity and/or personal calibration requirements.

#### 1.1.4.3 Direct model-based 3-D gaze reconstruction

Direct model-based 3-D gaze reconstruction methodologies use integrated mathematical models of the system and the eye to directly determine the position and orientation of the eye in space and, therefore, they can estimate the point-of-gaze in a 3-D scene while inherently tolerating

head movements. The mathematical models are based on the fundamental laws of geometrical optics (laws of reflection and refraction).

Although some preliminary ideas for direct model-based 3-D gaze reconstruction were suggested in [56] twenty years ago, the work in this methodology class that was available in the literature at the start of my Ph.D. program in May 2003 was limited [76]-[78].

The work in [76] had shown that using at least two cameras and at least two light sources it is theoretically possible to reconstruct the optic axis of the eye in 3-D space without knowing any subject-specific eye parameter. However, this work did not consider the fact that the point-of-gaze is located on the visual axis, which exhibits a subject-specific deviation from the optic axis of the eye, and did not provide experimental results. The simulation results included in [76] were limited to the quantification of the combined effect of errors in the estimation of the system parameters and errors in the estimation of the image coordinates of the eye features. These simulation results did not provide insight into the contribution of each source of error to the total error, the role of the system setup (relative positions of the cameras and light sources) in the methodology sensitivities, and the effect of the deviation of the shape of real corneas from the ideal spherical shape assumed in the model (corneal asphericity).

The work in [77] had shown the ability to estimate the point-of-gaze in the presence of head movements using one camera with an autofocus lens and one light source, after completing a personal calibration procedure in which the subject was required to fixate at least two target calibration points. Note that the focusing distance of the autofocus lens system was used as a parameter for the calculation of the point-of-gaze.

In my M.A.Sc. thesis [78], I had shown the ability to estimate the point-of-gaze in 3-D space in the presence of head movements with one fixed camera with a fixed lens and two light sources, using only the pupil and the corneal reflections extracted from the eye images. Similarly to the system described in [77], each subject had to complete a one-time personal calibration procedure in which (s)he was required to fixate multiple (at least two) target calibration points. The work in my M.A.Sc. thesis included a detailed analysis of the effects of errors in system parameters, errors in the estimation of the image coordinates of the eye features and corneal asphericity. A way to reduce the effect of corneal asphericity was also proposed and evaluated experimentally.

The state of the field at the start of my Ph.D. program, as described in the preceding paragraphs, led to the research goals summarized in the next section. During my Ph.D. program there was a significant growth in the amount of published work on direct model-based 3-D gaze reconstruction, with several research groups doing similar work simultaneously [73], [79]-[95]<sup>5</sup> ([83]-[87] include work derived from my M.A.Sc. and/or Ph.D. programs). References to these works are provided when appropriate throughout the detailed discussions in this thesis. Through those discussions, the advantages of direct model-based 3-D gaze reconstruction methodologies will become clear.

## 1.2 Goals

Given the state-of-the-art at the start of my Ph.D. program, the *general* goal of this thesis was set to develop accurate, remote, non-contact gaze estimation methodologies that tolerate natural head movements and require minimal subject cooperation to *enable* applications that were very difficult or impossible to carry out with previously available gaze estimation technologies. Towards this end, a series of more specific goals were set:

- The *first* goal was to develop a general mathematical model for gaze estimation systems that use the pupil and one or more corneal reflections extracted from images captured by one or more video cameras.
- The *second* goal was to create a simulation framework based on the general model that would serve as an essential tool to quantify model sensitivities and optimize system setups.
- The *third* goal was to use the general model to carry out a detailed analytical study of the point-of-gaze estimation problem for the full range of possible system configurations, from the simplest, that includes one camera and one light source, to the most complex, that include multiple cameras and multiple light sources. The purpose of this analysis was to provide fundamental insight into the limitations and potential of each configuration, including the

---

<sup>5</sup>The systems described in [77], [78], [80]-[88], [90]- [92] and [94] use the dark pupil effect, the system described in [79] uses the bright pupil effect, the systems described in [73], [89] and [95] use the bright-dark pupil difference method described in [52]-[55] (see beginning of Section 1.1.4), whereas [76] and [93] are simulation works.

personal calibration requirements. For each system configuration, different methods to calculate the point-of-gaze in 3-D space from the principles of geometrical optics represented by the general model would be considered since different methods could exhibit significantly different performances in non-ideal conditions (e.g., noise in the image coordinates of the eye features, corneal asphericity, errors in system parameters).

The analysis of the different system configurations would confirm that two cameras and two light sources is the simplest system configuration that can be used to reconstruct the optic axis of the eye in 3-D space from the pupil and corneal reflections without knowing any subject-specific eye parameter [76]. This analysis would also indicate that in order to estimate the point-of-gaze, which is located on the visual axis, a single-point personal calibration procedure would be needed to determine the angular deviation between the optic and visual axes.

- The *fourth* goal was then to use the simulation framework to evaluate the expected performance of different methods for point-of-gaze estimation using two cameras and at least two light sources as a function of the system setup (relative positions of the cameras and light sources, number of light sources). This simulation study would consider the effects of errors in the estimation of the image coordinates of the eye features, errors in the estimation of the system parameters and corneal asphericity. The results from this study would provide guidelines for optimal system implementation.
- The *fifth* goal was to implement a prototype system based on the guidelines from the simulation study to verify the concepts developed in this thesis. The prototype system would be tested with both adults and infants. The results with adults would provide an estimate of the expected system accuracy, while the results with infants would demonstrate the ability of a system that is based on the proposed methodologies to monitor infants' visual scanning behavior.

### 1.3 Organization of this thesis

Chapter 2 presents a detailed mathematical model for video-based gaze estimation systems that use the pupil and one or more corneal reflections extracted from images captured by one or more

video cameras. This mathematical model is based on the principles of geometrical optics (laws of reflection and refraction). Given the geometrical setup of a gaze estimation system (position of light sources, camera parameters), the mathematical model relates the position and orientation of the eye (or the position of the eye and the point-of-gaze) in 3-D space with the images of the pupil and corneal reflections captured by the camera(s). The mathematical model is used not only to estimate gaze in 3-D space but also as the basis of a simulation framework to generate synthetic eye feature data. The generation of synthetic eye feature data consists of determining the pupil boundary, the pupil center and the centers of the corneal reflections (eye features) in the images captured by the camera(s), given a set of eye parameters, the system parameters (position of the light sources, camera parameters) and the 3-D positions of the center of the eyeball and the point-of-gaze. The simulation framework is an invaluable tool that supports studies of the sensitivity of the estimation of the point-of-gaze to several different types of error such as errors in the estimation of the coordinates of the eye features in each image of the eye, the deviation of the actual corneal shape from the ideal spherical shape assumed for the estimation of the point-of-gaze (corneal asphericity), and errors in the system parameters. In order to study the effect of corneal asphericity, an ellipsoidal corneal model is developed in detail.

Chapter 3 provides a detailed analysis of the point-of-gaze estimation problem using a spherical corneal model. The analysis covers the full range of possible system configurations, from the simplest, that includes one camera and one light source, to the most complex, that include multiple cameras and multiple light sources. This analysis provides fundamental insight into the limitations and potential of each configuration, including the personal calibration requirements.

For each system configuration there is more than one way to calculate the point-of-gaze from the principles of geometrical optics. This is because there is more than one way to express the principles of geometrical optics as equations, and for a given system of nonlinear equations there is often more than one way of solving them. Most importantly, each way of calculating the point-of-gaze can exhibit a significantly different performance in non-ideal conditions (e.g., noise in the image coordinates of the eye features, corneal asphericity, errors in system parameters).

Although the analysis from Chapter 3 is rather lengthy, it is probably the most complete analysis of the problem of estimating the point-of-gaze from the pupil and corneal reflections available to date (an earlier version of this work can be found in [84]). It includes concepts developed

throughout my M.A.Sc. and Ph.D. programs as well as concepts published by other researchers between 2000 and 2008. In particular, the analysis from Chapter 3 shows that by using at least two cameras and at least two light sources it is possible to reconstruct the optic axis of the eye in 3-D space, and therefore measure relative eye movements, without the need for personal calibration. The visual axis and the point-of-gaze in 3-D space can be estimated after completing a one-time personal calibration procedure in which the subject is required to fixate only one known target calibration point. Because of its theoretical and practical interest, the remainder of the thesis focuses on the point-of-gaze estimation problem using a spherical corneal model with two cameras and two or more light sources.

Chapter 4 describes a detailed simulation study of different methods to estimate the point-of-gaze using two cameras and two or more light sources. The sensitivity of the point-of-gaze estimates to several sources of errors (noise in the image coordinates of the eye features, corneal asphericity, errors in system parameters) is evaluated for the different methods and for different system setups (number of light sources, relative positions of the cameras and light sources). The understanding of the sensitivities of the different methods and the role of the geometrical system setup is essential for a successful system implementation. No other study of this nature and magnitude is currently available in the literature.

Chapter 5 presents a series of preliminary, proof-of-concept experiments carried out with adults and infants using a prototype system that uses two cameras and four light sources to estimate the point-of-gaze on a computer screen. The prototype system tolerates moderate head movements, which are only limited by the field of view and depth of field of the cameras used. The experimental results with adults, obtained with a system setup that follows the guidelines for system implementation derived from the simulation study from Chapter 4, show that using the proposed methodologies it is possible to estimate the point-of-gaze with an accuracy of 0.4-0.6° of visual angle after completing a single-point personal calibration procedure. Such point-of-gaze estimation accuracy is comparable to the best commercially available systems, which require multiple-point personal calibration procedures. Furthermore, to the best of my knowledge, the point-of-gaze estimation accuracy obtained with the prototype system is better than that of any other working systems described in the literature that use a single-point personal calibration procedure. The experimental results with infants demonstrate the ability of a system that requires a single-point calibration procedure to estimate infants' point-of-gaze. The ability to monitor

infants' visual scanning behavior can enable many applications that currently rely on subjective measurements performed by trained human observers and are very difficult or impossible to carry out with systems that require multiple-point personal calibration procedures. Some examples of applications with infants are the study of the development and the objective assessment of functions of the visual and oculomotor systems of preverbal infants, the objective determination of attention allocation, and the study of cognitive development.

Chapter 6 summarizes the contributions of this work and outlines some possible future work directions.

## Chapter 2

### Mathematical Model

## 2.1 Introduction

This chapter presents a detailed mathematical model for video-based gaze estimation systems that use the pupil and one or more corneal reflections extracted from images captured by one or more video cameras. The corneal reflections (first Purkinje images, glints) are virtual images of light sources (usually infrared) that illuminate the eye (Fig. 1.2). These virtual images are created by reflection on the front surface of the cornea, which acts as a convex mirror. Given the geometrical setup of a gaze estimation system (position of light sources, camera parameters), the mathematical model relates the position and orientation of the eye (or the position of the eye and the point-of-gaze) in 3-D space with the images of the pupil and corneal reflections captured by the camera(s). In order to derive this mathematical model, models of the eye and the system components (cameras, light sources) are developed first and then they are integrated into a single model based on the laws of reflection and refraction.

## 2.2 Eye model

The eye is a complex optical system (Fig. 1.1). Before reaching the retina, light is refracted several times as it travels through the cornea, the aqueous humor, the lens and the vitreous humor. Refraction occurs at each optical interface, i.e., where there is a change in index of refraction: the front surface of the cornea (air-cornea interface), the back surface of the cornea (cornea-aqueous humor interface), the front surface of the lens (aqueous humor-lens interface) and the back surface of the lens (lens-vitreous humor interface). The cornea is a very powerful refractive surface, accounting for about 60-75 % of the total refractive power of the eye. The lens has the ability to change its shape and, therefore, its refractive power, in order to focus images on the retina (accommodation). The shape of the lens is changed through relaxation and contraction of the ciliary muscles (Fig. 1.1).

Before reaching the lens, light passes through the pupil. The pupil is the opening at the center of the iris, which acts as a diaphragm that controls the amount of light reaching the retina. The diameter of the pupil is controlled by the smooth muscle fibers in the iris in response to changes in light intensity, as well as other physiological factors, psychological factors and certain drugs. In addition to controlling the amount of light reaching the retina, the pupil has an important role in image quality: as the pupil size is decreased, the depth of field is increased and the effect of spherical aberration is reduced.

Since the eye is a compound optical system (cornea and lens), it has two nodal points (Fig. 2.1). The nodal points are points on the optic axis of the optical system such that if a ray is directed towards one of them, the ray will emerge on the other side of the optical system parallel to its entering direction and seeming to be emerging from the other nodal point (i.e., if the extension of the entering ray passes through one nodal point, the extension of the emerging ray passes through the other nodal point, parallel to the entering ray).

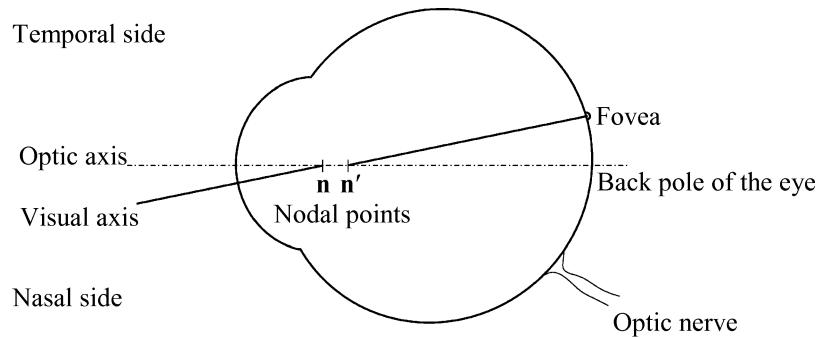


Fig. 2.1: Optic axis, visual axis and nodal points of the eye (right eye as viewed from above).  
(Adapted from [90].)

When a person with normal vision examines an object, his/her eyes are oriented in such a way that the image of the object falls at the center of the fovea of each eye. The fovea is the region of highest acuity of the retina corresponding to  $0.6\text{--}1^\circ$  of visual angle [27], and is displaced relative to the point of intersection of the optic axis and the retina (back pole of the eye). From the definition of the nodal points it follows that the line joining the point that the person is looking at (the point-of-gaze) with the front nodal point of the eye (**n**) will be parallel to the line joining the

rear nodal point ( $\mathbf{n}'$ ) and the center of the fovea (Fig. 2.1). The segment from the center of the fovea to the rear nodal point, and the line from the front nodal point to the point-of-gaze constitute the visual axis (Fig. 2.1) [96], [97]. Since the two nodal points of the eye are only 0.3 mm apart [27], for the sake of simplicity, they will be considered as a single nodal point (this is a common assumption in the gaze estimation literature –e.g., [27], [29], [61], [62], [64], [77]-[95]). With this simplification, the visual axis can be defined simply as the line passing through the center of the fovea and the nodal point of the eye. The point-of-gaze can then be defined as the intersection of the visual axis with the 3-D scene, or alternatively, as the intersection of the visual axes of both eyes.

For the gaze estimation problem, the elements that are relevant are the optic axis, the visual axis, the cornea and the pupil. In order to model these elements, an eye coordinate system that is rigidly attached to the eye is defined first.

### 2.2.1 Eye coordinate system (ECS)

The Eye Coordinate System (ECS) is a right-handed 3-D Cartesian coordinate system whose axes are represented by  $x_{\text{eye}}$ ,  $y_{\text{eye}}$ ,  $z_{\text{eye}}$ . The origin of the ECS is defined at the center of the eyeball and the  $z_{\text{eye}}$ -axis is coincident with the optic axis of the eye, pointing forward (Fig. 2.2). The  $x_{\text{eye}}$ -axis and  $y_{\text{eye}}$ -axis are defined such that, when the eye is in the primary position, the  $x_{\text{eye}}$ -axis is horizontal and the  $y_{\text{eye}}$ -axis points up. The primary position can be defined as the position of the eyes in the skull when the head is in a natural erect position and the visual axis is horizontal and perpendicular to the line connecting the centers of rotation of both eyes [96] (another definition is provided later in Appendix A, Section A.1).

Since there is no single point in the eye that remains stationary with respect to the skull as the eye rotates [96], [97], it is therefore not possible to define a center of rotation in strict sense. It is only possible to define an instantaneous center of rotation, which moves relative to the eye as the eye rotates [96], [97]. To avoid this complication, for the purpose of defining the orientation of the eye in space (Appendix A), it is assumed that the center of rotation of the eye is fixed and coincident with the geometrical center of the eyeball. It should be noted, however, that the gaze estimation methodologies presented in this thesis do not make any assumption about the position or behavior of the center of rotation of the eye.

## 2.2.2 Visual axis

The visual axis intersects the optic axis at the nodal point of the eye,  $\mathbf{n}$ , which is located at a distance  $D_N$  from the center of the eyeball,  $\mathbf{d}$  (Fig. 2.2). The deviation of the visual axis with respect to the optic axis of the eye can be described in the ECS by the horizontal and vertical angles  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , respectively, as shown in Fig. 2.2. From this figure, it follows that the unit vector in the direction of the visual axis with respect to the ECS can be expressed as

$$\mathbf{v}_{\text{ECS}} \equiv \mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = \begin{bmatrix} -\sin \alpha_{\text{eye}} \cos \beta_{\text{eye}} \\ \sin \beta_{\text{eye}} \\ \cos \alpha_{\text{eye}} \cos \beta_{\text{eye}} \end{bmatrix}. \quad (2.1)$$

In a typical adult human eye, the fovea falls about  $4\text{-}5^\circ$  temporally and about  $1.5^\circ$  below the point of intersection of the optic axis and the retina (back pole of the eye) [98]. It then follows that with the definition of  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  from Fig. 2.2,  $\alpha_{\text{eye}} < 0$  for the right eye,  $\alpha_{\text{eye}} > 0$  for the left eye, and  $\beta_{\text{eye}} > 0$  for both eyes ( $\alpha_{\text{eye}, \text{right}} = -5^\circ$ ,  $\alpha_{\text{eye}, \text{left}} = 5^\circ$ , and  $\beta_{\text{eye}} = 1.5^\circ$  are assumed as typical values).

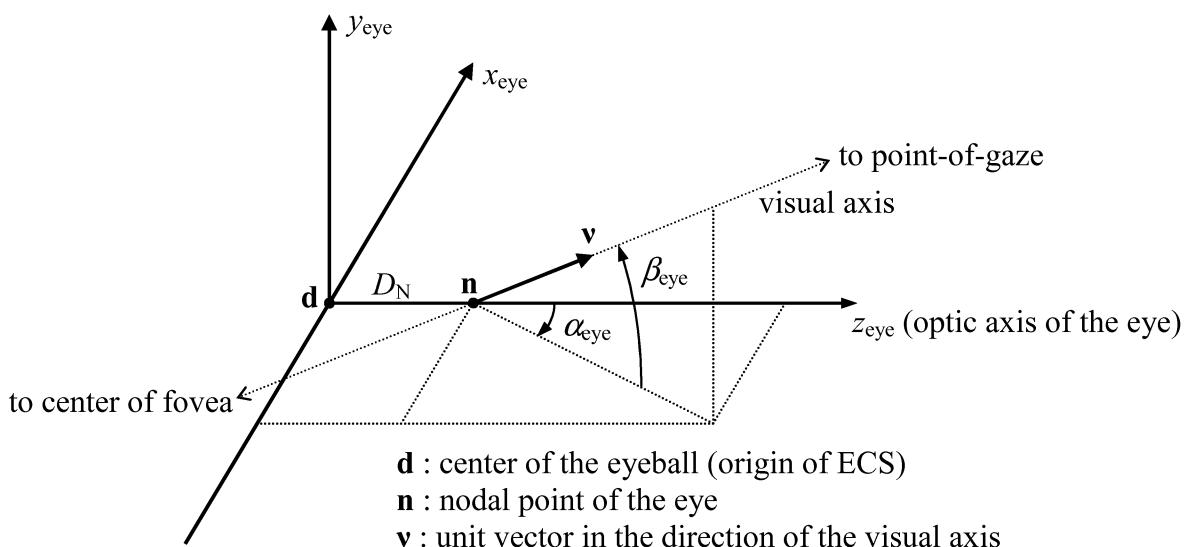


Fig. 2.2: Visual axis in the ECS.  
(Diagram is not to scale.)

### 2.2.3 Corneal model

In general terms, the front surface of the cornea can be described in the ECS by an implicit equation of the form

$$C_{\text{ECS}}(x_{\text{eye}}, y_{\text{eye}}, z_{\text{eye}}) = 0 , \quad (2.2)$$

where  $C_{\text{ECS}}(x_{\text{eye}}, y_{\text{eye}}, z_{\text{eye}})$  is a smooth function in the region of interest, i.e.,  $\nabla C_{\text{ECS}}(x_{\text{eye}}, y_{\text{eye}}, z_{\text{eye}})$  is a continuous function in the region of interest.

The simplest model of the front surface of the cornea is the spherical model, which has been the most common model in the gaze estimation literature (e.g., [27], [29], [30], [61], [62], [71], [73], [76]-[78], [80]-[95]). Next in the scale of complexity is the ellipsoidal model [79], [99]-[102], which includes the spherical model as a particular case. Other more general models (e.g., ellipsotoric [102], polynomial [78], [100]) have been proposed, but only the spherical and ellipsoidal models will be discussed here.

#### 2.2.3.1 Spherical corneal model

In the spherical corneal model, the front surface of the cornea is modeled as a spherical section of radius  $R$  whose center is located at a point  $\mathbf{c}$  (center of curvature of the cornea) in the optic axis of the eye. The center of curvature of the cornea is at a distance  $D_C$  from the center of the eyeball,  $\mathbf{d}$ . The normal to the spherical corneal surface, which is essential for the application of the laws of reflection and refraction, has the property of always going through the center of curvature of the cornea,  $\mathbf{c}$ .

Since the nodal point of the eye,  $\mathbf{n}$ , remains within 1 mm of the center curvature of the cornea,  $\mathbf{c}$ , for different degrees of eye accommodation [27], for the sake of simplicity, it is assumed hereafter that the nodal point is coincident with the center of corneal curvature, i.e.,  $\mathbf{n} \equiv \mathbf{c}$ , and, therefore  $D_N \equiv D_C \equiv D$ . This is another common assumption in the gaze estimation literature (e.g., [61], [62], [73], [77], [78], [80]-[95]).

### 2.2.3.2 Ellipsoidal corneal model

In the ellipsoidal corneal model, the front surface of the cornea is modeled as a section of an ellipsoid that has one of its axes coincident with the optic axis of the eye. Towards deriving the analytical expression for this model, the planar case will be considered first, which is equivalent to considering a transversal section containing the optic axis of the eye. This transversal section can be expressed as

$$\frac{(z_{\text{eye}} - A)^2}{a^2} + \frac{\tau^2}{b^2} = 1 . \quad (2.3)$$

It is convenient to express the parameters  $a$ ,  $b$ , and  $A$  as functions of anatomical eye parameters such as the radius of curvature at the apex of the cornea ( $\tau = 0$ ),  $R_0$ , the radius of curvature of the cornea at a distance  $\tau = \tau^*$  from the optic axis,  $R_{\tau^*}$ , and the distance between the apex of the cornea and the center of the eyeball,  $D_{\text{apex}}$ . Following the outline from Appendix B, it can then be shown that

$$b = \frac{R_0 \tau^*}{\sqrt{R_0^2 + \tau^{*2} - \sqrt[3]{R_0^4 R_{\tau^*}^2}}} , \quad (2.4)$$

$$a = \frac{b^2}{R_0} , \quad (2.5)$$

$$A = D_{\text{apex}} - a . \quad (2.6)$$

If  $\tau^2 = x_{\text{eye}}^2 + y_{\text{eye}}^2$  in (2.3), an ellipsoid of revolution is obtained (i.e., the cross-sections of the corneal surface perpendicular to the optic axis are circles). However, a more general model can be obtained if  $\tau^2 = \gamma_1 x_{\text{eye}}^2 + \gamma_2 x_{\text{eye}} y_{\text{eye}} + \gamma_3 y_{\text{eye}}^2$  so that the cross-sections of the corneal surface perpendicular to the optic axis are elliptical, allowing for the modeling of corneal astigmatism. The coefficients  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  can be expressed as functions of the ratio of the lengths of the major axis to the minor axis and the orientation of the major axis of those elliptical cross-sections. Towards this end, consider the following equation of an ellipse whose axes are coincident with some orthogonal coordinate axes  $x_o$  and  $y_o$ :

$$\frac{x_o^2}{a_{cs}^2} + \frac{y_o^2}{b_{cs}^2} = 1 \quad (2.7)$$

subject to the constraint

$$a_{cs} b_{cs} = 1 \quad (2.8)$$

so that the area remains constant regardless of the values of  $a_{cs}$  and  $b_{cs}$  (the area is the same as that of a circle of unit radius). Let  $a_{cs}$  be the half-length of the major axis and  $b_{cs}$  be the half-length of the minor axis. The ratio of the lengths of the major axis to the minor axis is then represented by

$$\xi = \frac{a_{cs}}{b_{cs}}, \quad \xi \geq 1. \quad (2.9)$$

After solving (2.8) and (2.9) for  $a_{cs}$  and  $b_{cs}$ , (2.7) can be rewritten as

$$\frac{x_o^2}{\xi} + \xi y_o^2 = 1. \quad (2.10)$$

Next, consider a rotation of the  $x_o$ - and  $y_o$ - axes relative to the  $x_{eye}$ - and  $y_{eye}$ - axes by an angle  $\delta$  as shown in Fig. 2.3. This angle  $\delta$  represents the orientation of the major axis of the ellipse with respect to the  $x_{eye}$ -axis. The  $x_o$ - and  $y_o$ - axes are related to the  $x_{eye}$ - and  $y_{eye}$ - axes by

$$\begin{aligned} x_o &= x_{eye} \cos \delta + y_{eye} \sin \delta \\ y_o &= y_{eye} \cos \delta - x_{eye} \sin \delta \end{aligned} \quad . \quad (2.11)$$

Substituting (2.11) into (2.10), an equation of the form

$$\gamma_1 x_{eye}^2 + \gamma_2 x_{eye} y_{eye} + \gamma_3 y_{eye}^2 = 1 \quad (2.12)$$

is obtained, where

$$\gamma_1 = \frac{\cos^2 \delta}{\xi} + \xi \sin^2 \delta, \quad (2.13)$$

$$\gamma_2 = 2 \cos \delta \sin \delta \left( \frac{1}{\xi} - \xi \right), \quad (2.14)$$

$$\gamma_3 = \frac{\sin^2 \delta}{\xi} + \xi \cos^2 \delta. \quad (2.15)$$

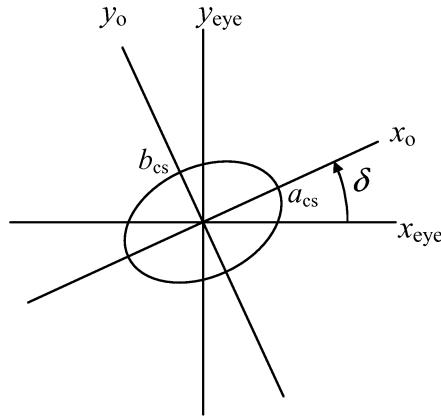


Fig. 2.3: Rotation of the  $x_o$ - and  $y_o$ - axes relative to the  $x_{eye}$ - and  $y_{eye}$ - axes by an angle  $\delta$ .

Substituting  $\tau^2 = \gamma_1 x_{eye}^2 + \gamma_2 x_{eye} y_{eye} + \gamma_3 y_{eye}^2$  into (2.3), the ellipsoidal corneal model can be expressed with respect to the ECS as

$$C_{ECS}(x_{eye}, y_{eye}, z_{eye}) = \frac{(z_{eye} - A)^2}{a^2} + \frac{\gamma_1 x_{eye}^2 + \gamma_2 x_{eye} y_{eye} + \gamma_3 y_{eye}^2}{b^2} - 1 = 0, \quad (2.16)$$

and the normal to the corneal surface is then given by

$$\nabla C_{ECS}(x_{eye}, y_{eye}, z_{eye}) = \begin{bmatrix} \frac{2\gamma_1 x_{eye} + \gamma_2 y_{eye}}{b^2} \\ \frac{\gamma_2 x_{eye} + 2\gamma_3 y_{eye}}{b^2} \\ \frac{2(z_{eye} - A)}{a^2} \end{bmatrix}, \quad (2.17)$$

where  $a$ ,  $b$ , and  $A$  are given by (2.4)-(2.6), respectively, and  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  are given by (2.13)-(2.15).

## 2.2.4 Pupil model

The pupil is modeled as a circular opening with radius  $r_{\text{pupil}}$ , whose center is located at a point  $\mathbf{p}$  (pupil center) on the optic axis of the eye, and such that the plane of the pupil is perpendicular to the optic axis of the eye (Fig. 2.4). The plane of the pupil is at a distance  $K$  from the nodal point of the eye / center of curvature of the spherical cornea. The pupil boundary can then be described in the ECS as

$$\mathbf{p}_{\text{b,ECS}}(\theta_{\text{pupil}}) = \begin{bmatrix} r_{\text{pupil}} \cos \theta_{\text{pupil}} \\ r_{\text{pupil}} \sin \theta_{\text{pupil}} \\ D + K \end{bmatrix}, \quad 0 \leq \theta_{\text{pupil}} < 360^\circ. \quad (2.18)$$

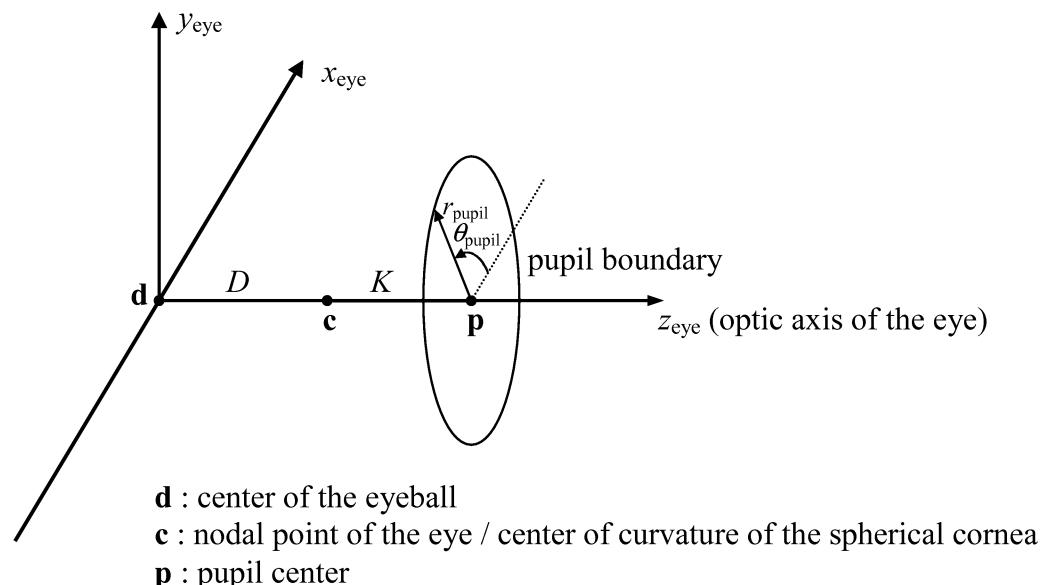


Fig. 2.4: Pupil model in the ECS.

## 2.3 Camera model

The model adopted for the cameras is the pinhole camera model extended with a lens distortion model that accounts for the image distortion produced by real lenses. In order to model the cameras, it is convenient to define two coordinate systems: the Camera Coordinate System and the Image Coordinate System.

### 2.3.1 Camera coordinate system (CCS)

The Camera Coordinate System (CCS) is a right-handed 3-D Cartesian coordinate system whose axes are represented by  $x_{\text{cam}}$ ,  $y_{\text{cam}}$ ,  $z_{\text{cam}}$  (Fig. 2.5). The origin of the CCS is coincident with the nodal point of the camera (**o**) and the  $z_{\text{cam}}$ -axis is perpendicular to the plane of the image sensor, coincident with the optic axis of the camera and pointing forward. The  $x_{\text{cam}}$ -axis is parallel to the rows of the image sensor and the  $y_{\text{cam}}$ -axis is parallel to the columns of the image sensor, pointing upwards. The coordinates in the CCS are measured in units of length.

### 2.3.2 Image coordinate system (ICS)

The Image Coordinate System (ICS) is a 2-D Cartesian coordinate system whose axes are represented by  $r$  and  $c$  (Fig. 2.5 –note that the image is inverted). The origin of the ICS is located at the upper-left corner of the image, and  $r$  and  $c$  represent, respectively, the row and column coordinates measured in pixels.

### 2.3.3 Perspective projection

The perspective projection, i.e., the projection of 3-D points onto the camera image plane, is illustrated in Fig. 2.5. In the pinhole camera model, the nodal point of the camera, **o**, is located at (the center of) the pinhole. The nodal point is the point on the optic axis where all lines that join object points (e.g., **q**) with their respective image points (e.g., **u**) intersect. The nodal point of a camera is also known as center of projection, camera center, and, sometimes, lens center.

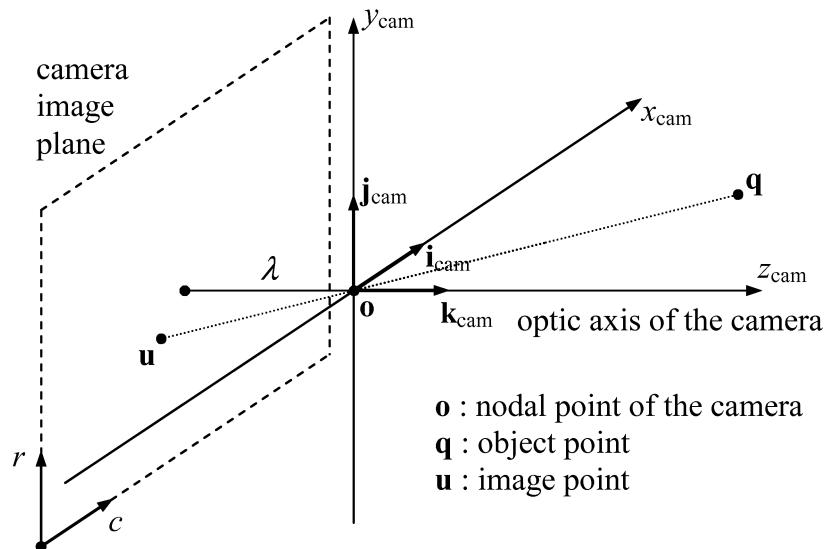
Let an object point and its corresponding image point be represented with respect to the CCS as  $\mathbf{q}_{\text{CCS}} = [q_{x,\text{cam}} \ q_{y,\text{cam}} \ q_{z,\text{cam}}]^T$  and  $\mathbf{u}_{\text{CCS}} = [u_{x,\text{cam}} \ u_{y,\text{cam}} \ u_{z,\text{cam}}]^T$ , respectively. Since the image plane is parallel to the  $x_{\text{cam}}y_{\text{cam}}$  plane, all image points have the same  $z_{\text{cam}}$  coordinate, i.e.,  $u_{z,\text{cam}} = -\lambda$ , where  $\lambda$  is the distance between the nodal point of the camera and the image plane. For a pinhole camera,  $\lambda$  equals the focal length. For a camera with a lens that has to be focused,  $\lambda$ , which is

hereafter referred to as effective focal length, is related to the focal length of the lens,  $f$ , by the Gaussian lens formula:

$$\frac{1}{q_{z,\text{cam}}} + \frac{1}{\lambda} = \frac{1}{f}. \quad (2.19)$$

From the properties of similar triangles, it follows that the coordinates of the object point,  $\mathbf{q}_{\text{CCS}}$ , and the coordinates of the image point,  $\mathbf{u}_{\text{CCS}}$ , are related by

$$\begin{aligned} \frac{q_{x,\text{cam}}}{q_{z,\text{cam}}} &= -\frac{u_{x,\text{cam}}}{\lambda}, \\ \frac{q_{y,\text{cam}}}{q_{z,\text{cam}}} &= -\frac{u_{y,\text{cam}}}{\lambda}. \end{aligned} \quad (2.20)$$



$x_{\text{cam}}, y_{\text{cam}}, z_{\text{cam}}$  : camera coordinate system  
 $rc$  : image coordinate system (note that image is inverted)

Fig. 2.5: Perspective projection and relation between the CCS and the ICS.

### 2.3.4 Lens distortion

Real lenses introduce a certain amount of distortion in the image. Depending on the characteristics of the lens, lens distortion ranges from negligible to significant. A typical model for lens distortion includes two components: radial distortion and tangential distortion [103], [104]. This distortion model can be expressed as

$$\begin{bmatrix} u_{x,\text{cam,DIST}} \\ u_{y,\text{cam,DIST}} \end{bmatrix} = \underbrace{\begin{bmatrix} u_{x,\text{cam}} \\ u_{y,\text{cam}} \end{bmatrix} + (k_{\text{r.d.1}}r^2 + k_{\text{r.d.2}}r^4 + k_{\text{r.d.3}}r^6) \begin{bmatrix} u_{x,\text{cam}} \\ u_{y,\text{cam}} \end{bmatrix}}_{\text{radial distortion}} + \underbrace{\begin{bmatrix} 2k_{\text{t.d.1}}u_{x,\text{cam}}u_{y,\text{cam}} + k_{\text{t.d.2}}(r^2 + 2u_{x,\text{cam}}^2) \\ k_{\text{t.d.1}}(r^2 + 2u_{y,\text{cam}}^2) + 2k_{\text{t.d.2}}u_{x,\text{cam}}u_{y,\text{cam}} \end{bmatrix}}_{\text{tangential distortion}},$$

$$r^2 = u_{x,\text{cam}}^2 + u_{y,\text{cam}}^2,$$
(2.21)

where  $(u_{x,\text{cam,DIST}}, u_{y,\text{cam,DIST}})$  are the coordinates of the image point with respect to the CCS after distortion,  $k_{\text{r.d.1}}$ ,  $k_{\text{r.d.2}}$  and  $k_{\text{r.d.3}}$  are the radial distortion coefficients, and  $k_{\text{t.d.1}}$  and  $k_{\text{t.d.2}}$  are the tangential distortion coefficients.

### 2.3.5 Transformation to the ICS

The general expression that relates the coordinates of the image point  $(u_{x,\text{cam,DIST}}, u_{y,\text{cam,DIST}})$  in the CCS (unit of length) with its coordinates  $(r, c)$  in the ICS (pixels) is

$$\begin{aligned} c &= (u_{x,\text{cam,DIST}} + s_{\text{cam}}u_{y,\text{cam,DIST}}) / \text{pixel pitch}_x + c_{\text{center}} \\ r &= u_{y,\text{cam,DIST}} / \text{pixel pitch}_y + r_{\text{center}} \end{aligned},$$
(2.22)

where  $c_{\text{center}}$  and  $r_{\text{center}}$  are, respectively, the column coordinate and row coordinate in pixels of the intersection of the optic axis of the camera with the image sensor (the principal point), and  $\text{pixel pitch}_x$  and  $\text{pixel pitch}_y$  are the distances in units of length between the centers of adjacent pixels in the  $x$  and  $y$  directions (across the columns and across the rows), respectively. The coefficient  $s_{\text{cam}}$ , which is usually referred to as the skew coefficient (or “skewness” coefficient) [104]-[108], encodes the angle between the sensor axes, allowing the camera model to account for cases where the pixels are not in a rectangular arrangement. Normally, however, the sensor axes are orthogonal (as assumed for the definition of the CCS –Section 2.3.1) and  $s_{\text{cam}} = 0$ . Furthermore, the pixels in digital cameras are typically square and, therefore,  $\text{pixel pitch}_x = \text{pixel}$

$pitch_y$ . With these simplifications,

$$\begin{bmatrix} c \\ r \end{bmatrix} = \frac{1}{pixel\ pitch} \begin{bmatrix} u_{x,cam,DIST} \\ u_{y,cam,DIST} \end{bmatrix} + \begin{bmatrix} c_{center} \\ r_{center} \end{bmatrix}. \quad (2.23)$$

The parameters  $\lambda$ ,  $k_{r.d.1}$ ,  $k_{r.d.2}$ ,  $k_{r.d.3}$ ,  $k_{t.d.1}$ ,  $k_{t.d.2}$ ,  $pixel\ pitch_x$ ,  $pixel\ pitch_y$ ,  $c_{center}$ ,  $r_{center}$ , and  $s_{cam}$  are collectively termed the intrinsic (or internal) camera parameters [103]-[109].

## 2.4 Light source model

The light sources that illuminate the eye and produce the corneal reflections are modeled as point sources. In practice, each light source consists of one or more near-infrared (NIR) light-emitting diodes (LEDs). When a light source consists of an array of LEDs, it is approximated by a point source located at the center of the array.

## 2.5 Integration of the eye, camera and light source models

The eye, camera and light source models are integrated into a single mathematical model that is based on the laws of reflection and refraction. Towards this end, it is convenient to define a World Coordinate System, with respect to which the position and orientation of the eye, the position and orientation of the camera(s), the position of the light source(s) and the location of objects in the visual scene are measured.

### 2.5.1 World coordinate system (WCS)

The World Coordinate System (WCS) is a right-handed 3-D Cartesian coordinate system whose axes are represented by  $X$ ,  $Y$ ,  $Z$ . The  $X$ -axis is horizontal, the  $Y$ -axis is vertical pointing up, and the  $Z$ -axis generally points towards the subject. When the scene is a vertical plane (e.g., a computer screen, a projection screen), the WCS is defined such that the  $XY$ -plane is coincident with the scene plane and the  $Z$ -axis extends outward from the scene plane. In the particular case of a computer screen or a projection screen, the origin of the WCS is set at the center of the screen.

### 2.5.2 Relation between the ECS, CCS and WCS

The relation between the coordinates of a point  $[x_{\text{eye}} \ y_{\text{eye}} \ z_{\text{eye}}]^T$  with respect to the ECS (Section 2.2.1) and its corresponding coordinates  $[X \ Y \ Z]^T$  with respect to the WCS is given by

$$\underbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}_{\text{WCS}} = \mathbf{R}_{\text{eye}} \underbrace{\begin{bmatrix} x_{\text{eye}} \\ y_{\text{eye}} \\ z_{\text{eye}} \end{bmatrix}}_{\text{ECS}} + \mathbf{d} , \quad (2.24)$$

where  $\mathbf{d}$  is the position of the center of the eyeball with respect to the WCS and  $\mathbf{R}_{\text{eye}} = [\mathbf{i}_{\text{eye}} \ \mathbf{j}_{\text{eye}} \ \mathbf{k}_{\text{eye}}]$  is the rotation matrix that describes the orientation of the ECS with respect to the WCS ( $\mathbf{i}_{\text{eye}}$ ,  $\mathbf{j}_{\text{eye}}$  and  $\mathbf{k}_{\text{eye}}$  are, respectively, the unit vectors in the direction of the  $x_{\text{eye}}$ -,  $y_{\text{eye}}$ - and  $z_{\text{eye}}$ - axes with respect to the WCS).

Similarly, the relation between the coordinates of a point  $[x_{\text{cam}} \ y_{\text{cam}} \ z_{\text{cam}}]^T$  with respect to the CCS (Section 2.3.1) and its corresponding coordinates  $[X \ Y \ Z]^T$  with respect to the WCS is given by

$$\underbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}_{\text{WCS}} = \mathbf{R}_{\text{cam}} \underbrace{\begin{bmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ z_{\text{cam}} \end{bmatrix}}_{\text{CCS}} + \mathbf{o} , \quad (2.25)$$

where  $\mathbf{o}$  is the position of the nodal point of the camera with respect to the WCS and  $\mathbf{R}_{\text{cam}} = [\mathbf{i}_{\text{cam}} \ \mathbf{j}_{\text{cam}} \ \mathbf{k}_{\text{cam}}]$  is the rotation matrix that describes the orientation of the CCS with respect to the WCS ( $\mathbf{i}_{\text{cam}}$ ,  $\mathbf{j}_{\text{cam}}$  and  $\mathbf{k}_{\text{cam}}$  are, respectively, the unit vectors in the direction of the  $x_{\text{cam}}$ -,  $y_{\text{cam}}$ - and  $z_{\text{cam}}$ -axes with respect to the WCS –see Fig. 2.5). The position of the nodal point,  $\mathbf{o}$ , and the rotation matrix,  $\mathbf{R}_{\text{cam}}$ , which encodes the orientation of the camera, are collectively termed the extrinsic (or external) camera parameters [103]-[109].

The above rotation matrices can be expressed in terms of three rotation angles that describe the pan, tilt and torsion of the ECS or CCS with respect to the WCS. This is particularly relevant for the ECS and is therefore described in detail in Appendix A. The case of the CCS is presented in Appendix C.

### 2.5.3 Eye model with respect to the WCS

The unit vector in the direction of the optic axis of the eye with respect to the WCS is given by  $\mathbf{k}_{\text{eye}}$ , the third column of  $\mathbf{R}_{\text{eye}} \equiv \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ , where  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$  are, respectively, the pan, tilt and torsion angles of the ECS with respect to the WCS as defined in Appendix A. Because of its importance in the mathematical model, the unit vector in the direction of the optic axis of the eye is assigned its own symbol,  $\boldsymbol{\omega}$ :

$$\boldsymbol{\omega} \equiv \boldsymbol{\omega}(\theta_{\text{eye}}, \varphi_{\text{eye}}) = \mathbf{k}_{\text{eye}} = \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin \theta_{\text{eye}} \cos \varphi_{\text{eye}} \\ \sin \varphi_{\text{eye}} \\ -\cos \theta_{\text{eye}} \cos \varphi_{\text{eye}} \end{bmatrix}. \quad (2.26)$$

Notice that the previous expression can also be obtained directly from Fig. 2.6, where the center of the WCS is translated to the center of the eyeball,  $\mathbf{d}$  (this is formally described by step 1) in Appendix A and (A.1)).

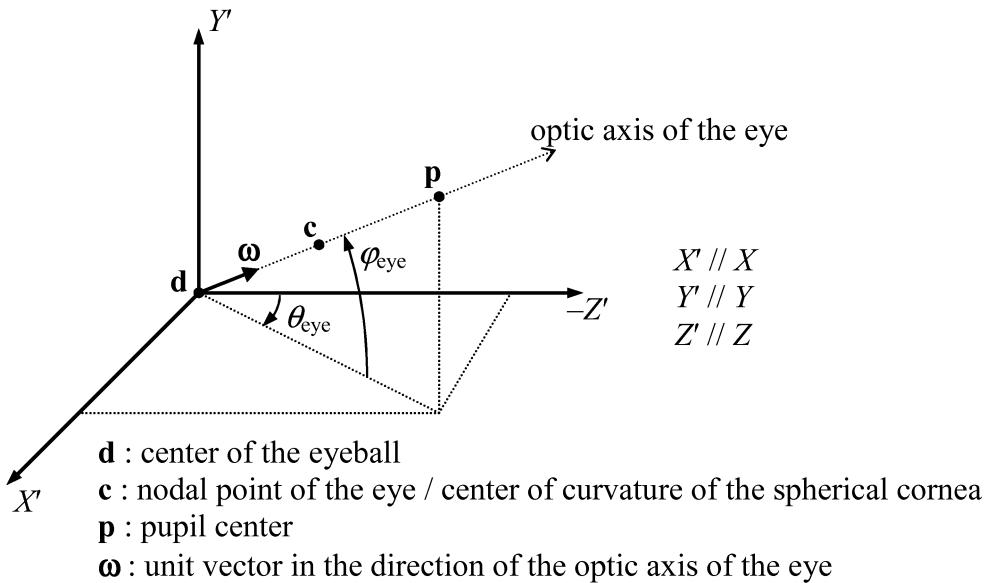


Fig. 2.6: Orientation of the optic axis of the eye with respect to the WCS.  
(Diagram is not to scale.)

The center of the eyeball,  $\mathbf{d}$ , the nodal point of the eye / center of curvature of the spherical cornea,  $\mathbf{c}$ , and the pupil center,  $\mathbf{p}$ , are related by

$$\mathbf{c} = \mathbf{d} + D\boldsymbol{\omega} \quad (2.27)$$

and

$$\mathbf{p} = \mathbf{c} + K\boldsymbol{\omega} = \mathbf{d} + (D + K)\boldsymbol{\omega} , \quad (2.28)$$

where  $D$  was defined in Section 2.2.3.1 and  $K$  was defined in Section 2.2.4.

The pupil boundary points are given by

$$\mathbf{p}_b(\mathbf{d}, \theta_{eye}, \varphi_{eye}, \kappa_{eye}, \theta_{pupil}) = \mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})\mathbf{p}_{b,ECS}(\theta_{pupil}) + \mathbf{d} , \quad 0 \leq \theta_{pupil} < 360^\circ , \quad (2.29)$$

where  $\mathbf{p}_{b,ECS}(\theta_{pupil})$  is given by (2.18) and  $\mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})$  is given by (A.6).

The unit vector in the direction of the visual axis with respect to the WCS is given by

$$\mathbf{v} \equiv \mathbf{v}(\theta_{eye}, \varphi_{eye}, \kappa_{eye}, \alpha_{eye}, \beta_{eye}) = \mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})\mathbf{v}_{ECS}(\alpha_{eye}, \beta_{eye}) , \quad (2.30)$$

where  $\mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})$  is given by (A.6) and  $\mathbf{v}_{ECS}(\alpha_{eye}, \beta_{eye})$  is given by (2.1). Since the visual axis goes through  $\mathbf{c}$ , it can then be described with respect to the WCS as

$$\mathbf{g} = \underbrace{\mathbf{d} + D\boldsymbol{\omega}(\theta_{eye}, \varphi_{eye})}_{\mathbf{c}} + k_g \underbrace{\mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})\mathbf{v}_{ECS}(\alpha_{eye}, \beta_{eye})}_{\mathbf{v}(\theta_{eye}, \varphi_{eye}, \kappa_{eye}, \alpha_{eye}, \beta_{eye})} , \quad k_g \in \Re , \quad (2.31)$$

where  $\boldsymbol{\omega}(\theta_{eye}, \varphi_{eye})$  is given by (2.26),  $\mathbf{R}_{eye}(\theta_{eye}, \varphi_{eye}, \kappa_{eye})$  is given by (A.6) and  $\mathbf{v}_{ECS}(\alpha_{eye}, \beta_{eye})$  is given by (2.1).

The corneal surface can be described with respect to the WCS, after using (2.2), (2.24) and a slight abuse of notation, as

$$C_{WCS}(X, Y, Z, \mathbf{d}, \mathbf{R}_{eye}) \equiv C_{WCS}([X \quad Y \quad Z]^T, \mathbf{d}, \mathbf{R}_{eye}) = C_{ECS}(\mathbf{R}_{eye}^T([X \quad Y \quad Z]^T - \mathbf{d})) = 0 , \quad (2.32)$$

$$\text{where } C_{ECS}([x_{eye} \quad y_{eye} \quad z_{eye}]^T) \equiv C_{ECS}(x_{eye}, y_{eye}, z_{eye}) .$$

The normal to the corneal surface is then given by

$$\begin{aligned}\nabla C_{\text{WCS}}(X, Y, Z, \mathbf{d}, \mathbf{R}_{\text{eye}}) &\equiv \nabla C_{\text{WCS}}([X \ Y \ Z]^T, \mathbf{d}, \mathbf{R}_{\text{eye}}) = \\ &= \mathbf{R}_{\text{eye}} \nabla C_{\text{ECS}}(\mathbf{R}_{\text{eye}}^T ([X \ Y \ Z]^T - \mathbf{d})) ,\end{aligned}\tag{2.33}$$

$$\text{where } \nabla C_{\text{ECS}}([x_{\text{eye}} \ y_{\text{eye}} \ z_{\text{eye}}]^T) \equiv \nabla C_{\text{ECS}}(x_{\text{eye}}, y_{\text{eye}}, z_{\text{eye}}) .$$

Note that since rotation matrices are orthonormal,  $\mathbf{R}_{\text{eye}}^{-1} = \mathbf{R}_{\text{eye}}^T$ .

## 2.5.4 Integrated model

Having developed models of the eye, the camera(s) and the light source(s), they are integrated into a single mathematical model that, given the system setup (position of light sources, camera parameters), relates the position and orientation of the eye (or the position of the eye and the point-of-gaze) in 3-D space with the pupil and corneal reflections in the images captured by the camera(s). Based on the laws of reflection and refraction, the mathematical model is derived first for the simplest case in which the front surface of the cornea is modeled as a spherical section and then it is generalized to non-spherical corneal models.

### 2.5.4.1 Integrated model for a spherical cornea

Under the modeling assumptions described in this chapter (i.e., the light sources are modeled as point sources and the video cameras are modeled as pinhole cameras), Fig. 2.7 presents a ray-tracing diagram of the system and the eye, where all points are represented as 3-D column vectors (bold font) in the WCS.

Consider a ray that comes from light source  $i$ ,  $\mathbf{l}_i$ , and reflects at a point  $\mathbf{q}_{ij}$  on the corneal surface such that the reflected ray passes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and intersects the camera image plane at a point  $\mathbf{u}_{ij}$ . The condition that the ray coming from the point of reflection  $\mathbf{q}_{ij}$  and passing through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , intersects the camera image plane at point  $\mathbf{u}_{ij}$ , can be expressed in parametric form as

$$\mathbf{q}_{ij} = \mathbf{o}_j + k_{q,ij} (\mathbf{o}_j - \mathbf{u}_{ij}) \text{ for some } k_{q,ij} .\tag{2.34}$$

If the corneal surface is modeled as a convex spherical mirror of radius  $R$ , the condition that  $\mathbf{q}_{ij}$  is on the corneal surface can be written as

$$\|\mathbf{q}_{ij} - \mathbf{c}\| = R , \quad (2.35)$$

where  $\mathbf{c}$  is the center of curvature of the cornea and  $\|\dots\|$  is the Euclidean norm.

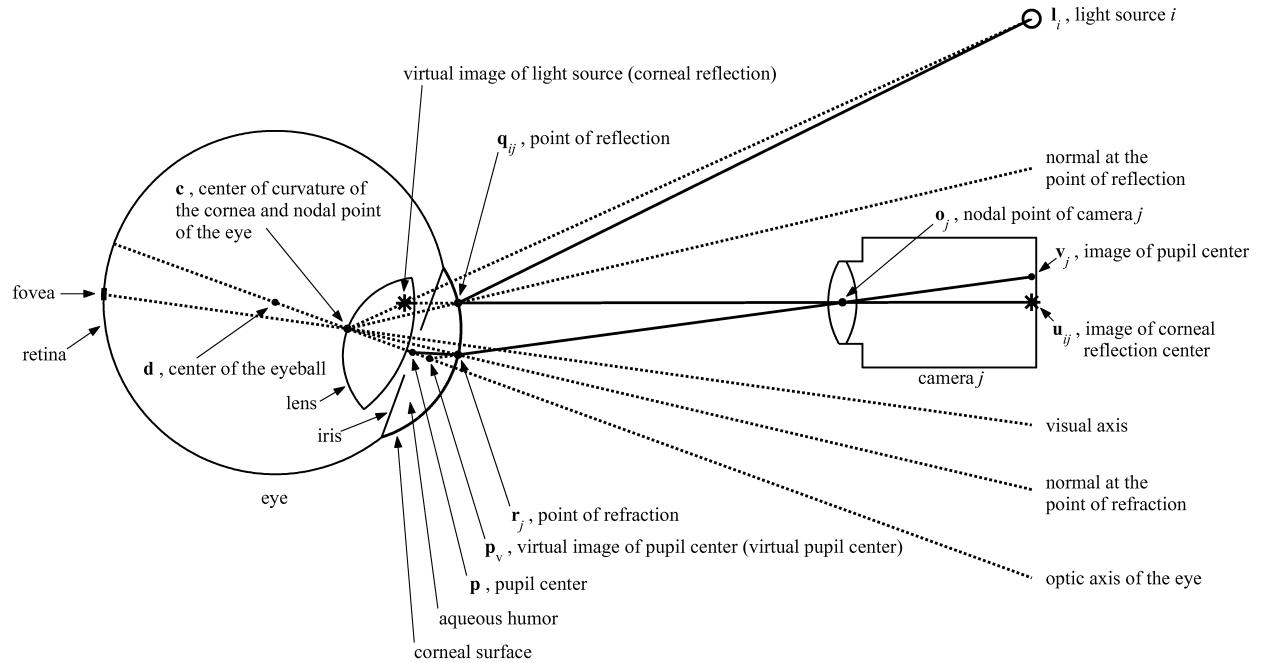


Fig. 2.7: Ray-tracing diagram, showing schematic representations of the eye, a camera and a light source. Note that eye parts are drawn to scale whereas the rest of the diagram is not to scale in order to be able to show all the elements of interest.

The law of reflection states two conditions: (1) the incident ray, the reflected ray and the normal at the point of reflection are coplanar; and (2) the angles of incidence and reflection are equal (the angles of incidence and reflection are measured with respect to the normal at the point of reflection). In order to express the law of reflection in analytical form, two different cases are considered.

The first, particular case is when the incident ray, the reflected ray and the normal at the point of reflection are coincident (i.e., the angles of incidence and reflection are nil, the incident ray reflects back on itself). Since vector  $(\mathbf{q}_{ij} - \mathbf{c})$  is normal to the spherical surface at the point of reflection  $\mathbf{q}_{ij}$ , this implies that points  $\mathbf{l}_i$ ,  $\mathbf{q}_{ij}$ ,  $\mathbf{o}_j$ ,  $\mathbf{u}_{ij}$ , and  $\mathbf{c}$  are collinear (i.e.,  $(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{c} - \mathbf{o}_j) = \mathbf{0}$ ,

$(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{u}_{ij} - \mathbf{o}_j) = \mathbf{0}$ . In this case, the law of reflection can be expressed analytically simply as

$$\frac{\mathbf{q}_{ij} - \mathbf{c}}{R} = \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} = \frac{\mathbf{u}_{ij} - \mathbf{o}_j}{\|\mathbf{u}_{ij} - \mathbf{o}_j\|} . \quad (2.36)$$

The typical system configuration that results in the incident and reflected rays being coincident is when the light source,  $\mathbf{l}_i$ , is located at the nodal point of the camera,  $\mathbf{o}_j$ . In practice, this is achieved in an approximate way by either (a) setting a single LED in the optic axis of the camera, in front of the lens and as close as possible to the lens, or (b) using a ring light source around the lens of the camera such that the center of the ring is coincident with the nodal point of the camera and the plane of the ring is perpendicular to the optic axis of the camera.

The second, more general case is when the incident ray, the reflected ray and the normal at the point of reflection are not coincident (i.e., the angles of incidence and reflection are not nil). In such case, condition (1) of the law of reflection implies that points  $\mathbf{l}_i$ ,  $\mathbf{q}_{ij}$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  are coplanar but not collinear. Noting that three coplanar vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  and  $\mathbf{a}_3$  satisfy  $\mathbf{a}_1 \times \mathbf{a}_2 \bullet \mathbf{a}_3 = 0$ , condition (1) can be formalized as, for example,

$$(\mathbf{l}_i - \mathbf{q}_{ij}) \times (\mathbf{o}_j - \mathbf{q}_{ij}) \bullet (\mathbf{q}_{ij} - \mathbf{c}) = 0 , \quad (2.37)$$

or

$$(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{q}_{ij} - \mathbf{o}_j) \bullet (\mathbf{c} - \mathbf{o}_j) = 0 . \quad (2.38)$$

Since the angle  $\theta$  between two vectors  $\mathbf{a}$  and  $\mathbf{b}$  can be obtained from  $\mathbf{a} \bullet \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta$ , condition (2) can be expressed as

$$\frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} \bullet \frac{\mathbf{q}_{ij} - \mathbf{c}}{R} = \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|} \bullet \frac{\mathbf{q}_{ij} - \mathbf{c}}{R} , \quad (2.39)$$

where the scalar factors  $R$  cancel out.

Finally, it should be noted that the following condition must be satisfied to avoid spurious solutions:

$$\|\mathbf{q}_{ij} - \mathbf{o}_j\| < \|\mathbf{c} - \mathbf{o}_j\| . \quad (2.40)$$

Noting that given two vectors  $\mathbf{a}$  and  $\mathbf{b}$  such that  $\|\mathbf{a}\| = \|\mathbf{b}\|$  and  $\mathbf{a} \neq -\mathbf{b}$ , vector  $\mathbf{a} + \mathbf{b}$  bisects the angle formed by  $\mathbf{a}$  and  $\mathbf{b}$ , (2.35) and (2.38)-(2.40) can be replaced by

$$\frac{\mathbf{q}_{ij} - \mathbf{c}}{R} = \frac{\frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|}}{\left\| \frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|} \right\|} , \quad (2.41)$$

which is valid regardless of the angle of incidence and guarantees unique solutions. Notice that the case “ $\mathbf{a} = -\mathbf{b}$ ” would happen if the line connecting the light source  $\mathbf{l}_i$  and the nodal point of the camera  $\mathbf{o}_j$  were tangent to the corneal surface at  $\mathbf{q}_{ij}$  (i.e., there is no actual reflection on the corneal surface), which does not happen in a practical situation.

Next, consider an imaginary ray that comes from the pupil center,  $\mathbf{p}$ , and refracts at point  $\mathbf{r}_j$  on the corneal surface such that the refracted ray passes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and intersects the camera image plane at a point  $\mathbf{v}_j$ . The condition that the ray coming from the point of refraction  $\mathbf{r}_j$  and passing through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , intersects the camera image plane at point  $\mathbf{v}_j$ , can be expressed in parametric form as

$$\mathbf{r}_j = \mathbf{o}_j + k_{r,j}(\mathbf{o}_j - \mathbf{v}_j) \text{ for some } k_{r,j} , \quad (2.42)$$

whereas the condition that  $\mathbf{r}_j$  is on the corneal surface can be written as

$$\|\mathbf{r}_j - \mathbf{c}\| = R . \quad (2.43)$$

The law of refraction states two conditions: (1) the incident ray, the refracted ray and the normal at the point of refraction are coplanar; and (2) the angle of incidence,  $\theta_1$ , and the angle of refraction,  $\theta_2$ , satisfy Snell's law (i.e.,  $n_1 \sin \theta_1 = n_2 \sin \theta_2$ , where  $n_1$  and  $n_2$  are the indices of refraction of mediums 1 and 2). In order to express the law of refraction in analytical form, two different cases are considered.

The first, particular case is when the incident ray, the refracted ray and the normal at the point of refraction are coincident (i.e., the angles of incidence and refraction are nil, the incident ray crosses the cornea-air optical interface without changing direction). Since vector  $(\mathbf{r}_j - \mathbf{c})$  is normal to the spherical surface at the point of refraction  $\mathbf{r}_j$ , this implies that points  $\mathbf{p}$ ,  $\mathbf{r}_j$ ,  $\mathbf{o}_j$ ,  $\mathbf{v}_j$  and  $\mathbf{c}$  are collinear (i.e.,  $(\mathbf{c} - \mathbf{o}_j) \times (\mathbf{p} - \mathbf{o}_j) = \mathbf{0}$ ,  $(\mathbf{c} - \mathbf{o}_j) \times (\mathbf{v}_j - \mathbf{o}_j) = \mathbf{0}$ , the optic axis of the eye goes through the nodal point of the camera,  $\mathbf{o}_j$ ). In this case, the law of refraction can be expressed analytically simply as

$$\frac{\mathbf{r}_j - \mathbf{c}}{R} = \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|} = \frac{\mathbf{v}_j - \mathbf{o}_j}{\|\mathbf{v}_j - \mathbf{o}_j\|}. \quad (2.44)$$

The second, more general case is when the incident ray, the refracted ray and the normal at the point of refraction are not coincident (i.e., the angles of incidence and refraction are not nil). In such case, condition (1) of the law of refraction implies that points  $\mathbf{p}$ ,  $\mathbf{r}_j$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  are coplanar, which can be formalized as

$$(\mathbf{r}_j - \mathbf{o}_j) \times (\mathbf{c} - \mathbf{o}_j) \bullet (\mathbf{p} - \mathbf{o}_j) = 0. \quad (2.45)$$

Since the sine of the angle  $\theta$  between two vectors  $\mathbf{a}$  and  $\mathbf{b}$  can be obtained from  $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta$ , condition (2) can be expressed as

$$n_1 \left\| \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} \times \frac{\mathbf{r}_j - \mathbf{c}}{R} \right\| = n_2 \left\| \frac{\mathbf{r}_j - \mathbf{c}}{R} \times \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} \right\|, \quad (2.46)$$

where the scalar factors  $R$  cancel out,  $n_1$  is the effective index of refraction of the aqueous humor and cornea combined ( $= 1.3375$  [110]) and  $n_2$  is the index of refraction of air ( $\cong 1$ ). In this model, the refraction at the aqueous humor-cornea interface is neglected since the difference in their indices of refraction is small relative to that of the cornea-air interface. Only the refraction at the cornea-air interface is taken into account and the aqueous humor and cornea are considered as a homogeneous medium.

Equation (2.46) is subject to the condition that the incident ray and the refracted ray are on opposite sides of the normal at the point of refraction, which can be expressed as

$$[(\mathbf{p} - \mathbf{r}_j) \times (\mathbf{r}_j - \mathbf{c})] \bullet [(\mathbf{r}_j - \mathbf{c}) \times (\mathbf{o}_j - \mathbf{r}_j)] > 0 . \quad (2.47)$$

Since points  $\mathbf{p}$ ,  $\mathbf{r}_j$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  are coplanar,  $[(\mathbf{p} - \mathbf{r}_j) \times (\mathbf{r}_j - \mathbf{c})]$  and  $[(\mathbf{r}_j - \mathbf{c}) \times (\mathbf{o}_j - \mathbf{r}_j)]$  are parallel, and, therefore, (2.45)-(2.47) can be integrated into a single equation as

$$n_1 \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} \times (\mathbf{r}_j - \mathbf{c}) = n_2 (\mathbf{r}_j - \mathbf{c}) \times \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} , \quad (2.48)$$

or, after rearranging terms,

$$\left( n_1 \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} + n_2 \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} \right) \times (\mathbf{r}_j - \mathbf{c}) = \mathbf{0} , \quad (2.49)$$

which is valid regardless of the angle of incidence.

Finally, it should be noted that the following condition must be satisfied to avoid spurious solutions:

$$\|\mathbf{r}_j - \mathbf{o}_j\| < \|\mathbf{c} - \mathbf{o}_j\| . \quad (2.50)$$

Note that the preceding discussion applies to the pupil boundary points by replacing  $\mathbf{p}$  by  $\mathbf{p}_b$ ,  $\mathbf{r}_j$  by  $\mathbf{r}_{b,j}$ , and  $\mathbf{v}_j$  by  $\mathbf{v}_{b,j}$ .

#### 2.5.4.2 Generalization of the integrated model to non-spherical corneas

For a non-spherical corneal surface described by  $C_{\text{WCS}}([X \ Y \ Z]^T, \mathbf{d}, \mathbf{R}_{\text{eye}}) = 0$ , with its normal given by  $\nabla C_{\text{WCS}}([X \ Y \ Z]^T, \mathbf{d}, \mathbf{R}_{\text{eye}})$  (see Section 2.5.3), the condition that the point of reflection  $\mathbf{q}_{ij}$  is on the corneal surface can be written as

$$C_{\text{WCS}}(\mathbf{q}_{ij}, \mathbf{d}, \mathbf{R}_{\text{eye}}) = 0 \quad (2.51)$$

and the conditions of the law of reflection can be written as

$$\frac{\nabla C_{\text{WCS}}(\mathbf{q}_{ij}, \mathbf{d}, \mathbf{R}_{\text{eye}})}{\|\nabla C_{\text{WCS}}(\mathbf{q}_{ij}, \mathbf{d}, \mathbf{R}_{\text{eye}})\|} = \frac{\frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|}}{\left\| \frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|} \right\|}, \quad (2.52)$$

where it is assumed that  $\nabla C_{\text{WCS}}(\mathbf{q}_{ij}, \mathbf{d}, \mathbf{R}_{\text{eye}})$  points out of the eye.

The condition that the point of refraction  $\mathbf{r}_j$  is on the corneal surface can be written as

$$C_{\text{WCS}}(\mathbf{r}_j, \mathbf{d}, \mathbf{R}_{\text{eye}}) = 0 \quad (2.53)$$

and the conditions of the law of refraction can be written as

$$\left( n_1 \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} + n_2 \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} \right) \times \nabla C_{\text{WCS}}(\mathbf{r}_j, \mathbf{d}, \mathbf{R}_{\text{eye}}) = \mathbf{0}, \quad (2.54)$$

subject to a constraint like (2.50), where  $\mathbf{c}$  represents the nodal point of the eye, to avoid spurious solutions.

The use of the mathematical model to estimate gaze in 3-D space using the spherical corneal model is discussed in great detail in the next chapters. The use of the mathematical model as the basis of a simulation framework to generate synthetic eye feature data is discussed in detail in the next section. The generation of synthetic eye feature data consists of determining the pupil boundary, the pupil center and the centers of the corneal reflections (eye features) in the images captured by the camera(s), given a set of eye parameters, the system parameters (position of the light sources, camera parameters) and the 3-D positions of the center of the eyeball,  $\mathbf{d}$ , and the point-of-gaze,  $\mathbf{g}$ . The simulation framework is an invaluable tool that supports studies of the sensitivity of the estimation of the point-of-gaze to several different types of error such as errors in the estimation of the coordinates of the eye features in each image of the eye, the deviation of the actual corneal shape from the ideal spherical shape assumed for the estimation of the point-of-gaze (corneal asphericity), and errors in the system parameters.

## 2.6 Generation of synthetic eye feature data

The first step is to find the orientation of the ECS with respect to the WCS from the 3-D positions of the center of the eyeball,  $\mathbf{d}$ , and the point-of-gaze,  $\mathbf{g}$ , and the deviation of the visual axis from the optic axis of the eye ( $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , or simply  $\mathbf{v}_{\text{ECS}}$  –see Section 2.2.2). Having determined the orientation of the ECS with respect to the WCS, the second step is to determine, for each camera, the points of reflection associated with each light source, and the points of refraction associated with the pupil center and each pupil boundary point. The last step is to project the points of reflection and refraction onto the camera image plane to obtain, respectively, the images of the centers of the corneal reflections, the pupil center and the pupil boundary points.

### 2.6.1 Orientation of the ECS with respect to the WCS

Given the 3-D positions of the center of the eyeball,  $\mathbf{d}$ , and the point-of-gaze,  $\mathbf{g}$ , and the deviation of the visual axis from the optic axis of the eye ( $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , or simply  $\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}})$ ), the determination of the orientation of the ECS with respect to the WCS ( $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ ) is based on (2.31). In this case, (2.31) is equivalent to 3 scalar equations with 4 scalar unknowns ( $k_g$ ,  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$ ). The scalar unknown  $k_g$  can be eliminated if (2.31) is rewritten as

$$\frac{\mathbf{g} - \mathbf{d} - D\boldsymbol{\omega}(\theta_{\text{eye}}, \varphi_{\text{eye}})}{\|\mathbf{g} - \mathbf{d} - D\boldsymbol{\omega}(\theta_{\text{eye}}, \varphi_{\text{eye}})\|} = \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}) \mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) . \quad (2.55)$$

Even though this vector equation is equivalent to 3 scalar equations with 3 scalar unknowns ( $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$ ), since the Euclidean norm of both the left-hand-side and the right-hand-side is always 1, the 3 scalar equations are not fully independent. As a consequence, (2.55) has, similarly to (2.31), infinite solutions for  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$ , even when  $-90^\circ < \theta_{\text{eye}} < 90^\circ$ ,  $-90^\circ < \varphi_{\text{eye}} < 90^\circ$ .

One way to solve the problem is to assume that  $\kappa_{\text{eye}} = 0$  and solve (2.55) numerically for  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ . An initial guess for  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$  can be obtained as follows. Recall that the visual axis goes through the nodal point of the eye,  $\mathbf{c}$ , and the point-of-gaze,  $\mathbf{g}$ . Since the distance between  $\mathbf{c}$  and the center of the eyeball,  $\mathbf{d}$ , is much shorter than the distance between  $\mathbf{c}$  and  $\mathbf{g}$ , the line connecting  $\mathbf{d}$  and  $\mathbf{g}$  is a reasonable approximation to the visual axis. In addition, since  $\kappa_{\text{eye}} = 0$

and  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  are relatively small, the pan and tilt angles of the visual axis with respect to the WCS can be approximated by  $(\theta_{\text{eye}} + \alpha_{\text{eye}})$  and  $(\varphi_{\text{eye}} + \beta_{\text{eye}})$ , respectively. With these approximations it follows that

$$\mathbf{v}_{\text{initial guess}} = \frac{\mathbf{g} - \mathbf{d}}{\|\mathbf{g} - \mathbf{d}\|} = \begin{bmatrix} \sin(\theta_{\text{eye, initial guess}} + \alpha_{\text{eye}}) \cos(\varphi_{\text{eye, initial guess}} + \beta_{\text{eye}}) \\ \sin(\varphi_{\text{eye, initial guess}} + \beta_{\text{eye}}) \\ -\cos(\theta_{\text{eye, initial guess}} + \alpha_{\text{eye}}) \cos(\varphi_{\text{eye, initial guess}} + \beta_{\text{eye}}) \end{bmatrix}, \quad (2.56)$$

where the right-hand-side expression has the same form as (2.26), which describes the orientation of the optic axis of the eye with respect to the WCS as a function of its pan and tilt angles,  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ , respectively.

Solving (2.56) for  $\theta_{\text{eye, initial guess}}$  and  $\varphi_{\text{eye, initial guess}}$ ,

$$\theta_{\text{eye, initial guess}} = -\arctan(v_{\text{initial guess}, X} / v_{\text{initial guess}, Z}) - \alpha_{\text{eye}}, \quad (2.57)$$

and

$$\varphi_{\text{eye, initial guess}} = \arcsin(v_{\text{initial guess}, Y}) - \beta_{\text{eye}}. \quad (2.58)$$

The orientation of the ECS with respect to the WCS can be determined without assuming that  $\kappa_{\text{eye}} = 0$  if some eye torsion model such as Listing's law is used. The determination of  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  using Listing's law is discussed in detail in Appendix A, Section A.1.1.1.

## 2.6.2 Points of reflection and refraction

### 2.6.2.1 Points of reflection and refraction on a spherical cornea

In order to find the points of reflection and refraction on a spherical cornea, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , with respect to the WCS must be known. Having determined the orientation of the ECS with respect to the WCS,  $\mathbf{c}$  is found with (2.26)-(2.27).

#### 2.6.2.1.1 Points of reflection on a spherical cornea

When the center of curvature of the cornea,  $\mathbf{c}$ , the light source,  $\mathbf{l}_i$ , and the nodal point of the camera,  $\mathbf{o}_j$ , are collinear, or when  $\mathbf{l}_i$  and  $\mathbf{o}_j$  are coincident (i.e.,  $(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{c} - \mathbf{o}_j) = \mathbf{0}$ ), the

incident ray, the reflected ray and the normal at the point of reflection are coincident (i.e., the incident ray reflects back on itself). Since, in this situation, the point of reflection,  $\mathbf{q}_{ij}$ , is collinear with  $\mathbf{c}$ ,  $\mathbf{l}_i$ , and  $\mathbf{o}_j$ , its calculation is particularly simple (see (2.36)):

$$\mathbf{q}_{ij} = \mathbf{c} + R \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \mathbf{c} + R \frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} . \quad (2.59)$$

Because  $\mathbf{q}_{ij}$  is collinear with  $\mathbf{c}$  and  $\mathbf{o}_j$ , both  $\mathbf{q}_{ij}$  and  $\mathbf{c}$  project onto the same point  $\mathbf{u}_{ij}$  on the camera image plane. This means that  $\mathbf{u}_{ij}$  can be obtained by simply projecting  $\mathbf{c}$  onto the camera image plane and, therefore, the calculation of  $\mathbf{q}_{ij}$  is not necessary to obtain  $\mathbf{u}_{ij}$ .

When  $\mathbf{c}$ ,  $\mathbf{l}_i$ , and  $\mathbf{o}_j$  are three distinct, non-collinear points, the point of reflection  $\mathbf{q}_{ij}$  can be obtained from (2.35) and (2.38)-(2.40) or from (2.41). The calculation of  $\mathbf{q}_{ij}$  in this manner, however, requires the simultaneous solution of 3 nonlinear scalar equations with 3 scalar unknowns (the 3 components of  $\mathbf{q}_{ij}$ ). The calculation of  $\mathbf{q}_{ij}$  can be made more efficient if  $\mathbf{q}_{ij}$  is expressed as a function of a single scalar parameter and the number of equations is reduced to one as follows.

The condition that points  $\mathbf{l}_i$ ,  $\mathbf{q}_{ij}$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  are coplanar (2.38) can be rewritten as

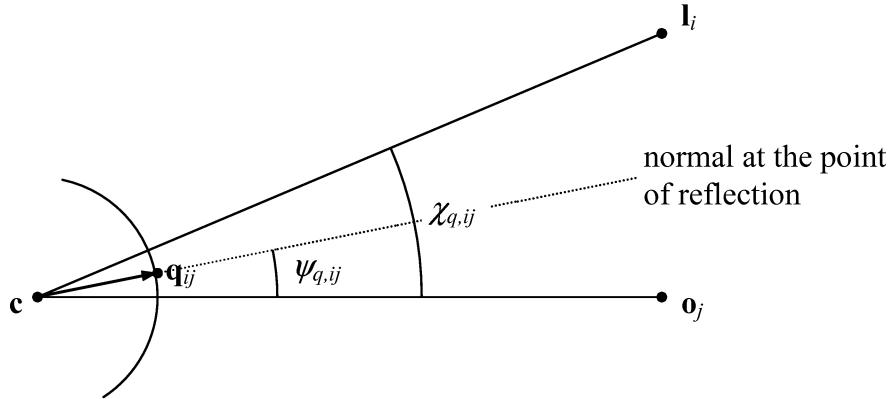
$$\mathbf{q}_{ij} - \mathbf{c} = a_{q,ij} \frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} + b_{q,ij} \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} , \quad a_{q,ij}, b_{q,ij} \in \Re . \quad (2.60)$$

The scalar factors  $a_{q,ij}$  and  $b_{q,ij}$  are found from the following conditions (Fig. 2.8):

$$(\mathbf{q}_{ij} - \mathbf{c}) \bullet \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = R \cos \psi_{q,ij} , \quad (2.61)$$

$$(\mathbf{q}_{ij} - \mathbf{c}) \bullet \frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} = R \cos(\chi_{q,ij} - \psi_{q,ij}) , \quad (2.62)$$

$$\frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} \bullet \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \cos \chi_{q,ij} . \quad (2.63)$$

Fig. 2.8: Representation of the point of reflection  $\mathbf{q}_{ij}$  in parametric form.

Substituting (2.60) into (2.61) and (2.62), and solving for  $a_{q,ij}$  and  $b_{q,ij}$ ,  $\mathbf{q}_{ij}$  is given as a function of the single scalar parameter  $\psi_{q,ij}$  as

$$\mathbf{q}_{ij}(\psi_{q,ij}) = \mathbf{c} + \frac{R}{1 - \cos^2 \chi_{q,ij}} \begin{cases} [\cos(\chi_{q,ij} - \psi_{q,ij}) - \cos \psi_{q,ij} \cos \chi_{q,ij}] \frac{\mathbf{l}_i - \mathbf{c}}{\|\mathbf{l}_i - \mathbf{c}\|} + \\ [\cos \psi_{q,ij} - \cos(\chi_{q,ij} - \psi_{q,ij}) \cos \chi_{q,ij}] \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} \end{cases}, \quad (2.64)$$

where  $\chi_{q,ij}$  is obtained from (2.63).

Substituting (2.64) into (2.39), solving numerically for  $\psi_{q,ij}$  and substituting the resulting value of  $\psi_{q,ij}$  back into (2.64), the coordinates of the point of reflection,  $\mathbf{q}_{ij}$ , are obtained. It should be noted that the equation resulting from the substitution of (2.64) into (2.39) has two solutions for  $\psi_{q,ij}$ : one valid solution and one spurious solution on the extension of the spherical corneal surface, “behind” the center of curvature of the cornea. However, it can be guaranteed that only the valid solution for  $\psi_{q,ij}$  is in the interval  $[0, \chi_{q,ij}]$ . A very good initial guess for  $\psi_{q,ij}$  is  $\chi_{q,ij}/2$ .

### 2.6.2.1.2 Points of refraction on a spherical cornea

First, recall that the coordinates of the pupil center,  $\mathbf{p}$ , with respect to the WCS are given by (2.28), whereas the coordinates of the pupil boundary points,  $\mathbf{p}_b$ , with respect to the WCS are given by (2.29). In the discussion that follows,  $\mathbf{p}$  and the term “pupil point” are used to refer to

both the pupil center and the pupil boundary points.

When the pupil point,  $\mathbf{p}$ , the center of curvature of the cornea,  $\mathbf{c}$ , and the nodal point of the camera,  $\mathbf{o}_j$ , are collinear (i.e.,  $(\mathbf{c} - \mathbf{o}_j) \times (\mathbf{p} - \mathbf{o}_j) = \mathbf{0}$ ; note that for the case of the pupil center this means that the optic axis of the eye goes through the nodal point of the camera), the incident ray, the refracted ray and the normal at the point of refraction are coincident (i.e., the incident ray crosses the cornea-air optical interface without changing its direction). Since, in this situation, the point of refraction,  $\mathbf{r}_j$ , is collinear with  $\mathbf{p}$ ,  $\mathbf{c}$ , and  $\mathbf{o}_j$ , its calculation is particularly simple (see (2.44)):

$$\mathbf{r}_j = \mathbf{c} + R \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \mathbf{c} + R \frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} . \quad (2.65)$$

Because  $\mathbf{r}_j$  is collinear with  $\mathbf{p}$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$ , the points  $\mathbf{r}_j$ ,  $\mathbf{p}$  and  $\mathbf{c}$  project onto the same point  $\mathbf{v}_j$  on the camera image plane. This means that  $\mathbf{v}_j$  can be obtained by simply projecting  $\mathbf{p}$  or  $\mathbf{c}$  onto the camera image plane and, therefore, the calculation of  $\mathbf{r}_j$  is not necessary to obtain  $\mathbf{v}_j$ .

When  $\mathbf{p}$ ,  $\mathbf{c}$ , and  $\mathbf{o}_j$  are non-collinear, the point of refraction  $\mathbf{r}_j$  can be obtained from (2.43), (2.49) and (2.50). The calculation of  $\mathbf{r}_j$  in this manner, however, requires the simultaneous solution of 3 nonlinear scalar equations with 3 scalar unknowns (the 3 components of  $\mathbf{r}_j$ ). The calculation of  $\mathbf{r}_j$  can be made more efficient if  $\mathbf{r}_j$  is expressed as a function of a single scalar parameter and the number of equations is reduced to one in a similar way as for the point of reflection.

The condition that points  $\mathbf{p}$ ,  $\mathbf{r}_j$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  are coplanar (2.45) can be rewritten as

$$\mathbf{r}_j - \mathbf{c} = a_{r,j} \frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} + b_{r,j} \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} , \quad a_{r,j}, b_{r,j} \in \Re . \quad (2.66)$$

The scalar factors  $a_{r,j}$  and  $b_{r,j}$  are found from the following conditions (Fig. 2.9):

$$(\mathbf{r}_j - \mathbf{c}) \bullet \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = R \cos \psi_{r,j} , \quad (2.67)$$

$$(\mathbf{r}_j - \mathbf{c}) \bullet \frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} = R \cos(\chi_{r,j} - \psi_{r,j}) , \quad (2.68)$$

$$\frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} \bullet \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} = \cos \chi_{r,j} . \quad (2.69)$$

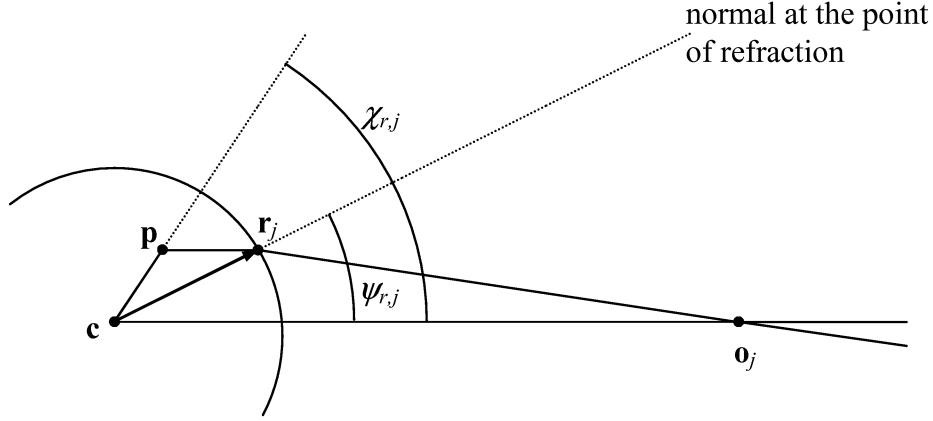


Fig. 2.9: Representation of the point of refraction  $\mathbf{r}_j$  in parametric form.

Substituting (2.66) into (2.67) and (2.68), and solving for  $a_{r,j}$  and  $b_{r,j}$ ,  $\mathbf{r}_j$  is given as a function of the single scalar parameter  $\psi_{r,j}$  as

$$\mathbf{r}_j(\psi_{r,j}) = \mathbf{c} + \frac{R}{1 - \cos^2 \chi_{r,j}} \begin{cases} [\cos(\chi_{r,j} - \psi_{r,j}) - \cos \psi_{r,j} \cos \chi_{r,j}] \frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} + \\ [\cos \psi_{r,j} - \cos(\chi_{r,j} - \psi_{r,j}) \cos \chi_{r,j}] \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} \end{cases}, \quad (2.70)$$

where  $\chi_{r,j}$  is obtained from (2.69).

The vector equation (2.49) can be transformed into a scalar equation as

$$\underbrace{\left[ \left( n_1 \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} + n_2 \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} \right) \times (\mathbf{r}_j - \mathbf{c}) \right]}_{\mathbf{v}} \bullet \underbrace{[(\mathbf{p} - \mathbf{c}) \times (\mathbf{o}_j - \mathbf{c})]}_{\mathbf{n}} = 0 , \quad (2.71)$$

where  $\mathbf{v}$  is the left-hand-side of (2.49) and  $\mathbf{n}$  is a vector normal to the plane defined by  $\mathbf{p}$ ,  $\mathbf{c}$  and  $\mathbf{o}_j$  (typically referred to as the plane of incidence). If  $\mathbf{r}_j$  is given by (2.70) (i.e., it is both in the plane

of incidence and on the corneal surface) the following can be observed: (a) when  $\mathbf{r}_j$  is a solution to (2.49),  $\mathbf{v}$  becomes the null vector and therefore  $\mathbf{r}_j$  is a solution to (2.71); (b) when  $\mathbf{r}_j$  is not a solution to (2.49), vector  $\mathbf{v}$  is parallel to  $\boldsymbol{\eta}$  but has opposite directions at each side of the solutions to (2.49). These observations imply that the left-hand-side of (2.71) with  $\mathbf{r}_j$  given by (2.70) is a scalar expression that crosses zero when  $\mathbf{r}_j$  is the actual point of refraction. Zero-crossing is a general requirement of numerical methods for solving scalar nonlinear equations in one unknown (e.g., bisection method, Van Wijngaarden-Dekker-Brent method, etc. [111]).

Substituting (2.70) into (2.71), solving numerically for  $\psi_{r,j}$  and substituting the resulting value of  $\psi_{r,j}$  back into (2.70), the coordinates of the point of refraction,  $\mathbf{r}_j$ , are obtained. It should be noted that the equation resulting from the substitution of (2.70) into (2.71) has two solutions for  $\psi_{r,j}$ : one valid solution and one spurious solution on the extension of the spherical corneal surface, “behind” the center of curvature of the cornea. However, it can be guaranteed that only the valid solution for  $\psi_{r,j}$  is in the interval  $[0, \chi_{r,j}]$ . A reasonable initial guess for  $\psi_{r,j}$  can be found as follows. First, an initial guess for the point of refraction  $\mathbf{r}_j$  is found as the intersection of the line connecting  $\mathbf{p}$  and  $\mathbf{o}_j$  with the corneal surface:

$$\begin{aligned} \mathbf{r}_{j, \text{initial guess}} &= \mathbf{p} + k_{r,j, \text{initial guess}} \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|}, \\ k_{r,j, \text{initial guess}} &= -(\mathbf{p} - \mathbf{c}) \bullet \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|} + \sqrt{\left[ (\mathbf{p} - \mathbf{c}) \bullet \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|} \right]^2 - \|\mathbf{p} - \mathbf{c}\|^2 + R^2}, \end{aligned} \quad (2.72)$$

where the expression for  $k_{r,j, \text{initial guess}}$  is derived using (2.43). Then, using (2.67) it follows that

$$\psi_{r,j, \text{initial guess}} = \arccos \left( \frac{\mathbf{r}_{j, \text{initial guess}} - \mathbf{c}}{R} \bullet \frac{\mathbf{o}_j - \mathbf{c}}{\|\mathbf{o}_j - \mathbf{c}\|} \right). \quad (2.73)$$

### 2.6.2.2 Points of reflection and refraction on an ellipsoidal cornea

The coordinates of the points of reflection and refraction on an ellipsoidal cornea could be calculated based on (2.16). However, that expression is a function of three variables ( $x_{\text{eye}}$ ,  $y_{\text{eye}}$ ,

$z_{\text{eye}}$ ). If, using (2.16),  $z_{\text{eye}}$  is expressed as a function of  $x_{\text{eye}}$  and  $y_{\text{eye}}$ , then points on the ellipsoidal cornea are defined by two variables. The limitation of such expression is that it is only real-valued when  $\gamma_1 x_{\text{eye}}^2 + \gamma_2 x_{\text{eye}} y_{\text{eye}} + \gamma_3 y_{\text{eye}}^2 \leq b^2$ , which can cause trouble during numerical optimization. That problem can be avoided if the corneal surface is described in parametric form with respect to the ECS as

$$\begin{bmatrix} x_{\text{eye}} \\ y_{\text{eye}} \end{bmatrix} = b \sin \theta_{\text{ellip}} \begin{bmatrix} \sqrt{\xi} \cos \delta & -\frac{1}{\sqrt{\xi}} \sin \delta \\ \sqrt{\xi} \sin \delta & \frac{1}{\sqrt{\xi}} \cos \delta \end{bmatrix} \begin{bmatrix} \cos \phi_{\text{ellip}} \\ \sin \phi_{\text{ellip}} \end{bmatrix}, \quad (2.74)$$

$$z_{\text{eye}} = a \cos \theta_{\text{ellip}} + A$$

where  $a, b, A, \xi$  and  $\delta$  are the parameters of the ellipsoidal corneal model (Section 2.2.3.2).

This expression is defined and real-valued for all values of the parameters  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$ . Even though this expression is periodic in both  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$ , because the period is  $2\pi$ , the periodicity does not introduce any difficulty. The coordinates of a point on the ellipsoidal cornea with respect to the WCS are obtained using (2.24). The normal to the corneal surface (i.e., the gradient) with respect to the WCS is obtained with (2.33) as  $\mathbf{R}_{\text{eye}} \nabla C_{\text{ECS}}$ , where  $\nabla C_{\text{ECS}}$  is given by (2.17) and  $\mathbf{R}_{\text{eye}}$  is given by (A.6).

### 2.6.2.2.1 Points of reflection on an ellipsoidal cornea

The calculation of the point of reflection consists of finding the values of the parameters  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$  in (2.74) such that the coordinates of the resulting point with respect to the WCS,  $\mathbf{q}_{ij}$ , satisfy (2.52). A reasonable initial guess for  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$  can be obtained by taking a spherical approximation with radius  $R_0$  (radius of curvature at the apex of the ellipsoidal cornea), whose center is in the optic axis of the eye such that the spherical approximation and the ellipsoidal surface are coincident at the apex of the cornea. Then, the center of this spherical approximation is at

$$\mathbf{c}^* = \mathbf{d} + (A + a - R_0) \mathbf{\omega}, \quad (2.75)$$

where  $\mathbf{\omega}$  is the unit vector in the direction of the optic axis of the eye and is given by (2.26).

The initial guess for  $\mathbf{q}_{ij}$  is then taken as the point on the spherical approximation such that the line defined by  $\mathbf{q}_{ij, \text{initial guess}}$  and  $\mathbf{c}^*$  bisects the angle  $\mathbf{l}_i \mathbf{c}^* \mathbf{o}_j$ :

$$\mathbf{q}_{ij, \text{initial guess}} = \mathbf{c}^* + R_0 \frac{\frac{\mathbf{l}_i - \mathbf{c}^*}{\|\mathbf{l}_i - \mathbf{c}^*\|} + \frac{\mathbf{o}_j - \mathbf{c}^*}{\|\mathbf{o}_j - \mathbf{c}^*\|}}{\left\| \frac{\mathbf{l}_i - \mathbf{c}^*}{\|\mathbf{l}_i - \mathbf{c}^*\|} + \frac{\mathbf{o}_j - \mathbf{c}^*}{\|\mathbf{o}_j - \mathbf{c}^*\|} \right\|}. \quad (2.76)$$

Next this initial guess is transformed into the ECS using (2.24) as

$$\begin{bmatrix} x_{\text{eye, initial guess}} \\ y_{\text{eye, initial guess}} \\ z_{\text{eye, initial guess}} \end{bmatrix} = \mathbf{R}_{\text{eye}}^T (\mathbf{q}_{ij, \text{initial guess}} - \mathbf{d}). \quad (2.77)$$

The initial guess for  $\theta_{\text{ellip}}$  is obtained from (2.74) as

$$\theta_{\text{ellip, initial guess}} = \arccos \frac{z_{\text{eye, initial guess}} - A}{a}. \quad (2.78)$$

If  $\theta_{\text{ellip, initial guess}} = 0$  (at the apex of the cornea), then the initial guess for  $\phi_{\text{ellip}}$  is set to 0. Otherwise, the initial guess for  $\phi_{\text{ellip}}$  is obtained as

$$\begin{aligned} \phi_{\text{ellip, initial guess}} &= \text{four-quadrant-arctan}(\varsigma_y, \varsigma_x), \\ \begin{bmatrix} \varsigma_x \\ \varsigma_y \end{bmatrix} &= \frac{1}{b \sin \theta_{\text{ellip, initial guess}}} \begin{bmatrix} \frac{1}{\sqrt{\xi}} \cos \delta & \frac{1}{\sqrt{\xi}} \sin \delta \\ -\frac{1}{\sqrt{\xi}} \sin \delta & \frac{1}{\sqrt{\xi}} \cos \delta \end{bmatrix} \begin{bmatrix} x_{\text{eye, initial guess}} \\ y_{\text{eye, initial guess}} \end{bmatrix}. \end{aligned} \quad (2.79)$$

### 2.6.2.2 Points of refraction on an ellipsoidal cornea

The calculation of the point of refraction consists of finding the values of the parameters  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$  in (2.74) such that the coordinates of the resulting point with respect to the WCS,  $\mathbf{r}_j$ , satisfy (2.54). A reasonable initial guess for  $\theta_{\text{ellip}}$  and  $\phi_{\text{ellip}}$  can be obtained by taking a spherical approximation to the corneal surface in the same way as for the point of reflection. In this case, the initial guess for  $\mathbf{r}_j$  is taken as the point of intersection of the line connecting  $\mathbf{p}$  and  $\mathbf{o}_j$  with the spherical approximation:

$$\mathbf{r}_{j, \text{initial guess}} = \mathbf{p} + k_{r,j, \text{initial guess}} \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|},$$

$$k_{r,j, \text{initial guess}} = -(\mathbf{p} - \mathbf{c}^*) \bullet \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|} + \sqrt{\left[ (\mathbf{p} - \mathbf{c}^*) \bullet \frac{\mathbf{o}_j - \mathbf{p}}{\|\mathbf{o}_j - \mathbf{p}\|} \right]^2 - \|\mathbf{p} - \mathbf{c}^*\|^2 + R_0^2}, \quad (2.80)$$

where  $\mathbf{c}^*$  is given by (2.75).

Next this initial guess,  $\mathbf{r}_{j, \text{initial guess}}$ , is transformed into the ECS using (2.77), where  $\mathbf{q}_{ij, \text{initial guess}}$  is substituted by  $\mathbf{r}_{j, \text{initial guess}}$ . Finally,  $\theta_{\text{ellip, initial guess}}$  and  $\phi_{\text{ellip, initial guess}}$  are obtained with (2.78) and (2.79).

It should be noted that the problem of the calculation of the point of refraction on an ellipsoidal cornea using (2.54) has two solutions: one valid solution and one spurious solution on the extension of the ellipsoidal corneal surface, inside the eye. The spurious solution can be avoided by setting a constraint on  $\theta_{\text{ellip}}$ , such as  $0 \leq \theta_{\text{ellip}} < \theta_{\text{ellip, MAX}}$ . The simplest constraint is obtained with  $\theta_{\text{ellip, MAX}} = 90^\circ$ , which corresponds to the front half of the ellipsoid. A tighter constraint can be set taking into account the geometrical parameters of the ellipsoidal model and the extent of the cornea as

$$\theta_{\text{ellip, MAX}} = \arccos \frac{\sqrt{b^2 - \xi \tau_{\text{MAX}}^2}}{b}, \quad (2.81)$$

where  $b$  and  $\xi$  are two of the parameters of the ellipsoidal corneal model (Section 2.2.3.2), and  $\tau_{\text{MAX}}$  represents the distance from the optic axis of the eye, which can be set to, say, 7 mm (the diameter of the cornea is 12-13 mm [27]).

### 2.6.3 Validation of the points of reflection and refraction

After finding the points of reflection and refraction on the corneal surface, it is necessary to confirm that they do not fall outside of the region that corresponds to the cornea of real eyes. A test to validate the points of reflection and refraction is to compute the distance between each of them and the optic axis of the eye and make sure that it is less than some value,  $\tau_{\text{VALID}}$ , which is

typically in the range of 5 to 6 mm (recall that the diameter of the cornea is 12-13 mm [27]) to ensure that the points of reflection and refraction are not too close to what corresponds to the cornea-sclera boundary (limbus). For example, for the points of reflection,  $\mathbf{q}_{ij}$ , this test can be formally expressed as

$$\sqrt{\|\mathbf{q}_{ij} - \mathbf{d}\|^2 - [(\mathbf{q}_{ij} - \mathbf{d}) \bullet \boldsymbol{\omega}]^2} \leq \tau_{\text{VALID}} , \quad (2.82)$$

where  $\boldsymbol{\omega}$  is the unit vector in the direction of the optic axis of the eye and is given by (2.26).

#### 2.6.4 Projection onto the camera image plane

The last step in the generation of the synthetic eye feature data is to project the points of reflection and refraction onto the camera image plane to obtain, respectively, the images of the centers of the corneal reflections, the pupil center and the pupil boundary points. In order to do this, the coordinates of the points of reflection and refraction are first transformed into the CCS using (2.25) noting that, since rotation matrices are orthonormal,  $\mathbf{R}_{\text{cam}}^{-1} = \mathbf{R}_{\text{cam}}^T$ . Next, the images of the points of reflection and refraction with respect to the CCS are obtained using (2.20). Following, if needed, the lens distortion model is applied using (2.21). Finally, the image coordinates with respect to the ICS are obtained using (2.22) or (2.23), depending on the geometry of the image sensor.

Throughout the development of the mathematical model several assumptions and simplifications were made and it is worth summarizing them at this point.

## 2.7 Summary of model assumptions and simplifications

The following assumptions and simplifications were made throughout the development of the mathematical model:

- The nodal points of the eye are assumed to be coincident, and, therefore, they are treated as a single nodal point. Furthermore, this single nodal point is assumed to be coincident with the center of curvature of the spherical corneal model.
- The center of the eyeball is assumed to be in the optic axis of the eye.

- The center of rotation of the eye is assumed to be coincident with the center of the eyeball.
- The deviation of the visual axis from the optic axis of the eye is assumed to be constant with respect to a coordinate system rigidly attached to the eye.
- The corneal models are symmetrical about the optic axis of the eye.
- The pupil is assumed to be circular. The pupil center is assumed to be in the optic axis of the eye. The plane of the pupil is assumed to be perpendicular to the optic axis of the eye.
- The refraction at the aqueous humor-cornea interface is neglected since the difference in their indices of refraction is small relative to that of the cornea-air interface. Only the refraction at the cornea-air interface is taken into account and the aqueous humor and cornea are considered as a homogeneous medium.
- The light sources are modeled as point sources.
- The camera model is the pinhole camera model extended with a lens distortion model.

In the next chapter, the estimation of the point-of-gaze using a spherical corneal model is analyzed in detail. The analysis covers the full range of possible system configurations, from the simplest, that includes one camera and one light source, to the most complex, that include multiple cameras and multiple light sources. That analysis provides fundamental insight into the limitations and potential of each configuration, including the personal calibration requirements.

## Chapter 3

# Gaze Estimation with a Spherical Corneal Model

### 3.1 Introduction

This chapter provides a detailed analysis of the point-of-gaze estimation problem using a spherical corneal model. The analysis covers the full range of possible system configurations, from the simplest, that includes one camera and one light source, to the most complex, that include multiple cameras and multiple light sources. This analysis provides fundamental insight into the limitations and potential of each configuration, including the personal calibration requirements.

For each system configuration there is more than one way to calculate the point-of-gaze from the principles of geometrical optics represented by the mathematical model developed in Chapter 2. This is because there is more than one way to express the principles of geometrical optics as equations, and for a given system of nonlinear equations there is often more than one way of solving them. Most importantly, each way of calculating the point-of-gaze can exhibit a significantly different performance in non-ideal conditions (e.g., noise in the image coordinates of the eye features, corneal asphericity, errors in system parameters).

Although the analysis from this chapter is rather lengthy, it is probably the most complete analysis of the problem of estimating the point-of-gaze from the pupil and corneal reflections available to date (an earlier version of this work can be found in [84]). It includes concepts developed throughout my M.A.Sc. and Ph.D. programs as well as concepts published by other researchers between 2000 and 2008.

Once the eye features (corneal reflections, pupil center / pupil boundary points) have been extracted from the images captured by the camera(s) and their coordinates transformed into the WCS, the gaze estimation problem can be divided into three steps. The first step consists of the determination of the 3-D coordinates of the center of curvature of the cornea using the corneal reflections. Having found the coordinates of the center of curvature of the cornea, the second step

consists of the reconstruction of the optic axis of the eye in 3-D space using the pupil center / pupil boundary points. The third step consists of the reconstruction of the visual axis from the optic axis of the eye and the determination of the point-of-gaze as the intersection of the visual axis with the scene. The methods for the first step depend on the number of cameras and the number of light sources, whereas the methods for the second step depend only on the number of cameras and the method for the third step is independent of the number of cameras and the number of light sources.

### 3.2 Determination of the 3-D coordinates of the center of curvature of the cornea

The equations that are relevant to the determination of the 3-D coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , are provided in Section 2.5.4.1 and are repeated here for convenience, either in their original form or with some modifications that make them more suitable for the problem at hand. As in Section 2.5.4.1, two situations are considered. The first, particular situation is when the incident ray (ray from light source  $i$ ,  $\mathbf{l}_i$ , to the point of reflection on the corneal surface,  $\mathbf{q}_{ij}$  –see Fig. 2.7), the reflected ray (ray from  $\mathbf{q}_{ij}$  that goes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and intersects the camera image plane at  $\mathbf{u}_{ij}$ ) and the normal at the point of reflection (line defined by  $\mathbf{c}$  and  $\mathbf{q}_{ij}$ ) are coincident (guaranteed when  $\mathbf{l}_i$  is located at  $\mathbf{o}_j$ ). In this case (2.36) is rewritten as

$$\mathbf{c} = \mathbf{o}_j + k_{c,ij} \frac{\mathbf{o}_j - \mathbf{u}_{ij}}{\|\mathbf{o}_j - \mathbf{u}_{ij}\|} \text{ for some } k_{c,ij}, \quad (3.1)$$

which defines  $\mathbf{c}$  up to the factor  $k_{c,ij}$  that represents the distance between the center of curvature of the cornea,  $\mathbf{c}$ , and the nodal point of the camera,  $\mathbf{o}_j$ .

The second, more general situation is when the incident ray, the reflected ray and the normal at the point of reflection are not coincident. In such case, (2.34), (2.35) and (2.38)-(2.40) or (2.34) and (2.41) are needed to solve the problem. Equation (2.34) is repeated here with a slight modification as

$$\mathbf{q}_{ij} = \mathbf{o}_j + k_{q,ij} \frac{\mathbf{o}_j - \mathbf{u}_{ij}}{\|\mathbf{o}_j - \mathbf{u}_{ij}\|} \text{ for some } k_{q,ij}, \quad (3.2)$$

where  $k_{q,ij}$  represents the distance between the point of reflection,  $\mathbf{q}_{ij}$ , and the nodal point of the camera,  $\mathbf{o}_j$ .

Equation (2.35) is repeated here without changes:

$$\|\mathbf{q}_{ij} - \mathbf{c}\| = R. \quad (3.3)$$

After substituting (2.34) into (2.38), (2.38) becomes

$$\underbrace{(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{u}_{ij} - \mathbf{o}_j)}_{\substack{\text{normal to the plane defined} \\ \text{by } \mathbf{l}_i, \mathbf{o}_j \text{ and } \mathbf{u}_{ij}}} \bullet (\mathbf{c} - \mathbf{o}_j) = 0. \quad (3.4)$$

This equation means that the center of curvature of the cornea,  $\mathbf{c}$ , belongs to each plane defined by the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , light source  $i$ ,  $\mathbf{l}_i$ , and the corresponding image of the corneal reflection center,  $\mathbf{u}_{ij}$ . Moreover, for each camera  $j$ , all those planes contain the line defined by points  $\mathbf{c}$  and  $\mathbf{o}_j$ .

Equation (2.39) is rewritten, after rearranging terms, as

$$\left( \frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} - \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|} \right) \bullet (\mathbf{q}_{ij} - \mathbf{c}) = 0. \quad (3.5)$$

The condition (2.40) is repeated here without any change:

$$\|\mathbf{q}_{ij} - \mathbf{o}_j\| < \|\mathbf{c} - \mathbf{o}_j\|. \quad (3.6)$$

Equation (2.41) is rewritten as

$$\mathbf{c} = \mathbf{c}_{ij} = \mathbf{q}_{ij} - R \frac{\frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|}}{\left\| \frac{\mathbf{l}_i - \mathbf{q}_{ij}}{\|\mathbf{l}_i - \mathbf{q}_{ij}\|} + \frac{\mathbf{o}_j - \mathbf{q}_{ij}}{\|\mathbf{o}_j - \mathbf{q}_{ij}\|} \right\|}. \quad (3.7)$$

The following sections discuss the determination of the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , for different system configurations.

### 3.2.1 One camera and one light source

A single camera and a single light source constitute the simplest possible system configuration. When the incident ray, the reflected ray and the normal at the point of reflection are coincident (e.g., the light source,  $\mathbf{l}_1$ , is located at the nodal point of the camera,  $\mathbf{o}_1$ )<sup>6</sup>, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , are related to the image of the corneal reflection center in the image plane of the camera,  $\mathbf{u}_{11}$ , by (3.1) with  $i = 1$  and  $j = 1$ , which defines  $\mathbf{c}$  up to the unknown factor  $k_{c,11}$ . Equation (3.1) with  $i = 1$  and  $j = 1$  is a vector equation that is equivalent to 3 scalar equations with 4 scalar unknowns (the 3 components of  $\mathbf{c}$  and the factor  $k_{c,11}$ ), and, therefore, there are more unknowns than equations.

When the incident ray, the reflected ray and the normal at the point of reflection are not coincident, the problem of determining the coordinates of  $\mathbf{c}$  is described by the system of equations constituted by (3.2)-(3.5) with constraint (3.6) or by the system of equations constituted by (3.2) and (3.7), with  $i = 1$  and  $j = 1$ . If  $R$  is known (e.g., obtained through a personal calibration procedure), either system of equations is equivalent to 6 scalar equations with 7 scalar unknowns (the 3 components of  $\mathbf{c}$ , the 3 components of  $\mathbf{q}_{11}$ , and the scale factor  $k_{q,11}$ ), and, therefore, there are, again, more unknowns than equations.

This analysis shows that using a single camera and a single light source it is not possible to determine the coordinates of  $\mathbf{c}$  from the corneal reflection unless another condition is introduced. Such condition could be that the distance between  $\mathbf{c}$  and the nodal point of the camera ( $\|\mathbf{c} - \mathbf{o}_1\|$ )

<sup>6</sup> Explanations about how this is accomplished in practice are provided in footnote 2 and in Section 2.5.4.1 immediately after (2.36).

is known (note that this is the factor  $k_{c,ij}$  in (3.1)). If no additional device is used, this condition generally means that the head must be fixed relative to the system. In such case, the distance between  $\mathbf{c}$  and  $\mathbf{o}_1$  could be estimated, at least approximately, through direct measurement with a ruler or as part of a personal calibration procedure. If an additional device such as an ultrasonic transducer or an auto-focus lens system is used to measure the distance between the eye and the camera, then the coordinates of  $\mathbf{c}$  can be estimated under head movements. A good example of this is given in [77], which uses an auto-focus lens system.

In summary, using a single camera and a single light source, the coordinates of  $\mathbf{c}$  can be estimated from the corneal reflection only if the head is stationary with respect to the system or if an additional device is used to provide further information about the relative position of the eye with respect to the system. This conclusion is consistent with the observations in Section 1.1.4.2.

The next section presents the simplest system configuration that allows for the estimation of the 3-D coordinates of  $\mathbf{c}$  without any constraints on head movements, using only the corneal reflections extracted from the images captured by the camera.

### 3.2.2 One camera and two light sources

Assuming that the incident ray, the reflected ray and the normal at the point of reflection are not coincident for either light source<sup>7</sup>, the problem of determining the coordinates of  $\mathbf{c}$  with one camera and two light sources is described by the system of equations constituted by (3.2)-(3.5) with constraint (3.6) or by the system of equations constituted by (3.2) and (3.7), with  $i = 1, 2$  and  $j = 1$  (since there is a single camera, the subscript  $j$  that identifies the camera is dropped in this section for simplicity of notation). The effective number of equations depends on whether the plane defined by  $\mathbf{o}$  ( $= \mathbf{o}_1$ ),  $\mathbf{l}_1$ , and  $\mathbf{u}_1$  ( $= \mathbf{u}_{11}$ ) and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  ( $= \mathbf{u}_{21}$ ) (see (3.4)) are distinct or coincident. If those two planes are distinct and  $R$  is known (e.g., obtained through a personal calibration procedure), either system of equations is equivalent to 12 scalar equations with 11 scalar unknowns (the 3 components of  $\mathbf{c}$ , the 3 components of  $\mathbf{q}_1$  ( $= \mathbf{q}_{11}$ ), the 3 components of  $\mathbf{q}_2$  ( $= \mathbf{q}_{21}$ ) and the scale factors  $k_{q,1}$  ( $= k_{q,11}$ ) and  $k_{q,2}$  ( $= k_{q,21}$ )). If the two planes are coincident (guaranteed if  $\mathbf{l}_1$ ,  $\mathbf{l}_2$ , and  $\mathbf{o}$  are collinear), then the effective number of equations is

<sup>7</sup> The case when the incident ray, the reflected ray and the normal at the point of reflection are coincident for one of the light sources is left for the interested reader.

reduced to 11 ((3.4) with  $i = 1$  and (3.4) with  $i = 2$  become the same equation). In such case, if  $R$  is known, there are 11 scalar equations with 11 scalar unknowns. This analysis shows that regardless of the relative positions of the camera and the light sources, if  $R$  is known, one camera and two light sources provide enough equations for the determination of the coordinates of  $\mathbf{c}$  using only the corneal reflections extracted from the images captured by the camera.

Since when the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  are distinct and  $R$  is known, the number of scalar equations exceeds the number of scalar unknowns by one, it seems that it might be possible to determine  $\mathbf{c}$  without knowing  $R$ . However, as it is shown in the example in Appendix D.1, it is not always possible to determine  $\mathbf{c}$  without knowing  $R$ . Even when it may be theoretically possible to determine  $\mathbf{c}$  without knowing  $R$ , as discussed in Appendix D.1 and Section 3.5, this is not recommended in practice. The determination of  $\mathbf{c}$  assuming that  $R$  has been obtained through a personal calibration procedure is discussed next.

When the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  are distinct, it follows from the interpretation of (3.4) that the center of curvature of the cornea,  $\mathbf{c}$ , is located in the line of intersection of those planes. Since the unit vector in the direction of such line is given by

$$\mathbf{b} = \pm \frac{[(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})] \times [(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})]}{\|[(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})] \times [(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})]\|}, \quad (3.8)$$

where  $[(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})]$  is the normal to the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_1$ ,  $[(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})]$  is the normal to the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_2$ , and the sign of  $\mathbf{b}$  is chosen so that it points from the camera to the eye, then

$$\mathbf{c} = \mathbf{o} + k_c \mathbf{b} \text{ for some } k_c, \quad (3.9)$$

where  $k_c$  represents the distance between the center of curvature of the cornea,  $\mathbf{c}$ , and the nodal point of the camera,  $\mathbf{o}$ .

If (3.2) and (3.9) are substituted into (3.3) and the resulting expression is solved for  $k_{q,i}$  subject to constraint (3.6), the factors  $k_{q,i}$  are expressed as a function of  $k_c$  as

$$k_{q,i} = k_c \frac{\mathbf{o} - \mathbf{u}_i}{\|\mathbf{o} - \mathbf{u}_i\|} \bullet \mathbf{b} - \sqrt{k_c^2 \left( \frac{\mathbf{o} - \mathbf{u}_i}{\|\mathbf{o} - \mathbf{u}_i\|} \bullet \mathbf{b} \right)^2 - k_c^2 + R^2} . \quad (3.10)$$

By substituting (3.10) into (3.2), the point of reflection  $\mathbf{q}_i$  is expressed as a function of the parameter  $k_c$ . Substituting the resulting expression for  $\mathbf{q}_i$  and (3.9) into (3.5) results in an expression that can be solved numerically for the parameter  $k_c$  for each of the light sources ( $i = 1$  and  $i = 2$ ). The estimates of  $k_c$  obtained for each of the light sources can be averaged to improve robustness to noise in the estimated coordinates of the corneal reflections. The resulting value of  $k_c$  is substituted back into (3.9) to finally obtain the coordinates of  $\mathbf{c}$ . This approach was originally described in [78] and experimental results using this approach are presented in [78], [83]-[85].

Alternatively, the coordinates of  $\mathbf{c}$  can be obtained using (3.2) and (3.7), with  $i = 1, 2, j = 1$ . If (3.2) is substituted into (3.7),  $\mathbf{c}$  can be expressed as a function of  $k_{q,i}$  ( $= k_{q,ij}$ ) and  $R$ :  $\mathbf{c}_i(k_{q,i}, R)$  ( $= \mathbf{c}_{ij}(k_{q,ij}, R)$ ,  $j = 1$ ). Ideally,  $\mathbf{c}_1(k_{q,1}, R) = \mathbf{c}_2(k_{q,2}, R)$ . However, in the presence of noise in the coordinates of the corneal reflections extracted from the eye images captured by the camera and when the cornea is not exactly spherical, it is very likely that  $\mathbf{c}_1(k_{q,1}, R) = \mathbf{c}_2(k_{q,2}, R)$  does not hold. In order to be able to solve the problem in such situation, the problem is restated as finding the values of  $k_{q,1}$  and  $k_{q,2}$  that minimize the (squared) distance between  $\mathbf{c}_1(k_{q,1}, R)$  and  $\mathbf{c}_2(k_{q,2}, R)$ :

$$\min_{k_{q,1}, k_{q,2}} \left\| \mathbf{c}_1(k_{q,1}, R) - \mathbf{c}_2(k_{q,2}, R) \right\|^2 , \quad (3.11)$$

where  $R$  is obtained through a personal calibration procedure. Having found the optimal values of  $k_{q,1}$  and  $k_{q,2}$ ,  $\mathbf{c}$  is finally obtained as

$$\mathbf{c} = \frac{\mathbf{c}_1(k_{q,1}, R) + \mathbf{c}_2(k_{q,2}, R)}{2} . \quad (3.12)$$

When the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  are coincident, the first approach to determine  $\mathbf{c}$  (with (3.8)-(3.9)) cannot be used. However, the second approach to determine  $\mathbf{c}$  (using (3.2) and (3.7), with  $i = 1, 2, j = 1$ , (3.11)-(3.12)) is still valid. Consequently, the second approach can be used regardless of the relative positions of the camera and light sources.

In summary, one camera and two light sources constitute the simplest system configuration that allows for the estimation of the 3-D coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , without any constraints on head movements and using only the corneal reflections extracted from the images captured by the camera. Furthermore, there is no theoretical constraint on the relative positions of the camera and the light sources.

### 3.2.3 One camera and more than two light sources

Increasing the number of light sources can contribute to improve the robustness of the calculation of the coordinates of  $\mathbf{c}$ , and might allow for the removal of the need to determine the value of  $R$  through a personal calibration procedure. The methods for the calculation of  $\mathbf{c}$  using more than two light sources are fairly straightforward extensions of the methods discussed in the previous section.

It should be noted, however, that as the number of light sources / corneal reflections is increased, the chances of interference of the corneal reflections with the extraction of the pupil boundary from the images captured by the cameras is also increased. A poor extraction of the pupil boundary points can result in a significant detriment to system performance. A reasonable trade-off is to use 3-4 light sources.

It is also important to note that the removal of the need to determine the value of  $R$  through a personal calibration procedure has, if any, a very limited practical benefit, as it is discussed later in Section 3.5. So, in general, when a single camera is used, calibrating for  $R$  is recommended.

### 3.2.4 Two cameras and one light source

Although a system configuration with two cameras and one light source is of limited practical interest, it is included here for the sake of completeness. This section can be skipped without affecting the understanding of the rest of the material.

The case of two cameras and one light source has a parallel with the case of one camera and two light sources. Assuming that the incident ray, the reflected ray and the normal at the point of

reflection are not coincident for either camera<sup>8</sup>, the problem of determining the coordinates of  $\mathbf{c}$  with two cameras and one light source is described by the system of equations constituted by (3.2)-(3.5) with constraint (3.6) or by the system of equations constituted by (3.2) and (3.7), with  $i = 1$  and  $j = 1, 2$  (since there is a single light source, the subscript  $i$  that identifies the light source is dropped in this section for simplicity of notation). The effective number of equations depends on whether the plane defined by  $\mathbf{l} (= \mathbf{l}_1)$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1 (= \mathbf{u}_{11})$  and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2 (= \mathbf{u}_{12})$  (see (3.4)) are distinct or coincident. If those two planes are distinct and  $R$  is known (e.g., obtained through a personal calibration procedure), either system of equations is equivalent to 12 scalar equations with 11 scalar unknowns (the 3 components of  $\mathbf{c}$ , the 3 components of  $\mathbf{q}_1 (= \mathbf{q}_{11})$ , the 3 components of  $\mathbf{q}_2 (= \mathbf{q}_{12})$  and the scale factors  $k_{q,1} (= k_{q,11})$  and  $k_{q,2} (= k_{q,12})$ ). If the two planes are coincident (guaranteed if  $\mathbf{o}_1$ ,  $\mathbf{o}_2$ , and  $\mathbf{l}$  are collinear), then the effective number of equations is reduced to 11 ((3.4) with  $j = 1$  and (3.4) with  $j = 2$  become the same equation). In such case, if  $R$  is known, there are 11 scalar equations with 11 scalar unknowns. This analysis shows that regardless of the relative positions of the cameras and the light source, if  $R$  is known, two cameras and one light source provide enough equations for the determination of the coordinates of  $\mathbf{c}$  using only the corneal reflection extracted from the images captured by the cameras.

Since when the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2$  are distinct and  $R$  is known, the number of scalar equations exceeds the number of scalar unknowns by one, it seems that it might be possible to determine  $\mathbf{c}$  without knowing  $R$ . However, as it is shown in the example in Appendix D.2, it is not always possible to determine  $\mathbf{c}$  without knowing  $R$ . Even when it may be theoretically possible to determine  $\mathbf{c}$  without knowing  $R$ , as discussed in Appendix D.2 and Section 3.5, this is not recommended in practice. The determination of  $\mathbf{c}$  assuming that  $R$  has been obtained through a personal calibration procedure is discussed next.

In this case, it is convenient to rewrite (3.4) as

$$\underbrace{(\mathbf{o}_j - \mathbf{l}_i) \times (\mathbf{u}_{ij} - \mathbf{l}_i)}_{\text{normal to the plane defined by } \mathbf{l}_i, \mathbf{o}_j \text{ and } \mathbf{u}_{ij}} \bullet (\mathbf{c} - \mathbf{l}_i) = 0 . \quad (3.13)$$

---

<sup>8</sup> The case when the incident ray, the reflected ray and the normal at the point of reflection are coincident for one of the cameras is left for the interested reader.

This equation means that the center of curvature of the cornea,  $\mathbf{c}$ , belongs to each plane defined by the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , light source  $i$ ,  $\mathbf{l}_i$ , and its corresponding image point  $\mathbf{u}_{ij}$ . Moreover, for each light source  $i$ , all those planes contain the line defined by points  $\mathbf{c}$  and  $\mathbf{l}_i$  (the subscript  $i$  that identifies the light source is dropped hereafter for simplicity of notation).

When the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2$  are distinct, it follows from the interpretation of (3.13) that the center of curvature of the cornea,  $\mathbf{c}$ , is located in the line of intersection of those planes. Since the unit vector in the direction of such line is given by

$$\mathbf{b} = \pm \frac{[(\mathbf{o}_1 - \mathbf{l}) \times (\mathbf{u}_1 - \mathbf{l})] \times [(\mathbf{o}_2 - \mathbf{l}) \times (\mathbf{u}_2 - \mathbf{l})]}{\|[(\mathbf{o}_1 - \mathbf{l}) \times (\mathbf{u}_1 - \mathbf{l})] \times [(\mathbf{o}_2 - \mathbf{l}) \times (\mathbf{u}_2 - \mathbf{l})]\|}, \quad (3.14)$$

where  $[(\mathbf{o}_1 - \mathbf{l}) \times (\mathbf{u}_1 - \mathbf{l})]$  is the normal to the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$  and  $\mathbf{u}_1$ ,  $[(\mathbf{o}_2 - \mathbf{l}) \times (\mathbf{u}_2 - \mathbf{l})]$  is the normal to the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$  and  $\mathbf{u}_2$ , and the sign of  $\mathbf{b}$  is chosen so that it points from the light source to the eye, then

$$\mathbf{c} = \mathbf{l} + k_c \mathbf{b} \text{ for some } k_c, \quad (3.15)$$

where  $k_c$  represents the distance between the center of curvature of the cornea,  $\mathbf{c}$ , and the light source,  $\mathbf{l}$ .

If (3.2) and (3.15) are substituted into (3.3) and the resulting expression is solved for  $k_{q,j}$  subject to constraint (3.6), the factors  $k_{q,j}$  are expressed as a function of  $k_c$ . By substituting the resulting expression of  $k_{q,j}$  as a function of  $k_c$  into (3.2), the point of reflection  $\mathbf{q}_j$  is expressed as a function of the parameter  $k_c$ . Substituting the resulting expression for  $\mathbf{q}_j$  and (3.15) into (3.5) results in an expression that can be solved numerically for the parameter  $k_c$  for each of the cameras ( $j = 1$  and  $j = 2$ ). The estimates of  $k_c$  obtained for each of the cameras can be averaged to improve robustness to noise in the estimated coordinates of the corneal reflection. The resulting value of  $k_c$  is substituted back into (3.15) to finally obtain the coordinates of  $\mathbf{c}$ .

Alternatively, the coordinates of  $\mathbf{c}$  can be obtained using (3.2) and (3.7), with  $i = 1, j = 1, 2$ . If (3.2) is substituted into (3.7),  $\mathbf{c}$  can be expressed as a function of  $k_{q,j}$  ( $= k_{q,ij}$ ) and  $R$ :  $\mathbf{c}_j(k_{q,j}, R)$  ( $= \mathbf{c}_{ij}(k_{q,ij}, R)$ ,  $i = 1$ ). Ideally,  $\mathbf{c}_1(k_{q,1}, R) = \mathbf{c}_2(k_{q,2}, R)$ . However, in the presence of noise in the coordinates of the corneal reflection extracted from the eye images captured by the cameras and

when the cornea is not exactly spherical, it is very likely that  $\mathbf{c}_1(k_{q,1}, R) = \mathbf{c}_2(k_{q,2}, R)$  does not hold. In order to be able to solve the problem in such situation, the problem is restated as finding the values of  $k_{q,1}$  and  $k_{q,2}$  that minimize the (squared) distance between  $\mathbf{c}_1(k_{q,1}, R)$  and  $\mathbf{c}_2(k_{q,2}, R)$  as described by (3.11), where the subscripts identify the cameras rather than the light sources. Having solved the optimization problem,  $\mathbf{c}$  is finally obtained with (3.12).

When the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2$  are coincident, the first approach to determine  $\mathbf{c}$  (with (3.14)-(3.15)) cannot be used. However, the second approach to determine  $\mathbf{c}$  (using (3.2) and (3.7), with  $i = 1, j = 1, 2$ , (3.11)-(3.12)) is still valid. Consequently, the second approach can be used regardless of the relative positions of the camera and light sources.

In summary, two cameras and one light source allow for the estimation of the 3-D coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , without any constraints on head movements, using only the corneal reflection extracted from the images captured by the cameras. Furthermore, there is no theoretical constraint on the relative positions of the cameras and the light source. Despite this, as discussed later, a system configuration with two cameras and a single light source is of limited practical interest.

### 3.2.5 Two cameras and two light sources

A system configuration with two cameras and two light sources is of particular theoretical and practical interest. Different variants are considered next depending on the relative positions of the cameras and the light sources. First, consider the particular situation when, for each camera, there is a light source such that the incident ray, the reflected ray and the normal at the point of reflection are coincident (e.g., there is one light source located at the nodal point of each camera [95], i.e.,  $\mathbf{l}_1 = \mathbf{o}_1$ ,  $\mathbf{l}_2 = \mathbf{o}_2$ ). In such case, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , satisfy (3.1),  $i = j = 1, 2$ . More specifically,  $\mathbf{c}$  is at the intersection of the line defined by (3.1) with  $i = j = 1$  and the line defined by (3.1) with  $i = j = 2$ . In practice, since the estimated coordinates of the images of the corneal reflections,  $\mathbf{u}_{ij}$ , are corrupted by noise, those lines may not intersect. Therefore,  $\mathbf{c}$  is found as the midpoint of the shortest segment defined by a point belonging to each of those lines. It can be easily shown that, in such case,  $\mathbf{c}$  is given by

$$\mathbf{c} = \frac{1}{2} [\mathbf{b}_1 \quad \mathbf{b}_2] \begin{bmatrix} \mathbf{b}_1 \bullet \mathbf{b}_1 & -\mathbf{b}_1 \bullet \mathbf{b}_2 \\ -\mathbf{b}_1 \bullet \mathbf{b}_2 & \mathbf{b}_2 \bullet \mathbf{b}_2 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{b}_1 \bullet (\mathbf{o}_1 - \mathbf{o}_2) \\ \mathbf{b}_2 \bullet (\mathbf{o}_1 - \mathbf{o}_2) \end{bmatrix} + \frac{1}{2} (\mathbf{o}_1 + \mathbf{o}_2), \quad (3.16)$$

where

$$\mathbf{b}_j = \frac{\mathbf{o}_j - \mathbf{u}_{ij}}{\|\mathbf{o}_j - \mathbf{u}_{ij}\|}, i = j. \quad (3.17)$$

Next, consider the more general situation when the incident ray, the reflected ray and the normal at the point of reflection are not coincident for any camera-light source combination<sup>9</sup> [76], [80], [84], [86], [87], [94]. In such case, if for each camera  $j$ , the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ , and the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$  are not coincident, it follows from the interpretation of (3.4) that the center of curvature of the cornea,  $\mathbf{c}$ , is located in the line of intersection of those planes. Since, for each camera  $j$ , the unit vector in the direction of such line is given by

$$\mathbf{b}_j = \pm \frac{[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)] \times [(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]}{\|[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)] \times [(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]\|}, \quad (3.18)$$

where  $[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)]$  is the normal to the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ ,  $[(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]$  is the normal to the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$ , and the sign of  $\mathbf{b}_j$  is chosen so that it points from camera  $j$  to the eye, then

$$\mathbf{c} = \mathbf{o}_j + k_{c,j} \mathbf{b}_j \text{ for some } k_{c,j}, \quad (3.19)$$

where  $k_{c,j}$  represents the distance between the center of curvature of the cornea,  $\mathbf{c}$ , and the nodal point of camera  $j$ ,  $\mathbf{o}_j$  (note that (3.18)-(3.19) is a generalization of (3.8)-(3.9)). Having two cameras, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , can be found as the intersection of the line defined by (3.18)-(3.19) with  $j = 1$  and the line defined by (3.18)-(3.19) with  $j = 2$ . Similarly to the previous case,  $\mathbf{c}$  is found using (3.16) with  $\mathbf{b}_j$  given by (3.18) [86], [87].

---

<sup>9</sup> The mixed case in which the incident ray, the reflected ray and the normal at the point of reflection are coincident for one camera-light source combination and not coincident for all others is left for the interested reader.

Alternatively, (3.4) with  $i = 1, 2$  and  $j = 1, 2$ , can be written in matrix form [84] as

$$\underbrace{\begin{bmatrix} [(\mathbf{l}_1 - \mathbf{o}_1) \times (\mathbf{u}_{11} - \mathbf{o}_1)]^T \\ [(\mathbf{l}_1 - \mathbf{o}_2) \times (\mathbf{u}_{12} - \mathbf{o}_2)]^T \\ [(\mathbf{l}_2 - \mathbf{o}_1) \times (\mathbf{u}_{21} - \mathbf{o}_1)]^T \\ [(\mathbf{l}_2 - \mathbf{o}_2) \times (\mathbf{u}_{22} - \mathbf{o}_2)]^T \end{bmatrix}}_{\mathbf{M}} \mathbf{c} = \underbrace{\begin{bmatrix} (\mathbf{l}_1 - \mathbf{o}_1) \times (\mathbf{u}_{11} - \mathbf{o}_1) \bullet \mathbf{o}_1 \\ (\mathbf{l}_1 - \mathbf{o}_2) \times (\mathbf{u}_{12} - \mathbf{o}_2) \bullet \mathbf{o}_2 \\ (\mathbf{l}_2 - \mathbf{o}_1) \times (\mathbf{u}_{21} - \mathbf{o}_1) \bullet \mathbf{o}_1 \\ (\mathbf{l}_2 - \mathbf{o}_2) \times (\mathbf{u}_{22} - \mathbf{o}_2) \bullet \mathbf{o}_2 \end{bmatrix}}_{\mathbf{m}}, \quad (3.20)$$

after applying the distributive property for the dot product, rearranging terms and noting that, when  $\mathbf{a}$  and  $\mathbf{b}$  are column vectors,  $\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^T \mathbf{b}$ . If  $\mathbf{M}$  has rank 3,  $\mathbf{c}$  can be obtained from (3.20) by using the left pseudoinverse of  $\mathbf{M}$  as

$$\mathbf{c} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{m}. \quad (3.21)$$

If only 3 linearly independent rows of  $\mathbf{M}$  and the corresponding rows of  $\mathbf{m}$  are considered in (3.20), then (3.21) is reduced to

$$\mathbf{c} = \mathbf{M}^{-1} \mathbf{m}. \quad (3.22)$$

Note that other expressions for  $\mathbf{c}$  can be obtained from the extension of (3.14)-(3.15) to two light sources or by writing (3.13), with  $i = 1, 2$  and  $j = 1, 2$ , in a matrix form similar to (3.20).

Another method to calculate the coordinates of  $\mathbf{c}$  that is of particular interest consists of using (3.2) and (3.7), with  $i = 1, 2, j = 1, 2$ . If (3.2) is substituted into (3.7),  $\mathbf{c}$  can be expressed as a function of  $k_{q,ij}$  and  $R$ :  $\mathbf{c}_{ij}(k_{q,ij}, R)$ . Ideally, all the  $\mathbf{c}_{ij}(k_{q,ij}, R)$  should be equal to each other. However, in the presence of noise in the coordinates of the corneal reflections extracted from the eye images captured by the cameras and when the cornea is not exactly spherical, it is very likely that the condition that all the  $\mathbf{c}_{ij}(k_{q,ij}, R)$  are equal to each other does not hold. In order to be able to solve the problem in such situation, the problem is restated as finding the value of  $R$  and the values of  $k_{q,ij}$  that minimize the sum of the squared distances between all possible pairs of  $\mathbf{c}_{ij}(k_{q,ij}, R)$ :

$$\min_{\substack{k_{q,ij} (i=1,2, j=1,2), \\ R}} \sum_{\substack{i_a \leq i_b, j_a \leq j_b \\ (i_a, j_a) \neq (i_b, j_b)}} \left\| \mathbf{c}_{i_a j_a}(k_{q,i_a j_a}, R) - \mathbf{c}_{i_b j_b}(k_{q,i_b j_b}, R) \right\|^2. \quad (3.23)$$

Since  $R$  is treated as a parameter to optimize, it does not need to be known a priori. Having obtained the optimum values of  $R$  and  $k_{q,ij}$ ,  $\mathbf{c}$  is finally obtained as the mean of all the  $\mathbf{c}_{ij}(k_{q,ij}, R)$ :

$$\mathbf{c} = \frac{1}{N} \sum_{i,j} \mathbf{c}_{ij}(k_{q,ij}, R) , \quad (3.24)$$

where  $N = 4$  for the case of two cameras and two light sources.

A generalized form of (3.23)-(3.24) can be expressed as

$$\min_{k_{q,ij}, R} \sum_{\substack{i_a \leq i_b, j_a \leq j_b \\ (i_a, j_a) \neq (i_b, j_b)}} w_{i_a j_a} w_{i_b j_b} \left\| \mathbf{c}_{i_a j_a}(k_{q,i_a j_a}, R) - \mathbf{c}_{i_b j_b}(k_{q,i_b j_b}, R) \right\|^2 \quad (3.25)$$

and

$$\mathbf{c} = \text{mean}_{i,j} (w_{ij} \mathbf{c}_{ij}(k_{q,ij}, R)) , \quad (3.26)$$

where the  $w_{ij}$  are weight factors. The use of binary weight factors to select a subset of all available corneal reflection images is discussed in Chapter 4.

If for one of the cameras the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ , and the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$  are coincident, (3.18) cannot be used. If for both cameras the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ , and the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$  are coincident (guaranteed if  $\mathbf{o}_1$ ,  $\mathbf{o}_2$ ,  $\mathbf{l}_1$  and  $\mathbf{l}_2$  are collinear), (3.20) cannot be used either since  $\mathbf{M}$  has rank 1. However, the method to solve for the coordinates of  $\mathbf{c}$  using (3.2) and (3.7), with  $i = 1, 2, j = 1, 2$ , (3.23)-(3.24) (or (3.25)-(3.26)) is still valid. Consequently, this last method can be used regardless of the relative positions of the cameras and light sources.

The preceding discussion shows that two cameras and two light sources constitute the simplest system configuration that allows for the determination of the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , without knowing  $R$  (i.e., without the need to determine the radius of curvature of the cornea through a personal calibration procedure), in the presence of head movements and regardless of the relative positions of the cameras and the light sources.

Although all the preceding methods for the determination of the 3-D coordinates of the center of curvature of the cornea can provide identical results in ideal conditions, they can exhibit significantly different behaviors in the presence of noise in the image coordinates of the corneal reflections, corneal asphericity and errors in system parameters. This is studied in detail in Chapter 4.

### 3.2.6 Two cameras and more than two light sources

Increasing the number of light sources can contribute to improve the robustness of the calculation of the coordinates of  $\mathbf{c}$ . The methods for the calculation of  $\mathbf{c}$  using more than two light sources are fairly straightforward extensions of the methods discussed in the previous section. The benefits of using more than two light sources will become apparent in Chapter 4.

It should be noted, again, that as the number of light sources / corneal reflections is increased, the chances of interference of the corneal reflections with the extraction of the pupil boundary from the images captured by the cameras is also increased. A poor extraction of the pupil boundary points can result in a significant detriment to system performance. A reasonable trade-off is to use 3-4 light sources.

### 3.2.7 Beyond two cameras (with multiple light sources)

Using more than two cameras could contribute to an improvement in the robustness of the system at the expense of a considerable increase in complexity and cost. Even though specific practical situations may benefit from the use of more cameras, from the theoretical point of view it is of limited interest. Nevertheless, if the coordinates of the center of curvature of the cornea were to be determined using more than two cameras, this would be done with a straightforward extension of the methods described in Section 3.2.5 (equations (3.20)-(3.21), or (3.23)-(3.24) or (3.25)-(3.26)).

## 3.3 Reconstruction of the optic axis of the eye

Having discussed the determination of the coordinates of the center of curvature of the cornea, the next step is to consider different methods for the reconstruction of the optic axis of the eye in 3-D space from the coordinates of the pupil center or pupil boundary points extracted from the

eye images captured by the camera(s). Three of the methods are based on the pupil center and one is based on the pupil boundary. The two methods based on the pupil center that are described in Sections 3.3.1 and 3.3.2 are meant to reconstruct the optic axis of the eye as the line defined by the center of curvature of the cornea ( $\mathbf{c}$ ) and the center of the true pupil ( $\mathbf{p}$ ) using the model from Section 2.5.4.1. Since the model from Section 2.5.4.1 explicitly considers the refraction at the front corneal surface, the methods from Sections 3.3.1 and 3.3.2 are referred to as using an “explicit refraction model”. The third method based on the pupil center, which is described in Section 3.3.3, assumes that the optic axis of the eye can be reconstructed as the line defined by the center of curvature of the cornea and the center of the entrance pupil (virtual image of the pupil created by refraction at the front corneal surface,  $\mathbf{p}_v$  –see Fig. 2.7). Given the way in which the refraction at the cornea is considered in this method, this method is referred to as using an “implicit refraction model” (note that some researchers refer to this approach as “refraction-ignored” [93]). The method based on the pupil boundary that is described in Section 3.3.4 follows the explicit refraction model from Section 2.5.4.1 to reconstruct the optic axis of the eye as the line defined by  $\mathbf{c}$  and  $\mathbf{p}$ .

The methods based on the pupil center assume that the coordinates of the image of the pupil center ( $\mathbf{v}_j$ ) are known. Since, in general, as discussed later in Section 3.3.4, the image of the pupil center does not coincide with the center of the pupil image in the images captured by the camera(s), using the center of the pupil image as an estimate of the image of the pupil center makes the methods based on the pupil center approximate. On the other hand, the method based on the pupil boundary is exact within the assumption that the corneal surface is spherical (in reality, of course, all methods are approximate to some extent since real corneas are not exactly spherical).

The methods based on the pupil center are not only more common but also simpler and significantly less computationally expensive than the method based on the pupil boundary. The significant difference in computational expense is due to the fact that the methods based on the pupil center provide closed-form solutions whereas the method based on the pupil boundary relies on numerical optimization. The methods based on the pupil center, due to their simplicity, provide valuable insight into the problem, particularly regarding the subject-specific eye parameters that need to be obtained a priori through a personal calibration procedure.

The sensitivity of the four different methods to noise in the image coordinates of the pupil-iris boundary points (the center of the pupil image is determined as the center of the ellipse that is fitted in least-squares sense to these boundary points), corneal asphericity and errors in system parameters, as well as the “inherent” bias of the methods that are based on the pupil center, are studied in detail through numerical simulations in Chapter 4 for the two-camera case.

### 3.3.1 Based on the pupil center with explicit refraction model – Method I (two cameras only)

As explained in Section 2.5.4.1, the condition of the law of refraction which states that the incident ray, the refracted ray and the normal at the point of refraction are coplanar implies that, for each camera, the pupil center ( $\mathbf{p}$ ), the point of refraction ( $\mathbf{r}_j$ ), the nodal point of the camera ( $\mathbf{o}_j$ ), the image of the pupil center ( $\mathbf{v}_j$ ), and the center of curvature of the cornea ( $\mathbf{c}$ ) are coplanar. It then follows that

$$\begin{aligned} (\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1) \bullet (\mathbf{p} - \mathbf{c}) &= 0 \\ (\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2) \bullet (\mathbf{p} - \mathbf{c}) &= 0 \end{aligned} \quad (3.27)$$

Since the optic axis of the eye is defined by points  $\mathbf{c}$  and  $\mathbf{p}$ , if  $\mathbf{c}$ ,  $\mathbf{o}_j$  and  $\mathbf{v}_j$  are not collinear for either of the two cameras (i.e.,  $[(\mathbf{o}_j - \mathbf{v}_j) \times (\mathbf{c} - \mathbf{o}_j)] \neq \mathbf{0}$ ,  $j = 1, 2$ ), these equations mean that the optic axis of the eye belongs to the plane defined by points  $\mathbf{c}$ ,  $\mathbf{o}_1$  and  $\mathbf{v}_1$  (normal given by  $[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)]$ ) and to the plane defined by points  $\mathbf{c}$ ,  $\mathbf{o}_2$  and  $\mathbf{v}_2$  (normal given by  $[(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]$ ). If, in addition, those two planes are not coincident (i.e.,  $[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)]$  and  $[(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]$  are not parallel), the optic axis of the eye is given by the line of the intersection of those two planes and its direction is given by the unit vector

$$\boldsymbol{\omega} = \pm \frac{[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)] \times [(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]}{\|[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)] \times [(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]\|}, \quad (3.28)$$

where the sign of  $\boldsymbol{\omega}$  is chosen so that it points forward (coming out of the eye through the cornea). In this way, the optic axis of the eye can be reconstructed in 3-D space as the line going through  $\mathbf{c}$  with direction given by  $\boldsymbol{\omega}$ . If  $\mathbf{c}$  is found using at least two light sources in conjunction with the two cameras, by using this method it is then possible to reconstruct the optic axis of the eye in 3-D space without actually knowing any eye parameter [76], [80], [84].

In order to obtain (3.28), it was assumed that  $[(\mathbf{o}_j - \mathbf{v}_j) \times (\mathbf{c} - \mathbf{o}_j)] \neq \mathbf{0}$ ,  $j = 1, 2$ , and that  $[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)]$  and  $[(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]$  are not parallel. If any of these conditions does not hold, then  $[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)] \times [(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)] = \mathbf{0}$  and (3.28) cannot be used. The condition that  $[(\mathbf{o}_j - \mathbf{v}_j) \times (\mathbf{c} - \mathbf{o}_j)] = \mathbf{0}$  implies that the center of curvature of the cornea,  $\mathbf{c}$ , the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and the image of the pupil center,  $\mathbf{v}_j$ , are collinear. Since any line passing through point  $\mathbf{c}$  is normal to the spherical corneal surface, the line defined by points  $\mathbf{o}_j$  and  $\mathbf{v}_j$  is normal to the corneal surface. Since points  $\mathbf{o}_j$  and  $\mathbf{v}_j$  are on the refracted ray coming from the pupil center,  $\mathbf{p}$ , the refracted ray is normal to the corneal surface (i.e., there is no refraction) and therefore point  $\mathbf{p}$  is collinear with points  $\mathbf{o}_j$ ,  $\mathbf{v}_j$  and  $\mathbf{c}$ . Since the optic axis of the eye is defined by points  $\mathbf{c}$  and  $\mathbf{p}$ , it implies that the optic axis of the eye goes through the nodal point of the camera,  $\mathbf{o}_j$ . In summary, if the optic axis of the eye passes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , then  $[(\mathbf{o}_j - \mathbf{v}_j) \times (\mathbf{c} - \mathbf{o}_j)] = \mathbf{0}$  and hence (3.28) cannot be used. Note that in this particular case, the direction of the optic axis of the eye is simply given by  $\omega = (\mathbf{v}_j - \mathbf{o}_j) / \|\mathbf{v}_j - \mathbf{o}_j\|$ .

The condition that  $[(\mathbf{o}_1 - \mathbf{v}_1) \times (\mathbf{c} - \mathbf{o}_1)]$  and  $[(\mathbf{o}_2 - \mathbf{v}_2) \times (\mathbf{c} - \mathbf{o}_2)]$  are parallel implies that points  $\mathbf{c}$ ,  $\mathbf{o}_1$ ,  $\mathbf{v}_1$ ,  $\mathbf{o}_2$  and  $\mathbf{v}_2$  are all on a single plane. Equation (3.27) implies then that the optic axis of the eye, defined by the center of curvature of the cornea,  $\mathbf{c}$ , and the pupil center,  $\mathbf{p}$ , is also in that plane. Consequently, this situation occurs when the optic axis of the eye is coplanar with the line defined by the nodal points of the cameras,  $\mathbf{o}_1$  and  $\mathbf{o}_2$ .

Since the condition that the optic axis of the eye passes through the nodal point of a camera is a particular case of the condition that the optic axis of the eye is coplanar with the line connecting the nodal points of the cameras, and since in practice the condition that the optic axis of the eye is parallel to the line connecting the nodal points of the cameras is unrealistic, the above discussion can be summarized by saying that (3.28) is valid as long as the optic axis of the eye does not intersect the line defined by the nodal points of the cameras [76], [80], [84]. This restriction does not exist in the other methods presented next.

### 3.3.2 Based on the pupil center with explicit refraction model – Method II

In this case, the reconstruction of the optic axis of the eye consists of determining first the 3-D coordinates of the pupil center,  $\mathbf{p}$ , and then reconstructing the optic axis as the line defined by  $\mathbf{c}$

and  $\mathbf{p}$ . By substituting (2.42) into (2.43) and applying constraint (2.50), it follows that the coordinates of the point of refraction,  $\mathbf{r}_j$ , are given by (2.42) with

$$k_{r,j} = \frac{-(\mathbf{o}_j - \mathbf{c}) \bullet (\mathbf{o}_j - \mathbf{v}_j) - \sqrt{[(\mathbf{o}_j - \mathbf{c}) \bullet (\mathbf{o}_j - \mathbf{v}_j)]^2 - \|\mathbf{o}_j - \mathbf{v}_j\|^2 (\|\mathbf{o}_j - \mathbf{c}\|^2 - R^2)}}{\|\mathbf{o}_j - \mathbf{v}_j\|^2}. \quad (3.29)$$

Now, let

$$\mathbf{t}_j = \frac{\mathbf{p} - \mathbf{r}_j}{\|\mathbf{p} - \mathbf{r}_j\|} \quad (3.30)$$

be the unit vector in the direction of the incident ray originating at the pupil center (see Fig. 2.7),

$$\boldsymbol{\zeta}_j = \frac{\mathbf{o}_j - \mathbf{r}_j}{\|\mathbf{o}_j - \mathbf{r}_j\|} = \frac{\mathbf{v}_j - \mathbf{o}_j}{\|\mathbf{v}_j - \mathbf{o}_j\|} \quad (3.31)$$

be the unit vector in the direction of the refracted ray passing through the nodal point of the camera, and

$$\boldsymbol{\eta}_j = \frac{\mathbf{r}_j - \mathbf{c}}{\|\mathbf{r}_j - \mathbf{c}\|} = \frac{\mathbf{r}_j - \mathbf{c}}{R} \quad (3.32)$$

be the unit vector normal to the corneal surface at the point of refraction. It can then be shown [77] that, from the law of refraction, given  $\boldsymbol{\zeta}_j$  and  $\boldsymbol{\eta}_j$ ,  $\mathbf{t}_j$  is given by

$$\mathbf{t}_j = \frac{n_2}{n_1} \left\{ \left[ \boldsymbol{\eta}_j \bullet \boldsymbol{\zeta}_j - \sqrt{\left( \frac{n_1}{n_2} \right)^2 - 1 + (\boldsymbol{\eta}_j \bullet \boldsymbol{\zeta}_j)^2} \right] \boldsymbol{\eta}_j - \boldsymbol{\zeta}_j \right\}, \quad (3.33)$$

where  $n_1$  is the effective index of refraction of the aqueous humor and cornea combined (= 1.3375 [110]) and  $n_2$  is the index of refraction of air ( $\cong 1$ ).

It then follows that

$$\mathbf{p} = \mathbf{r}_j + k_{p,j} \mathbf{t}_j \text{ for some } k_{p,j}, \quad (3.34)$$

which provides a line in which  $\mathbf{p}$  is located (more specifically, it is the incident ray in 3-D space). The final step in the calculation of the coordinates of  $\mathbf{p}$  depends on the number of cameras.

### 3.3.2.1 One camera

In the case of one camera ( $j = 1$ ), (3.34) defines  $\mathbf{p}$  up to the factor  $k_p$  ( $= k_{p,j}, j = 1$  –since there is a single camera, the subscript  $j$  that identifies the camera is dropped for simplicity of notation) and another condition is needed to uniquely determine  $\mathbf{p}$ . This condition is that the distance between the pupil center and the center of curvature of the cornea is given by  $K$  (see Section 2.2.4 and equation (2.28)), i.e.,

$$\|\mathbf{p} - \mathbf{c}\| = K . \quad (3.35)$$

Substituting (3.34) into (3.35) and solving for  $k_p$  with the constraint that

$$\|\mathbf{p} - \mathbf{o}\| < \|\mathbf{c} - \mathbf{o}\| , \quad (3.36)$$

results in

$$k_p = -(\mathbf{r} - \mathbf{c}) \bullet \mathbf{t} - \sqrt{[(\mathbf{r} - \mathbf{c}) \bullet \mathbf{t}]^2 - (R^2 - K^2)} . \quad (3.37)$$

Using this value of  $k_p$  in (3.34) gives the coordinates of  $\mathbf{p}$ . Having determined the coordinates of  $\mathbf{c}$  and  $\mathbf{p}$ , the unit vector in the direction of the optic axis of the eye is given by

$$\boldsymbol{\omega} = \frac{\mathbf{p} - \mathbf{c}}{\|\mathbf{p} - \mathbf{c}\|} . \quad (3.38)$$

In summary, using this method, it is possible to reconstruct the optic axis of the eye using a single camera as long as the subject-specific eye parameters  $R$  and  $K$  are known (i.e., estimated through a personal calibration procedure). Experimental results using this method are presented in [83]-[85].

### 3.3.2.2 Two cameras

In the case of two cameras,  $\mathbf{p}$  can be found as the intersection of the line described by (3.34) with  $j = 1$  and the line defined by (3.34) with  $j = 2$  [94]. Since, in practice, due to noise, those lines

may not intersect,  $\mathbf{p}$  is found as the midpoint of the shortest segment defined by a point belonging to each of those lines:

$$\mathbf{p} = \frac{1}{2} [\mathbf{l}_1 \quad \mathbf{l}_2] \begin{bmatrix} \mathbf{l}_1 \bullet \mathbf{l}_1 & -\mathbf{l}_1 \bullet \mathbf{l}_2 \\ -\mathbf{l}_1 \bullet \mathbf{l}_2 & \mathbf{l}_2 \bullet \mathbf{l}_2 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{l}_1 \bullet (\mathbf{r}_1 - \mathbf{r}_2) \\ \mathbf{l}_2 \bullet (\mathbf{r}_1 - \mathbf{r}_2) \end{bmatrix} + \frac{1}{2} (\mathbf{r}_1 + \mathbf{r}_2), \quad (3.39)$$

where  $\mathbf{r}_j$  is given by (2.42) and (3.29), and  $\mathbf{l}_j$  is given by (3.31)-(3.33). Having found  $\mathbf{p}$ , the unit vector in the direction of the optic axis of the eye is given by (3.38).

In this case, the only subject-specific eye parameter that needs to be known is the radius of curvature of the cornea,  $R$  (needed to calculate  $\mathbf{r}_j$  with (2.42) and (3.29)). When at least two light sources are used in conjunction with the two cameras, however,  $R$  does not need to be obtained through a personal calibration procedure: it can be obtained either during the calculation of  $\mathbf{c}$  (e.g., using (3.23)-(3.24) or (3.25)-(3.26)) or after the calculation of  $\mathbf{c}$  when  $R$  is not part of the calculations to determine  $\mathbf{c}$  (e.g., when using (3.16) and (3.18), or (3.20) and (3.21)/(3.22)). In the latter case,  $R$  can be found as follows. For the image of the corneal reflection due to light source  $i$  in camera  $j$ ,  $\mathbf{u}_{ij}$ , the corresponding point of reflection on the corneal surface,  $\mathbf{q}_{ij}$ , is given by (3.2) as a function of the scalar parameter  $k_{q,ij}$ . By substituting (3.2) into (3.5) and solving for  $k_{q,ij}$ , and then substituting the resulting value of  $k_{q,ij}$  back into (3.2), the coordinates of  $\mathbf{q}_{ij}$  are determined. Then, the radius of curvature of the cornea can be estimated for each  $i$  and  $j$  as

$$\hat{R}_{ij} = \|\mathbf{q}_{ij} - \mathbf{c}\|. \quad (3.40)$$

Finally, all the  $\hat{R}_{ij}$ 's are averaged to reduce the effect of image noise, model approximations and numerical errors:

$$R = \text{mean}_{i,j}(\hat{R}_{ij}). \quad (3.41)$$

In summary, using this method, it is possible to reconstruct the optic axis of the eye in 3-D space using two cameras and at least two light sources without the need to determine any subject-specific eye parameter through a personal calibration procedure, and without any constraint on the orientation of the optic axis of the eye.

### 3.3.3 Based on the pupil center with implicit refraction model

Due to the refraction at the cornea, the cameras observe a virtual image of the pupil (entrance pupil) rather than the true pupil. As in Section 2.5.4.1 (see also Fig. 2.7), consider an imaginary ray that comes from the pupil center,  $\mathbf{p}$ , and refracts at a point  $\mathbf{r}_j$  on the corneal surface such that the refracted ray passes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and intersects the camera image plane at a point  $\mathbf{v}_j$  (image of the pupil center). This refraction results in the formation of a virtual image of the pupil center (virtual pupil center),  $\mathbf{p}_{v,j}$ , located between the pupil center and the corneal surface, in the extension of the refracted ray:

$$\mathbf{p}_{v,j} = \mathbf{o}_j + k_{p_{v,j}} \underbrace{(\mathbf{o}_j - \mathbf{v}_j)}_{\mathbf{h}_j} \text{ for some } k_{p_{v,j}}. \quad (3.42)$$

The steps that follow in the reconstruction of the optic axis of the eye depend on the number of cameras.

#### 3.3.3.1 One camera

In the case of one camera ( $j = 1$ ), (3.42) defines  $\mathbf{p}_v$  ( $= \mathbf{p}_{v,1}$ ) up to the factor  $k_{p_v}$  ( $= k_{p_{v,1}}$  –since there is a single camera, the subscript  $j$  that identifies the camera is dropped for simplicity of notation) and another condition is needed to uniquely determine  $\mathbf{p}_v$ . This condition is that the distance between the virtual pupil center and the center of curvature of the cornea is given by  $K_v$ , that is,

$$\|\mathbf{p}_v - \mathbf{c}\| = K_v. \quad (3.43)$$

Substituting (3.42) into (3.43) and solving for  $k_{p_v}$  subject to the constraint that

$$\|\mathbf{p}_v - \mathbf{o}\| < \|\mathbf{c} - \mathbf{o}\|, \quad (3.44)$$

results in

$$k_{p_v} = \frac{-(\mathbf{o} - \mathbf{c}) \bullet (\mathbf{o} - \mathbf{v}) - \sqrt{[(\mathbf{o} - \mathbf{c}) \bullet (\mathbf{o} - \mathbf{v})]^2 - \|\mathbf{o} - \mathbf{v}\|^2 (\|\mathbf{o} - \mathbf{c}\|^2 - K_v^2)}}{\|\mathbf{o} - \mathbf{v}\|^2}. \quad (3.45)$$

Using this value of  $k_p$ , in (3.42) gives the coordinates of  $\mathbf{p}_v$ . Assuming that the virtual pupil center,  $\mathbf{p}_v$ , is in the optic axis of the eye, the unit vector in the direction of the optic axis of the eye is then given by

$$\mathbf{\omega} = \frac{\mathbf{p}_v - \mathbf{c}}{\|\mathbf{p}_v - \mathbf{c}\|} . \quad (3.46)$$

In summary, this method can be used to reconstruct the optic axis of the eye using a single camera as long as the parameter  $K_v$  is estimated through a personal calibration procedure. Even though the radius of curvature of the cornea,  $R$ , is not used for the calculation of the coordinates of  $\mathbf{p}_v$ , it is typically used in the calculation of the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ .

Despite the fact that, in strict terms, the value of  $K_v$  varies with the relative position and orientation of the eye with respect to the camera ([61], [62]) and with the pupil diameter, good experimental results using this method are presented in [78].

### 3.3.3.2 Two cameras

In strict terms, the spatial location of  $\mathbf{p}_{v,j}$  depends on the position of the nodal point of the camera,  $\mathbf{o}_j$ , relative to the eye. Therefore, in general, the spatial location of  $\mathbf{p}_{v,j}$  is slightly different for each of the two cameras. Despite this, an approximate virtual pupil center,  $\mathbf{p}_v$ , can be found as the midpoint of the shortest segment defined by a point belonging to each of the lines given by (3.42),  $j = 1, 2$ , i.e.,

$$\mathbf{p}_v = \frac{1}{2} [\mathbf{h}_1 \quad \mathbf{h}_2] \begin{bmatrix} \mathbf{h}_1 \bullet \mathbf{h}_1 & -\mathbf{h}_1 \bullet \mathbf{h}_2 \\ -\mathbf{h}_1 \bullet \mathbf{h}_2 & \mathbf{h}_2 \bullet \mathbf{h}_2 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{h}_1 \bullet (\mathbf{o}_1 - \mathbf{o}_2) \\ \mathbf{h}_2 \bullet (\mathbf{o}_1 - \mathbf{o}_2) \end{bmatrix} + \frac{1}{2} (\mathbf{o}_1 + \mathbf{o}_2) . \quad (3.47)$$

Assuming that  $\mathbf{p}_v$  is in the optic axis of the eye, the unit vector in the direction of the optic axis of the eye is the given by (3.46) [73], [86], [87], [95]. Note that the calculation of  $\mathbf{p}_v$  is independent of the calculation of  $\mathbf{c}$  and independent of any eye parameter.

In summary, using this method, it is also possible to reconstruct the optic axis of the eye in 3-D space using two cameras and at least two light sources without the need to determine any subject-specific eye parameter through a personal calibration procedure, and without any

constraint on the orientation of the optic axis of the eye.

### 3.3.4 Based on the pupil boundary with explicit refraction model

The methods discussed previously assume that the image of the pupil center,  $\mathbf{v}_j$ , can be determined. In practice, however,  $\mathbf{v}_j$  is estimated as the center of the pupil image in the camera image plane. As pointed out in [77], [81], [90], in general, due to the refraction at the front corneal surface followed by the perspective projection at the camera, the image of the pupil center does not coincide with the center of the pupil image. Consequently, using the center of the pupil image in place of the image of the pupil center results in point-of-gaze estimation bias.

The method discussed next overcomes this problem. It relies on the reconstruction of the pupil boundary in 3-D space from the pupil boundary points extracted from the images captured by the camera(s), under the assumptions that the front corneal surface is spherical, the pupil is circular and the optic axis of the eye contains the pupil center and is perpendicular to the plane of the pupil.

Let  $\mathbf{p}_{b,jk}$  be a point in the pupil boundary such that a ray of light originating from it refracts at the front corneal surface at point  $\mathbf{r}_{b,jk}$ , goes through the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , and intersects the camera image plane at point  $\mathbf{v}_{b,jk}$ , the  $k$ -th boundary point of the pupil image captured by camera  $j$ . Given  $\mathbf{c}$ ,  $\mathbf{o}_j$  and  $\mathbf{v}_{b,jk}$ ,  $\mathbf{r}_{b,jk}$  can be obtained as

$$\begin{aligned} \mathbf{r}_{b,jk} &= \mathbf{o}_j + k_{r_{b,jk}}(\mathbf{o}_j - \mathbf{v}_{b,jk}) , \\ k_{r_{b,jk}} &= \frac{-(\mathbf{o}_j - \mathbf{c}) \bullet (\mathbf{o}_j - \mathbf{v}_{b,jk}) - \sqrt{[(\mathbf{o}_j - \mathbf{c}) \bullet (\mathbf{o}_j - \mathbf{v}_{b,jk})]^2 - \|\mathbf{o}_j - \mathbf{v}_{b,jk}\|^2 (\|\mathbf{o}_j - \mathbf{c}\|^2 - R^2)}}{\|\mathbf{o}_j - \mathbf{v}_{b,jk}\|^2} , \end{aligned} \quad (3.48)$$

which is a generalized version of (2.42) and (3.29).

Now, let

$$\mathbf{l}_{b,jk} = \frac{\mathbf{p}_{b,jk} - \mathbf{r}_{b,jk}}{\|\mathbf{p}_{b,jk} - \mathbf{r}_{b,jk}\|} \quad (3.49)$$

be the unit vector in the direction of the incident ray originating at the pupil boundary point,

$$\zeta_{b,jk} = \frac{\mathbf{o}_j - \mathbf{r}_{b,jk}}{\|\mathbf{o}_j - \mathbf{r}_{b,jk}\|} = \frac{\mathbf{v}_{b,jk} - \mathbf{o}_j}{\|\mathbf{v}_{b,jk} - \mathbf{o}_j\|} \quad (3.50)$$

be the unit vector in the direction of the refracted ray passing through the nodal point of the camera, and

$$\eta_{b,jk} = \frac{\mathbf{r}_{b,jk} - \mathbf{c}}{\|\mathbf{r}_{b,jk} - \mathbf{c}\|} = \frac{\mathbf{r}_{b,jk} - \mathbf{c}}{R} \quad (3.51)$$

be the unit vector normal to the corneal surface at the point of refraction. It then follows that, given  $\zeta_{b,jk}$  and  $\eta_{b,jk}$ ,  $\mathbf{l}_{b,jk}$  is given by

$$\mathbf{l}_{b,jk} = \frac{n_2}{n_1} \left\{ \left[ \eta_{b,jk} \bullet \zeta_{b,jk} - \sqrt{\left( \frac{n_1}{n_2} \right)^2 - 1 + (\eta_{b,jk} \bullet \zeta_{b,jk})^2} \right] \eta_{b,jk} - \zeta_{b,jk} \right\}. \quad (3.52)$$

Note that (3.49)-(3.52) are generalized versions of (3.30)-(3.33). Having  $\mathbf{r}_{b,jk}$  and  $\mathbf{l}_{b,jk}$ ,  $\mathbf{p}_{b,jk}$  is given by

$$\mathbf{p}_{b,jk} = \mathbf{r}_{b,jk} + k_{p_{b,jk}} \mathbf{l}_{b,jk} \text{ for some } k_{p_{b,jk}}. \quad (3.53)$$

Since the plane of the pupil is perpendicular to the optic axis of the eye and the distance between the center of curvature of the cornea,  $\mathbf{c}$ , and the pupil plane is  $K$  (see Section 2.2.4 and equation (2.28)),  $\mathbf{p}_{b,jk}$  has to satisfy

$$(\mathbf{p}_{b,jk} - \mathbf{c}) \bullet \boldsymbol{\omega} = K, \quad (3.54)$$

where  $\boldsymbol{\omega}$  is the unit vector in the direction of the optic axis of the eye (Section 2.5.3; equation (2.26)). Substituting (3.53) into (3.54), it follows that

$$k_{p_{b,jk}} = \frac{K - (\mathbf{r}_{b,jk} - \mathbf{c}) \bullet \boldsymbol{\omega}}{\mathbf{l}_{b,jk} \bullet \boldsymbol{\omega}}. \quad (3.55)$$

Since the pupil is assumed to be circular,  $\mathbf{p}_{b,jk}$  also has to satisfy

$$\|\mathbf{p}_{b,jk} - \mathbf{p}\| = r_{\text{pupil}} , \quad (3.56)$$

where the pupil center,  $\mathbf{p}$ , is given by (2.28) and  $r_{\text{pupil}}$  is the radius of the pupil.

The problem of reconstructing the optic axis of the eye consists then of finding  $\omega \equiv \omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$  (see (2.26)) such that all the  $\mathbf{p}_{b,jk}$ 's obtained using (3.48), (3.50)-(3.53) and (3.55) satisfy (3.56) for all  $j$  and all  $k$ . Since in practice the coordinates of  $\mathbf{v}_{b,jk}$ , and therefore the coordinates of  $\mathbf{r}_{b,jk}$  and the components of  $\mathbf{l}_{b,jk}$ , are corrupted by noise, the condition expressed by equation (3.56) for all  $j$  and all  $k$  has to be turned into a minimization problem, whose actual expression depends on the number of cameras. For the same reason, even though, theoretically, three pupil boundary points are enough to reconstruct the pupil (three non-collinear points define a circle), many more pupil boundary points need to be used in practice.

Note that the preceding system of equations depends on the subject-specific eye parameters  $R$  and  $K$ , as well as the time-varying parameter  $r_{\text{pupil}}$ . The need for the a priori determination of  $R$  and  $K$  through a personal calibration procedure also depends on the system configuration, as it is discussed next.

### 3.3.4.1 One camera

For the case of a single camera ( $j = 1$ ) the subscript that identifies the camera can be dropped for simplicity of notation ( $\mathbf{p}_{b,k} = \mathbf{p}_{b,1k}$ ,  $\mathbf{v}_{b,k} = \mathbf{v}_{b,1k}$ ,  $\mathbf{r}_{b,k} = \mathbf{r}_{b,1k}$ ,  $\mathbf{l}_{b,k} = \mathbf{l}_{b,1k}$ ) and the orientation of the optic axis of the eye,  $\omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$ , is found by solving the following minimization problem:

$$\begin{aligned} \min_{\theta_{\text{eye}}, \varphi_{\text{eye}}, r_{\text{pupil}}} & \sum_k \left( \|\mathbf{p}_{b,k}(\theta_{\text{eye}}, \varphi_{\text{eye}}, K) - \mathbf{p}(\theta_{\text{eye}}, \varphi_{\text{eye}}, K) \| - r_{\text{pupil}} \right)^2 = \\ \min_{\theta_{\text{eye}}, \varphi_{\text{eye}}, r_{\text{pupil}}} & \sum_k \left( \left\| \left( \mathbf{r}_{b,k} - \mathbf{c} \right) + \frac{K - (\mathbf{r}_{b,k} - \mathbf{c}) \bullet \omega(\theta_{\text{eye}}, \varphi_{\text{eye}})}{\mathbf{l}_{b,k} \bullet \omega(\theta_{\text{eye}}, \varphi_{\text{eye}})} \mathbf{l}_{b,k} - K \omega(\theta_{\text{eye}}, \varphi_{\text{eye}}) \right\| - r_{\text{pupil}} \right)^2 \end{aligned} \quad (3.57)$$

where the last expression is obtained using (2.28), (3.53) and (3.55). Note that the optimization over  $r_{\text{pupil}}$  can be avoided by substituting

$$r_{\text{pupil}} = \frac{1}{N_{\text{boundary points}}} \sum_k \|\mathbf{p}_{b,k} - \mathbf{p}\| \quad (3.58)$$

into (3.57).

To solve the above minimization problem,  $R$  and  $K$  have to be obtained a priori through a personal calibration procedure ( $R$  is needed to obtain the  $\mathbf{r}_{b,k}$ 's for (3.57) using (3.48)). Even though when multiple light sources are used it might be theoretically possible, under certain conditions, to determine the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , without the a priori knowledge of the value of  $R$  and then estimate  $R$  either during or after the calculation of  $\mathbf{c}$ , not calibrating for  $R$  is not recommended in practice.

Note that the essence of this method is that of the method presented in [90]-[93].

### 3.3.4.2 Two cameras

For the case of two cameras ( $j = 1, 2$ ) the orientation of the optic axis of the eye ( $\omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$ ) is found by solving the following minimization problem:

$$\begin{aligned} \min_{\theta_{\text{eye}}, \varphi_{\text{eye}}, K, r_{\text{pupil}}} & \sum_j \sum_k \left( \|\mathbf{p}_{b,jk}(\theta_{\text{eye}}, \varphi_{\text{eye}}, K) - \mathbf{p}(\theta_{\text{eye}}, \varphi_{\text{eye}}, K)\| - r_{\text{pupil}} \right)^2 = \\ \min_{\theta_{\text{eye}}, \varphi_{\text{eye}}, K, r_{\text{pupil}}} & \sum_j \sum_k \left( \left\| \left( \mathbf{r}_{b,jk} - \mathbf{c} \right) + \frac{K - (\mathbf{r}_{b,jk} - \mathbf{c}) \bullet \omega(\theta_{\text{eye}}, \varphi_{\text{eye}})}{\mathbf{l}_{b,jk} \bullet \omega(\theta_{\text{eye}}, \varphi_{\text{eye}})} \mathbf{l}_{b,jk} - K \omega(\theta_{\text{eye}}, \varphi_{\text{eye}}) \right\| - r_{\text{pupil}} \right)^2 \end{aligned} \quad (3.59)$$

where the last expression is obtained using (2.28), (3.53) and (3.55). Note that, in this case, the optimization over  $r_{\text{pupil}}$  can be avoided by substituting

$$r_{\text{pupil}} = \frac{1}{N_{\text{boundary points}} \text{ (total of all cameras)}} \sum_j \sum_k \|\mathbf{p}_{b,jk} - \mathbf{p}\| \quad (3.60)$$

into (3.59).

Note that for the case of multiple cameras,  $K$  can be treated as a parameter to be optimized (see (3.59)) and therefore it does not need to be obtained a priori through a personal calibration procedure. In order to obtain the  $\mathbf{r}_{b,jk}$ 's for (3.59) (using (3.48)), the radius of curvature of the

cornea,  $R$ , must be known. When multiple light sources are used, the value of  $R$  is readily obtained during or after the determination of the coordinates of the center of curvature of the cornea,  $\mathbf{c}$  (see discussion after (3.39)). This implies that  $R$  does not need to be obtained a priori through a personal calibration procedure either. In summary, if at least two cameras and at least two light sources are used, the optic axis of the eye in 3-D space can be reconstructed without the need to obtain  $R$  and  $K$  through a personal calibration procedure.

### 3.3.5 Brief summary of the different methods

Regardless of which particular method is used to reconstruct the optic axis of the eye in 3-D space, the following brief, general summary can be provided. Using one camera and at least two light sources, the optic axis of the eye can be reconstructed if two subject-specific eye parameters (either  $R$  and  $K$ , or  $R$  and  $K_v$ ) are obtained a priori through a personal calibration procedure. Using two cameras and one light source the optic axis of the eye can be reconstructed if  $R$  has been obtained through personal calibration. Using two cameras and at least two light sources, the optic axis of the eye can be reconstructed without the need to obtain any subject-specific eye parameter through personal calibration.

A comparative simulation study of the different methods to reconstruct the optic axis of the eye with a single camera is presented in [93]. A comparative simulation study of the different methods to reconstruct the optic axis of the eye with two cameras is presented in Chapter 4. The simulation study in Chapter 4 focuses on the sensitivity to noise in the image coordinates of the pupil-iris boundary points (the center of the pupil image is determined as the center of the ellipse that is fitted in least-squares sense to these boundary points), corneal asphericity and errors in system parameters, as well as the “inherent” bias of the methods that are based on the pupil center.

## 3.4 Reconstruction of the visual axis and estimation of the point-of-gaze

Following the above methodologies, the pan and tilt angles of the optic axis of the eye with respect to the WCS,  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ , can be determined (if  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$  are not found directly, they can be obtained from  $\boldsymbol{\omega}$  (the unit vector in the direction of the optic axis of the eye with respect

to the WCS) and (2.26) using (A.19) and (A.20)). If the torsion of the eye about its optic axis,  $\kappa_{\text{eye}}$ , is measured independently, estimated with some eye torsion model such as Listing's law (the estimation of the point-of-gaze using Listing's law is discussed in detail in Appendix A, Sections A.1.1 and A.1.1.2), or simply assumed to be nil, then the orientation of the eye is fully specified with its 3 degrees of freedom. Having  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$ , the rotation matrix that describes the orientation of the eye with respect to the WCS,  $\mathbf{R}_{\text{eye}} \equiv \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ , can be obtained with (A.2)-(A.6). If the orientation of the visual axis with respect to the ECS (described by the horizontal and vertical angles,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , or simply by  $\mathbf{v}_{\text{ECS}}$  given by (2.1)) is known, then the unit vector in the direction of the visual axis of the eye with respect to the WCS,  $\mathbf{v}$ , is found with (2.30) and the entire visual axis in 3-D space is given by (2.31). The point-of-gaze is then given by (2.31) with a value of  $k_g$  such that the visual axis intersects the scene. In the case of a planar scene at  $Z = 0$ , that value of  $k_g$  is given by

$$k_g = -\frac{c_z}{v_z} . \quad (3.61)$$

### 3.5 Summary of the analysis of the different system configurations and discussion of their personal calibration requirements

A single camera and a single light source allow for the estimation of the point-of-gaze from the pupil and the corneal reflection only if the head is stationary with respect to the system or if an additional device (e.g., an ultrasonic transducer or an auto-focus system) is used to provide further information about the relative position of the eye with respect to the system, and as long as the subject-specific eye parameters  $R$ ,  $K$  (or  $K_v$ ),  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  are obtained through a personal calibration procedure.

One camera and two light sources constitute the simplest system configuration that allows for the estimation of the point-of-gaze without any constraints on head movements, using only the pupil and the corneal reflections extracted from the images captured by the camera. Increasing the number of light sources can contribute to improve the robustness of the point-of-gaze estimation as long as the increased number of corneal reflections does not interfere with the extraction of

the pupil-iris boundary from the images captured by the camera. To be able to estimate the point-of-gaze, the subject-specific eye parameters  $R$ ,  $K$  (or  $K_v$ ),  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  should be obtained through a personal calibration procedure. Even though it might be theoretically possible, under certain conditions, to avoid calibrating for  $R$ , this is not recommended in practice.

Two cameras and one light source also allow for the estimation of the point-of-gaze without any constraints on head movements, using only the pupil and the corneal reflection extracted from the images captured by the cameras. To be able to estimate the point-of-gaze, the subject-specific eye parameters  $R$ ,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  should be obtained through a personal calibration procedure (the need to calibrate for  $K$  (or  $K_v$ ) is removed). Even though it might be theoretically possible, under certain conditions, to avoid calibrating for  $R$ , this is not recommended in practice. By adding a second light source, the need to calibrate for  $R$  can be removed regardless of the relative positions of the eye, the cameras and the light sources. For this reason, a system configuration with two cameras and a single light source is of limited interest.

Two cameras and two light sources constitute the simplest system configuration that allows for the reconstruction of the optic axis of the eye in 3-D space in the presence of head movements, using only the pupil and the corneal reflections extracted from the images captured by the cameras and without knowing any subject-specific eye parameter. The ability to reconstruct the optic axis of the eye without the need for a personal calibration procedure is particularly important as it can enable the measurement of eye movements in situations in which subject cooperation cannot be expected (e.g., infants, mentally challenged individuals, non-human subjects) or in which subjects are unable to fixate a point accurately (e.g., Age-related Macular Degeneration) or simply cannot fixate (e.g., oscillopsia). Using two cameras and two light sources the point-of-gaze can be estimated after obtaining the subject-specific eye parameters  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  through a personal calibration procedure. Increasing the number of light sources can contribute to improve the robustness of the point-of-gaze estimation as long as the increased number of corneal reflections does not interfere with the extraction of the pupil-iris boundary from the images captured by the cameras.

A personal calibration procedure that estimates  $R$  and/or  $K$  (or  $K_v$ ),  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  (3-4 degrees of freedom; applicable to systems with a single camera or systems with two cameras and a single light source) requires the subject to fixate at least 2 known points. Since fixations are not perfect,

in practice, more calibration points should be used to obtain a reasonably robust calibration. A reasonable number of calibration points ranges from 4 to 9. In recent publications [82], [92], it is suggested that using one camera and multiple light sources it is possible to estimate the point-of-gaze after completing a calibration procedure in which the subject has to fixate a single point. However, to do that, the methodologies presented in [82] and [92] need to make assumptions that can result in significant point-of-gaze estimation errors for a large fraction of the population. In fact, it can be shown ([61], [62], [78]) that in a properly chosen Cartesian coordinate system,  $R$  and  $K$  (or  $K_v$ ) have predominantly a gain-like behavior, while  $\alpha_{\text{eye}}$  has predominantly a horizontal offset-like behavior and  $\beta_{\text{eye}}$  has predominantly a vertical offset-like behavior, supporting the notion that at least 2 calibration points are required to calibrate for these subject-specific eye parameters. Typically, to obtain  $R$  and/or  $K$  (or  $K_v$ ),  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  through a multiple-point calibration procedure, these eye parameters are optimized to minimize the sum of the square distances between the target calibration points and the estimated points-of-gaze.

A personal calibration procedure that estimates only  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  (applicable to systems with two or more cameras and two or more light sources) requires the subject to fixate a single known point. The basic procedure to find  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  is as follows. Suppose that when the subject fixates the calibration point  $\mathbf{g}$ , the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , are determined as discussed in Section 3.2.5, and the orientation of the optic axis of the eye with respect to the WCS (described either by its pan and tilt angles,  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ , or by its unit vector  $\mathbf{\omega}$ ) is found as discussed in Section 3.3 (particularly Sections 3.3.1, 3.3.2.2, 3.3.3.2 or 3.3.4.2). Having  $\mathbf{c}$  and  $\mathbf{g}$ , the unit vector in the direction of the visual axis with respect to the WCS,  $\mathbf{v}$ , is found as

$$\mathbf{v} = \frac{\mathbf{g} - \mathbf{c}}{\|\mathbf{g} - \mathbf{c}\|} . \quad (3.62)$$

If the torsion of the eye about its optic axis,  $\kappa_{\text{eye}}$ , is measured independently, estimated with some eye torsion model such as Listing's law (the single-point calibration using Listing's law is discussed in detail in Appendix A, Sections A.1.1 and A.1.1.3), or simply assumed to be nil, then the orientation of the eye is fully specified with its 3 degrees of freedom. Having  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$ , the rotation matrix that describes the orientation of the eye with respect to the WCS,  $\mathbf{R}_{\text{eye}} \equiv \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ , can be obtained with (A.2)-(A.6). Then, the unit vector in the direction of

the visual axis with respect to the ECS,  $\mathbf{v}_{\text{ECS}}$ , can be obtained from (2.30) as

$$\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = \mathbf{R}_{\text{eye}}^T(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}) \mathbf{v}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}, \alpha_{\text{eye}}, \beta_{\text{eye}}) . \quad (3.63)$$

Finally,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  can be obtained from  $\mathbf{v}_{\text{ECS}}$  and (2.1) (note that in most cases, it is enough to find  $\mathbf{v}_{\text{ECS}}$ ).

In practice, whenever multiple-point personal calibration procedures are possible, system configurations with a single camera and multiple light sources are preferred. Not only are these system configurations simpler and less costly than system configurations with multiple cameras but also the use of multiple calibration points can help to improve system robustness. Since the multiple-point personal calibration procedures are associated with more degrees of freedom in the mathematical model, the multiple-point personal calibration can help to absorb errors in model parameters [78] and deviations from modeling assumptions (especially those that have offset-like and gain-like effects). Furthermore, the multiple-point personal calibration can be exploited to fit mapping functions (e.g., a homographic transformation) to refine the point-of-gaze estimation accuracy in the presence of errors in system parameters that cannot be compensated for through the calibration of the eye model parameters and/or when there are significant deviations in the shape of the corneal surface from the ideal spherical shape assumed in the mathematical model (e.g., strong corneal asphericity, especially when it is aggravated by corneal astigmatism). A key benefit of the multiple-point personal calibration is that, in many practical situations, the requirements on the system calibration (measurement / estimation of system parameters) and certain requirements on the system geometry can be relaxed.

When multiple-point personal calibration procedures are not possible, system configurations with two (or more) cameras and multiple light sources are to be used. As mentioned previously, these system configurations allow for the reconstruction of the optic axis of the eye in 3-D space, and therefore the measurement of eye movements, without personal calibration. In order to estimate the point-of-gaze, a single-point personal calibration procedure is needed. The drawbacks of the use of multiple cameras include not only the obvious increase of system complexity and cost, but also significantly more stringent requirements on the system geometry and the system calibration. Nevertheless, systems with two cameras and multiple light sources can enable applications that are very difficult or impossible to carry out using systems that require multiple-

point personal calibration procedures. Furthermore, as it is shown later, by optimizing the geometry of the system setup and using specific point-of-gaze estimation methods, systems with two cameras and multiple light sources using a single-point personal calibration procedure can attain a point-of-gaze estimation accuracy of 0.4-0.6° of visual angle. This accuracy is comparable to the best commercially available systems, which use multiple-point personal calibration procedures. The remainder of this thesis concentrates on the estimation of the point-of-gaze using two cameras and two or more light sources.

## Chapter 4

### Simulation Study

#### 4.1 Introduction

This chapter presents a detailed simulation study of different point-of-gaze estimation methods that are based on a spherical corneal model and use two cameras and two or more light sources. The sensitivity of the point-of-gaze estimates to several sources of errors (noise in the image coordinates of the eye features, corneal asphericity, errors in system parameters) is evaluated for the different methods and for different system setups (number of light sources, relative positions of the cameras and the light sources). This simulation study is used to determine combinations of system setups and gaze estimation methods that result in improved point-of-gaze estimation robustness, thus providing guidelines for optimal system implementation.

The following section summarizes the different point-of-gaze estimation methods that are considered in the simulation study. Section 4.3 describes the general simulation methodology and Section 4.4 summarizes the model parameters that are used in the simulations. Section 4.5 presents a brief evaluation of the point-of-gaze estimation bias that is inherent to some of the point-of-gaze estimation methods due to certain model assumptions and simplifications. Section 4.6 studies the sensitivities of the different point-of-gaze estimation methods to noise in the image coordinates of the eye features, corneal asphericity and error in a selected system parameter for four different setups of a system with two cameras and two light sources (recall that, as it is shown in Chapter 3, two cameras and two light sources is the simplest system configuration that can be used to reconstruct the optic axis of the eye in 3-D space without personal calibration, and to estimate the point-of-gaze after completing a single-point personal calibration procedure). The simulation study from Section 4.6 provides several general observations about the performance of the different methods and the role of the relative positions of the cameras and the light sources. Based on these observations, Section 4.7 focuses attention on a subset of the point-of-gaze estimation methods and system setups with two cameras and four light sources (the additional light sources help to improve system robustness and reduce the

effect of corneal asphericity). Section 4.7.1 studies the sensitivities of the subset of point-of-gaze estimation methods to noise in the image coordinates of the eye features and corneal asphericity as a function of the camera separation and the position of two of the light sources. Based on the findings from Section 4.7.1 and some practical considerations, a system setup that minimizes the point-of-gaze estimation error is selected. The selected system setup is evaluated through simulations in Section 4.7.2. The simulations in Sections 4.5 through 4.7.1 are carried out for a single fixed position of the center of the eyeball to simplify the analysis of the results and keep the computational expense at a reasonable level. The simulations in Section 4.7.2 are carried out for several different positions of the center of the eyeball spanning the entire region of operation of the system setup studied in order to evaluate system performance under head movements.

## 4.2 Point-of-gaze estimation methods studied

A total of 5 methods for the calculation of the 3-D coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , and 5 methods for the subsequent reconstruction of the optic axis of the eye in 3-D space are considered. These methods are summarized in Sections 4.2.1 and 4.2.2, respectively.

### 4.2.1 Calculation of the 3-D coordinates of the center of curvature of the cornea

The calculation of the 3-D coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , using two cameras and two light sources, is described in Section 3.2.5. The 5 methods for the calculation of  $\mathbf{c}$  to be studied are the following:

- 1) Using (3.20) and (3.21). In this case,  $\mathbf{c}$  is found as the intersection (in least squares sense using a matrix left-pseudoinverse) of the 4 planes defined by the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , light source  $i$ ,  $\mathbf{l}_i$ , and its corresponding corneal reflection image,  $\mathbf{u}_{ij}$ , with  $i = 1, 2$ ,  $j = 1, 2$ . This method will thus be referred to as the “calculation of  $\mathbf{c}$  using *Plane Intersections with matrix PseudoInverse with 4 Corneal Reflection images*” (two from each camera), or, in abbreviated form, as “*CPIIntPsInv-4CR*”.
- 2) Using (3.20) with 3 rows and (3.22). In this case, which can be considered as a variant of the preceding one,  $\mathbf{c}$  is found as the intersection (using a matrix inverse) of 3 of the 4 planes defined by the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , light source  $i$ ,  $\mathbf{l}_i$ , and its corresponding corneal

reflection image,  $\mathbf{u}_{ij}$ , with  $i = 1, 2$ ,  $j = 1, 2$ . The 3 selected planes correspond to the 3 corneal reflection images that are closest to the center of the pupil image in the respective eye image (i.e., the 3  $\mathbf{u}_{ij}$ 's that have the smallest  $\|\mathbf{u}_{ij} - \hat{\mathbf{v}}_j\|$ , where  $\hat{\mathbf{v}}_j$  is the center of the pupil image in the image plane of camera  $j$ ). As it will be shown later, by selecting the corneal reflection images in this manner the point-of-gaze estimation bias due to corneal asphericity can be reduced. This is because, typically, the spherical corneal model is a good approximation to real corneas only for the central part of the cornea [27]. This method will be referred to as the “calculation of  $\mathbf{c}$  using *Plane Intersections with matrix Inverse with 3 Corneal Reflection images*”, or, in abbreviated form, as “*CPlIntInv-3CR*”.

- 3) Using (3.16) and (3.18). In this case, first, for each camera  $j$ , the line connecting the nodal point,  $\mathbf{o}_j$ , with the center of curvature of the cornea,  $\mathbf{c}$ , is found as the intersection of the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ , and the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$ . Then,  $\mathbf{c}$  is determined as the intersection of the resulting lines for  $j = 1, 2$  (triangulation technique). This method will thus be referred to as the “calculation of  $\mathbf{c}$  using *Plane Intersections and Triangulation with 4 Corneal Reflection images*” (two from each camera), or, in abbreviated form, as “*CPlIntTriang-4CR*”.
- 4) Using (3.2), (3.7), (3.23) and (3.24). In this case, the calculation of  $\mathbf{c}$  is not based on plane intersections and both corneal reflection images from both cameras are used. This method will thus be referred to as the “calculation of  $\mathbf{c}$  with *No Plane Intersections with 4 Corneal Reflection images*” (two from each camera), or, in abbreviated form, as “*CNoPliInt-4CR*”.
- 5) Using (3.2), (3.7), (3.25) and (3.26) with 3 of the 4 corneal reflection images (variant of the previous method). The 3 corneal reflection images are selected as in 2), that is, the 3 corneal reflection images that are closest to the center of the pupil image in the respective eye image. Each weight factor  $w_{ij}$  in (3.25) and (3.26) is set to 1 if the image of the corneal reflection due to light source  $i$  in camera  $j$ ,  $\mathbf{u}_{ij}$ , is used, or 0 otherwise. This method will thus be referred to as the “calculation of  $\mathbf{c}$  with *No Plane Intersections with 3 Corneal Reflection images*”, or, in abbreviated form, as “*CNoPliInt-3CR*”.

## 4.2.2 Subsequent reconstruction of the optic axis of the eye in 3-D space

The subsequent reconstruction of the optic axis of the eye in 3-D space is described in Section 3.3. The 5 methods for the reconstruction of the optic axis of the eye to be studied are as follows:

- 1) The method described in Section 3.3.1. In this case, the optic axis of the eye is found as the intersection of the plane defined by  $\mathbf{c}$ ,  $\mathbf{o}_1$  and  $\mathbf{v}_1$  and the plane defined by  $\mathbf{c}$ ,  $\mathbf{o}_2$  and  $\mathbf{v}_2$  (recall that  $\mathbf{o}_j$  is the nodal point of camera  $j$  and  $\mathbf{v}_j$  is the image of the pupil center in camera  $j$ ). This method will thus be referred to as “based on the *Pupil Center* using *Plane Intersection*”, or, in abbreviated form, as “*PCPIInt*”.
- 2) The method described in Section 3.3.2.2. In this case, the pupil center,  $\mathbf{p}$ , is found using a triangulation technique after refracting at the corneal surface the rays obtained by back-projection of the pupil center images. Then, the optic axis is reconstructed as the line defined by  $\mathbf{c}$  and  $\mathbf{p}$ . This method will thus be referred to as “based on the *Pupil Center* using *Triangulation with Refraction*”, or, in abbreviated form, as “*PCTriangRefr*”.
- 3) The method described in Section 3.3.3.2. In this case, the virtual pupil center,  $\mathbf{p}_v$ , is found through direct triangulation using the back-projection of the pupil center images. Then, the optic axis is reconstructed as the line defined by  $\mathbf{c}$  and  $\mathbf{p}_v$ . This method will thus be referred to as “based on the *Pupil Center* using direct *Triangulation*”, or, in abbreviated form, as “*PCTriang*”.
- 4) The method described in Section 3.3.4.2. In this case, the reconstruction of the optic axis of the eye is “based on the *Pupil Boundary* points extracted from the eye images”, and, therefore, this method will be referred to in abbreviated form as “*PB*”.
- 5) The method described in Section 3.3.4.2 combined with ellipse fitting. In this case, which is a variant of the previous one, instead of using the pupil boundary points extracted from the eye images directly, an ellipse is fitted (in least squares sense) to the pupil boundary points extracted from each image and a set of points sampled at regular intervals from the fitted ellipse is then used for the reconstruction of the optic axis of the eye. This method will thus be referred to as “based on the *Pupil Boundary* with *Ellipse Fitting*”, or, in abbreviated form as “*PB-EF*”. The interest in this method stems from the facts that (a) it can be more convenient to record only 5 ellipse parameters for each pupil image rather than a variable

number of pupil boundary points (the number of pupil boundary points varies with the pupil diameter and the relative position and orientation of the eye with respect to each camera, among other factors); and (b) even though the perspective projection of a circle is an ellipse, in general, there is no guarantee that after refraction at the cornea the image of a circular pupil is truly elliptical (the study of this method can therefore provide an indirect evaluation of the assumption that the pupil image is elliptical). Note that the center of the fitted ellipse is taken as the center of the pupil image, which, in turn, is used as an estimate of the image of the pupil center for the methods that are based on the pupil center (“PC...”) –this is usual practice in the gaze estimation field.

### 4.3 General simulation methodology

Given a set of eye parameters, the system parameters (position of the light sources, camera parameters) and the 3-D positions of the center of the eyeball,  $\mathbf{d}$ , and the point-of-gaze,  $\mathbf{g}$ , the coordinates of the images of the corneal reflections and 12 uniformly spaced pupil boundary points are calculated for each camera with respect to the ICS (Section 2.3.2) as discussed in detail in Section 2.6. The calculations using a spherical corneal model are described in Section 2.6.2.1 whereas the calculations using an ellipsoidal corneal model are described in Section 2.6.2.2. The eye features are validated as described in Section 2.6.3 with  $\tau_{\text{VALID}} = 6$  mm (the validation stage determines whether the calculated points of reflection associated with the corneal reflections and the calculated points of refraction associated with the pupil boundary are within the extent of the cornea –whenever the validation stage fails for the simulations of this chapter, it will be expressly indicated). The eye and system parameters used for the generation of the synthetic eye feature data are provided in Section 4.4.

For simulations with noise in the image coordinates of the eye features, each coordinate of the corneal reflection images is corrupted with additive zero-mean Gaussian noise with a standard deviation of 0.1 pixel whereas each coordinate of the pupil boundary point images is corrupted with additive zero-mean Gaussian noise with a standard deviation of 0.3 pixel. This noise in the pupil boundary point images results in a standard deviation of approximately 0.1 pixel in the center of the pupil image. These noise characteristics are similar to those observed with the prototype system described in Chapter 5 (Section 5.2) using a stationary artificial eye. For each

position of the center of the eyeball and each point-of-gaze, 100 independent noise realizations are used.

An ellipse is fitted (in least squares sense) to the 12 pupil boundary point images from each camera. The center of the fitted ellipse (center of the pupil image) is used as an estimate of the image of the pupil center for the methods for the reconstruction of the optic axis of the eye that are based on the pupil center (“*PC...*” –see Section 4.2.2). When the *PB-EF* method (Section 4.2.2, 5)) for the reconstruction of the optic axis of the eye is used, 6 sample points from the fitted ellipse are taken every 60°.

The torsion of the eye about its optic axis,  $\kappa_{\text{eye}}$ , is assumed to be nil for all the calculations in Chapter 4 and Chapter 5.

## 4.4 Model parameters for the simulations

This section summarizes the eye and system parameters used for the simulations.

### 4.4.1 Eye parameters

The eye parameters used for the generation of the synthetic eye feature data are summarized in Table 4.1.

Table 4.1: Eye parameters for the simulations

Parameter	Description	Value	References / notes
<i>Common parameters</i>			
$D$	Distance between the center of the eyeball / center of rotation of the eye ( <b>d</b> ) and the nodal point of the eye / center of curvature of the spherical cornea ( <b>c</b> )	5.3 mm	[27]
$K$	Distance between the center of the pupil ( <b>p</b> ) and the nodal point of the eye / center of curvature of the spherical cornea ( <b>c</b> )	4.2 mm	[64]
$r_{\text{pupil}}$	Pupil radius	4 mm	Note (a)
$n_1$	Effective relative index of refraction of the cornea and aqueous humor combined (a.k.a. Standard Keratometric Index, SKI)	1.3375	[110]
$\alpha_{\text{eye}}$	Horizontal angular deviation of the visual axis from the optic axis of the eye (w.r.t. the ECS –see Section 2.2.2)	-5° (right eye)	[64], [98]
$\beta_{\text{eye}}$	Vertical angular deviation of the visual axis from the optic axis of the eye (w.r.t. the ECS –see Section 2.2.2)	1.5°	[64], [98]
<i>Spherical corneal model (Section 2.2.3.1)</i>			
$R$	Radius of curvature of the cornea	7.8 mm	[27], [64]
<i>Ellipsoidal corneal model (Section 2.2.3.2)</i>			
$R_0$	Radius of curvature at the apex of the cornea	7.8 mm	[27], [64]
$R_{\tau^*} = R_6$	Radius of curvature of the cornea at distance $\tau^* = 6$ mm from the optic axis of the eye (nominally, at the limbus)	12 mm	Note (b)
$D_{\text{apex}}$	Distance between the apex of the cornea and the center of the eyeball	13.1 mm	[27]
$\xi$	Ratio between the major and minor axes of the elliptical cross-section of the cornea perpendicular to the optic axis	1	Note (c)
$\delta$	Orientation of the major axis of the above elliptical cross-section of the cornea (irrelevant if $\xi = 1$ )	0°	

Notes:

- (a) The pupil diameter typically varies between 2 and 8 mm. Since the deviation of the center of the pupil image from the image of the pupil center increases with the pupil diameter, using a 4 mm pupil radius represents the worst-case scenario for the methods for the reconstruction of the optic axis of the eye (Section 4.2.2) that are based on the pupil center (“PC...”).
- (b) Based on previous experience [78].
- (c)  $\xi > 1$  is used to model corneal astigmatism;  $\xi = 1$  means no astigmatism (the cross-sections of the cornea perpendicular to the optic axis are circular).

## 4.4.2 System parameters

The system parameters used for the simulations are based on the prototype system described in Section 5.2, which estimates the point-of-gaze on a computer screen.

### 4.4.2.1 Cameras

#### 4.4.2.1.1 Extrinsic parameters

The cameras are positioned below the screen to minimize the interference from the upper eyelid. The coordinates of the nodal points of the cameras are given by

$$\mathbf{o}_1 = \begin{bmatrix} -o_x \\ -210 \\ 20 \end{bmatrix} \text{ mm}, \quad \mathbf{o}_2 = \begin{bmatrix} o_x \\ -210 \\ 20 \end{bmatrix} \text{ mm}, \quad o_x > 0. \quad (4.1)$$

The optic axes of the cameras intersect at

$$\mathbf{d}^* = \begin{bmatrix} 0 \\ 125 \\ 650 \end{bmatrix} \text{ mm}. \quad (4.2)$$

Consequently, the unit vector in the direction of the optic axis of camera  $j$  (the  $z_{\text{cam}}$ -axis of the CCS –see Section 2.3.1) with respect to the WCS,  $\mathbf{k}_{\text{cam},j}$ , can be found as

$$\mathbf{k}_{\text{cam},j} = \frac{\mathbf{d}^* - \mathbf{o}_j}{\|\mathbf{d}^* - \mathbf{o}_j\|}. \quad (4.3)$$

The  $x_{\text{cam}}$ -axis of the CCS is set to be horizontal while the  $y_{\text{cam}}$ -axis of the CCS points upwards. Therefore, the unit vector in the direction of the  $x_{\text{cam}}$ -axis of the CCS with respect to the WCS,  $\mathbf{i}_{\text{cam},j}$ , is given by

$$\mathbf{i}_{\text{cam},j} = \frac{[0 \ 1 \ 0]^T \times \mathbf{k}_{\text{cam},j}}{\|[0 \ 1 \ 0]^T \times \mathbf{k}_{\text{cam},j}\|}, \quad (4.4)$$

and the unit vector in the direction of the  $y_{\text{cam}}$ -axis of the CCS with respect to the WCS,  $\mathbf{j}_{\text{cam},j}$ , is given by

$$\mathbf{j}_{\text{cam},j} = \mathbf{k}_{\text{cam},j} \times \mathbf{i}_{\text{cam},j} . \quad (4.5)$$

The rotation matrix that describes the orientation of the CCS of camera  $j$  with respect to the WCS (Section 2.5.2),  $\mathbf{R}_{\text{cam},j}$ , is then given by

$$\mathbf{R}_{\text{cam},j} = [\mathbf{i}_{\text{cam},j} \quad \mathbf{j}_{\text{cam},j} \quad \mathbf{k}_{\text{cam},j}] . \quad (4.6)$$

#### 4.4.2.1.2 Intrinsic parameters

The prototype system described in Section 5.2 uses two Scorpion SCOR-14SOM cameras from Point Grey Research (Richmond, BC, Canada). These cameras are based on the monochrome CCD image sensor ICX267AL from Sony, which has a maximum resolution of 1392 pixels x 1040 pixels (square pixels). The cameras are assumed to be configured in a custom imaging mode at a resolution of 1392 pixels x 900 pixels (at the maximum horizontal resolution while being able to reach a frame rate of 20 Hz) using a region-of-interest centered in the image sensor. The principal point is assumed to be at the center of the image sensor. The camera lenses in the simulations are distortion-free. The lenses are focused at  $\mathbf{d}^*$ , and, therefore, the effective focal length given by (2.19) is

$$\lambda_j = \frac{1}{\frac{1}{f} - \frac{1}{\|\mathbf{d}^* - \mathbf{o}_j\|}} , \quad (4.7)$$

where  $f$  is the focal length of the lens.

The intrinsic camera parameters are summarized in Table 4.2.

Table 4.2: Intrinsic camera parameters for the simulations

<i>Parameter</i>	<i>Description</i>	<i>Value</i>
$\text{pixel pitch}_x$ , $\text{pixel pitch}_y$	Distances in units of length between the centers of adjacent pixels in the $x$ and $y$ directions (across the columns and across the rows), respectively	4.65 $\mu\text{m}/\text{pixel}$
$c_{\text{center}}$	Column coordinate in pixels of the principal point (the intersection of the optic axis of the camera with the image sensor)	695.5 pixels *
$r_{\text{center}}$	Row coordinate in pixels of the principal point (the intersection of the optic axis of the camera with the image sensor)	449.5 pixels *
$s_{\text{cam}}$	Skew coefficient (or “skewness” coefficient)	0
$f$	Focal length of the lens	35 mm
$\lambda$	Effective focal length of the lens	Eq. (4.7)
$k_{\text{r.d.1}}, k_{\text{r.d.2}}, k_{\text{r.d.3}},$ $k_{\text{t.d.1}}, k_{\text{t.d.2}}$	Lens distortion coefficients	0

\* The top left pixel of the image has coordinates (0, 0) with respect to the ICS (Section 2.3.2)

#### 4.4.2.2 Light sources

All the system setups tested through the simulations have two light sources symmetrically positioned at the sides of the computer monitor:

$$\mathbf{I}_1 = \begin{bmatrix} -250 \\ l_{1-2,Y} \\ 0 \end{bmatrix} \text{mm}, \quad \mathbf{I}_2 = \begin{bmatrix} 250 \\ l_{1-2,Y} \\ 0 \end{bmatrix} \text{mm}, \quad (4.8)$$

where  $l_{1-2,Y}$  is an adjustable parameter.

System setups with four light sources have an additional light source centered below the screen and an additional light source centered above the screen:

$$\mathbf{I}_3 = \begin{bmatrix} 0 \\ -215 \\ 0 \end{bmatrix} \text{mm}, \quad \mathbf{I}_4 = \begin{bmatrix} 0 \\ 215 \\ 0 \end{bmatrix} \text{mm}. \quad (4.9)$$

#### 4.4.2.3 Scene

The scene in the prototype system described in Section 5.2 is the screen of a 19" LCD monitor with a resolution of 1280 pixels x 1024 pixels (5:4 aspect ratio). The viewable area of the screen is 376.32 mm x 301.06 mm. The dimensions of the screen are rounded for the simulations to 380 mm x 300 mm:

$$\begin{aligned} -190 \text{ mm} &\leq X_{\text{scene}} \leq 190 \text{ mm} \\ -150 \text{ mm} &\leq Y_{\text{scene}} \leq 150 \text{ mm} \\ Z_{\text{scene}} &= 0 \end{aligned} \quad (4.10)$$

## 4.5 Evaluation of the inherent point-of-gaze estimation bias

In the absence of noise in the eye features and errors in the system parameters, if the front corneal surface is spherical, all 5 methods for the calculation of the 3-D coordinates of the center of curvature of the cornea that are summarized in Section 4.2.1 are exact. On the other hand, out of the 5 methods for the subsequent reconstruction of the optic axis of the eye in 3-D space that are summarized in Section 4.2.2, only the *PB* method can be considered exact. The methods that are based on the pupil center (*PC*...) and use the center of the ellipse that is fitted to the pupil boundary image points as an estimate of the image of the pupil center, and the *PB-EF* method, which uses samples of the fitted ellipse, are expected to exhibit point-of-gaze estimation bias. The purpose of this section is to quantify this bias through numerical simulations.

These simulations are carried out for a right eye with a spherical cornea, a single fixed position of the center of the eyeball ( $\mathbf{d} = \mathbf{d}^* = [0 \ 125 \ 650]^T \text{ mm}$ ), 25 points-of-gaze arranged in a 5-by-5 rectangular grid spanning the entire screen (including the center and the four corners) and two different separations between the nodal points of the cameras (100 mm, i.e.,  $o_X = 50 \text{ mm}$ , and 300 mm, i.e.,  $o_X = 150 \text{ mm}$ ). These two different camera separations correspond, approximately, to the minimum and maximum separations that are possible in the prototype system described in Section 5.2. The point-of-gaze is estimated using the actual values of the system parameters and the actual values of  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ . For the *PCTriangRefr*, *PB* and *PB-EF* methods, the actual value of  $R$  is also used. Note that, for these particular simulations, the point-of-gaze is estimated using the coordinates of  $\mathbf{c}$  obtained during the generation of the synthetic eye feature data instead

of calculating  $\mathbf{c}$  from the corneal reflection images. The goal is to quantify the inherent bias of the methods for the subsequent reconstruction of the optic axis of the eye in 3-D space under “perfect” conditions.

The root-mean-square (RMS) and maximum (MAX) bias for each case are given in Table 4.3 and Table 4.4. The inherent bias of the *PCPIInt* method increases with the camera separation but is completely negligible. The inherent bias of the *PCTriang* method is the largest of all and it decreases slightly with the camera separation. Recall that this method not only uses the center of the pupil image as an estimate of the image of the pupil center (in general, these two points are not coincident –see Section 3.3.4) but also assumes that the virtual pupil center is in the optic axis of the eye (Section 3.3.3.2). The comparison of the bias of this method with the bias of the other methods that also use the center of the pupil image as an estimate of the image of the pupil center (i.e., *PCPIInt* and *PCTriangRefr*) suggests that the assumption that the virtual pupil center is in the optic axis of the eye is the primary source of the point-of-gaze estimation bias of the *PCTriang* method. The inherent bias of the *PCTriangRefr* method, which also decreases with the camera separation, is the second largest but is significantly smaller than that of the *PCTriang* method. The inherent bias of the *PB* method is completely negligible and it can be attributed to numerical errors since, theoretically, it is exact for the present simulation conditions. Nevertheless, it is worth noting that the numerical error, as well as the computation time of the numerical optimization used in this method, decrease with the camera separation. This is an expected result since increasing the camera separation implies increasing the angle of convergence of the cameras, which should improve the numerical stability of any 3-D reconstruction technique based on stereo vision. The inherent bias of the *PB-EF* method is relatively small and it increases with the camera separation. Since the *PB-EF* method differs from the *PB* method in the ellipse fitting-sampling step (see Section 4.2.2, 5)), the bias exhibited by the *PB-EF* method suggests that the boundary of the pupil image is not exactly elliptical.

Table 4.3 and Table 4.4 also provide the inherent point-of-gaze estimation bias at the screen center. Typically,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  are calibrated by having the subject fixate a target shown at the center of the screen. The relatively small values of the bias at the screen center (compared to the MAX Bias and the RMS Bias) suggest that calibrating for  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  would only have a marginal effect on the calculated values of the MAX Bias and the RMS Bias.

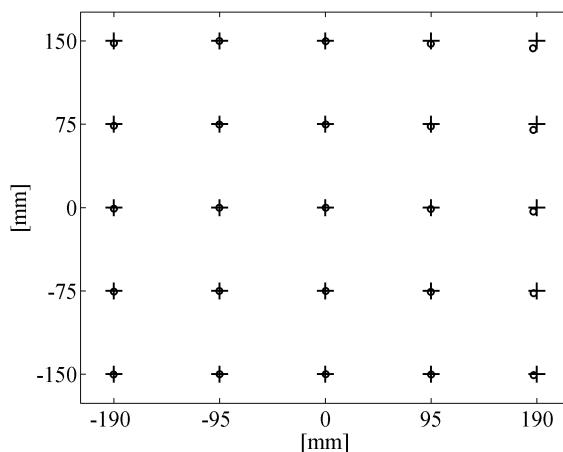
Table 4.3: Inherent point-of-gaze estimation bias when  $o_X = 50$  mm and the pupil diameter is 8 mm

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>RMS Bias</i> [mm]	$4.06 \cdot 10^{-6}$	2.91	1.05	$3.23 \cdot 10^{-5}$	$1.24 \cdot 10^{-1}$
<i>MAX Bias</i> [mm]	$1.15 \cdot 10^{-5}$	8.21	2.94	$7.74 \cdot 10^{-5}$	$3.37 \cdot 10^{-1}$
<i>Bias at screen center</i> [mm]	$1.51 \cdot 10^{-8}$	$1.95 \cdot 10^{-1}$	$7.01 \cdot 10^{-2}$	$3.65 \cdot 10^{-5}$	$4.10 \cdot 10^{-3}$

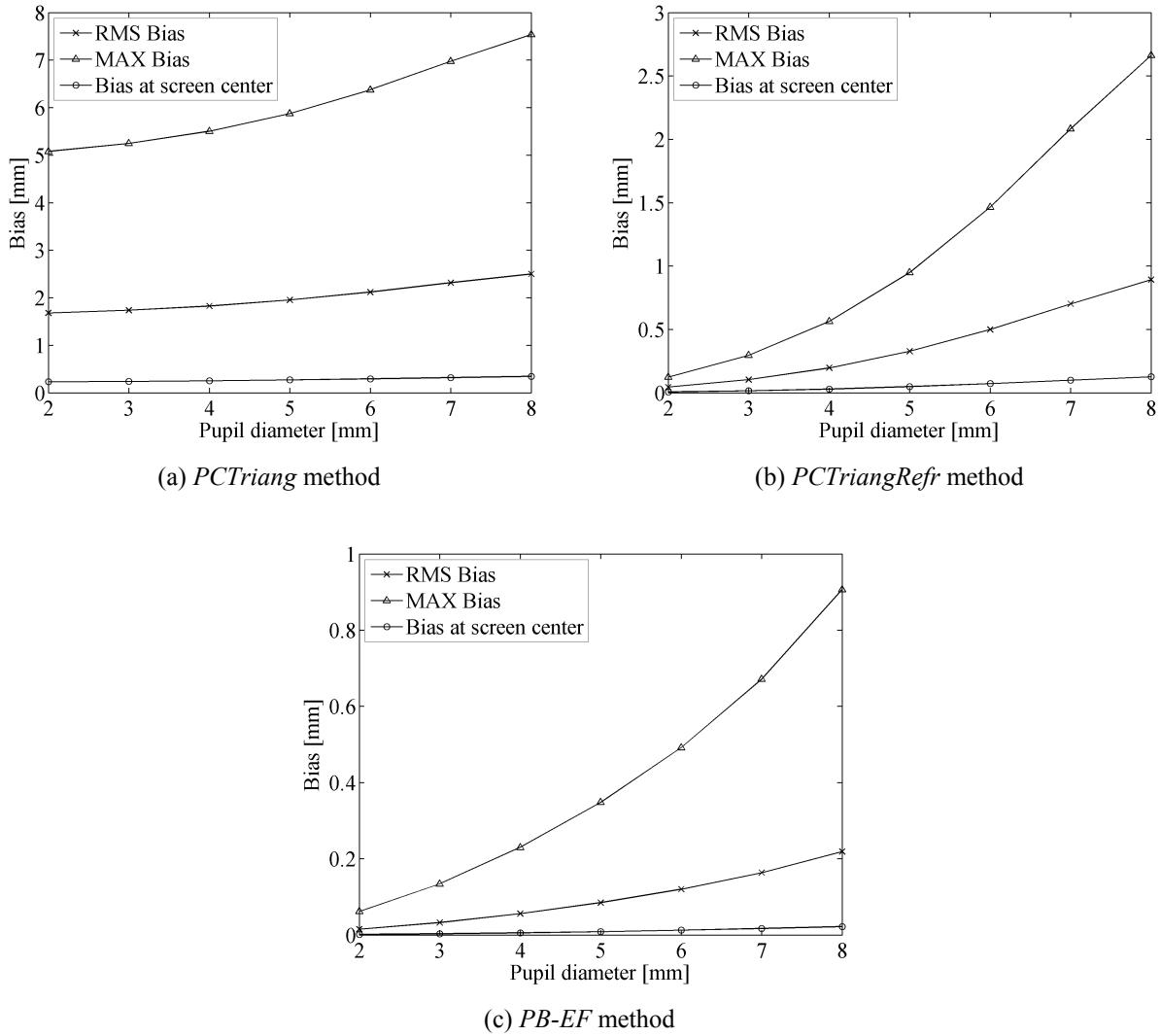
Table 4.4: Inherent point-of-gaze estimation bias when  $o_X = 150$  mm and the pupil diameter is 8 mm

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>RMS Bias</i> [mm]	$2.31 \cdot 10^{-5}$	2.50	$8.93 \cdot 10^{-1}$	$6.86 \cdot 10^{-6}$	$2.19 \cdot 10^{-1}$
<i>MAX Bias</i> [mm]	$1.04 \cdot 10^{-4}$	7.54	2.66	$2.07 \cdot 10^{-5}$	$9.06 \cdot 10^{-1}$
<i>Bias at screen center</i> [mm]	$4.94 \cdot 10^{-8}$	$3.46 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$	$5.10 \cdot 10^{-6}$	$2.18 \cdot 10^{-2}$

Fig. 4.1 shows the inherent point-of-gaze estimation bias of the *PCTriang* method when  $o_X = 150$  mm and the pupil diameter is 8 mm. As it can be seen in this figure, the largest bias occurs when the simulated right eye fixates the top right corner of the screen (largest angular deviation between the optic axis of the eye and the optic axes of the cameras for this simulation). Fig. 4.2 illustrates how the inherent bias of the *PCTriang*, *PCTriangRefr* and *PB-EF* methods increases with the pupil diameter, as mentioned earlier.

Fig. 4.1: Inherent point-of-gaze estimation bias of the *PCTriang* method when  $o_X = 150$  mm and the pupil diameter is 8 mm.

(+ : actual point-of-gaze, ○ : estimated point-of-gaze)

Fig. 4.2: Inherent point-of-gaze estimation bias as a function of the pupil diameter when  $o_X = 150$  mm.

## 4.6 Sensitivities of the point-of-gaze estimation with two cameras and two light sources

This section presents a simulation study of the sensitivity of the different point-of-gaze estimation methods to several sources of errors (noise in the image coordinates of the eye features, corneal asphericity, error in a selected system parameter) for a system with two cameras and two light sources. As it is shown in Chapter 3, two cameras and two light sources is the simplest system configuration that can be used to reconstruct the optic axis of the eye in 3-D

space without knowing any subject-specific eye parameter, and to estimate the point-of-gaze after completing a single-point personal calibration procedure.

All 25 combinations of the 5 methods for the calculation of the 3-D coordinates of the center of curvature of the cornea (Section 4.2.1) and the 5 methods for the subsequent reconstruction of the optic axis of the eye in 3-D space (Section 4.2.2) are considered. The simulations in this section are carried out for a single fixed position of the center of the eyeball ( $\mathbf{d} = \mathbf{d}^* = [0 \ 125 \ 650]^T$  mm), 25 points-of-gaze arranged in a 5-by-5 rectangular grid spanning the entire screen (including the center and the four corners), two different separations between the nodal points of the cameras and two different values of the vertical coordinate of the light sources for each camera separation. The two different camera separations correspond, approximately, to the minimum and maximum separations that are possible in the prototype system described in Section 5.2. Increasing the camera separation while keeping the point of intersection of the optic axes of the cameras fixed implies increasing the angle of convergence of the cameras, which should improve the robustness of any 3-D reconstruction technique based on stereo vision. The vertical coordinate of the light sources ( $l_{1,2,Y}$  in (4.8)) should have a significant effect on the methods for the calculation of the 3-D coordinates of the center of curvature of the cornea (Section 4.2.1) that are based on the intersection of planes defined by the nodal point of camera  $j$ ,  $\mathbf{o}_j$ , light source  $i$ ,  $\mathbf{l}_i$ , and its corresponding corneal reflection image,  $\mathbf{u}_{ij}$  ( $CPlInt\dots$ ) but a marginal effect on the methods that do not depend on plane intersections ( $CNoPlInt\dots$ ). The 4 system setups considered are summarized in Table 4.5, where  $\theta_{\text{opt.ax.cam}}$ 's is the angle of convergence of the optic axes of the cameras, and  $\theta_{\text{l.s.}}$  is the angle subtended by the two light sources,  $\mathbf{l}_1$  and  $\mathbf{l}_2$ , at the nodal point of each camera,  $\mathbf{o}_j$ . The angle  $\theta_{\text{l.s.}}$  is an approximation to the angle between the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_1$  and  $\mathbf{u}_{1j}$ , and the plane defined by  $\mathbf{o}_j$ ,  $\mathbf{l}_2$  and  $\mathbf{u}_{2j}$ . Of course,  $\theta_{\text{l.s.}} = 90^\circ$ , should result in the lowest sensitivity of the  $CPlInt\dots$  methods to noise in the estimated coordinates of the centers of the corneal reflection images.

Table 4.5: System setups

Setup #	$o_X$ [mm]	$l_{1-2,Y}$ [mm]	$\theta_{\text{opt.ax.cam's.}}$ [ $^\circ$ ]	$\theta_{\text{l.s.}}$ [ $^\circ$ ]
I	50	-173	8.02	160
II	50	34	8.02	90
III	150	-190	23.74	160
IV	150	-11	23.74	90

#### 4.6.1 Sensitivity to noise in the image coordinates of the eye features

In these simulations, the synthetic eye features are generated with the spherical corneal model and noise is added to the eye features as described in Section 4.3. The point-of-gaze is estimated using the actual values of the system parameters and the actual values of  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ . To quantify the dispersion of the point-of-gaze estimates due to noise by itself, the distance between the noisy point-of-gaze estimates and the corresponding noise-free point-of-gaze estimates is determined. Since the dispersion of the point-of-gaze estimates depends on the coordinates of the fixation point, the RMS distance between the noisy point-of-gaze estimates and the corresponding noise-free point-of-gaze estimates is determined for each of the 25 fixation points tested (RMS Dispersion ea. fix. pt.). The maximum value of the 25 RMS dispersions (MAX(RMS Dispersion ea. fix. pt.)) for each point-of-gaze estimation method is provided in Table 4.6 (setup I), Table 4.8 (setup II), Table 4.10 (setup III), and Table 4.12 (setup IV). The RMS distance between the noisy point-of-gaze estimates and the corresponding noise-free point-of-gaze estimates for all fixation points (RMS Dispersion) for each point-of-gaze estimation method is provided in Table 4.7 (setup I), Table 4.9 (setup II), Table 4.11 (setup III), and Table 4.13 (setup IV). Since for each system setup all the methods estimate the point-of-gaze for the same noisy eye feature data, and since the pseudo-random number generator is reset for each system setup, the simulation results can be compared both across methods and across setups.

From the results of these simulations, the following observations can be made:

- As expected, increasing the angle between the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's.}}$ ) reduces the sensitivity to noise of all methods. For example, compare Fig. 4.3(a) with Fig. 4.3(c), Fig. 4.3(b) with Fig. 4.3(d), Fig. 4.4(a) with Fig. 4.4(c), Fig. 4.4(b) with Fig. 4.4(d), and Fig. 4.5

with Fig. 4.6. Note also how the sensitivity to noise depends on the coordinates of the fixation point.

- Calculation of the 3-D coordinates of the center of curvature of the cornea:
  - The sensitivity to noise of the *CPlInt...* methods is strongly dependent on the angle  $\theta_{l.s.}$  subtended by the two light sources at the nodal point of each camera (Fig. 4.3) whereas the sensitivity to noise of the *CNoPlInt...* methods is largely independent of  $\theta_{l.s.}$  (Fig. 4.4), as predicted earlier.
  - In general terms, the *CNoPlInt...* methods perform better than the *CPlInt...* methods and the “-4CR” methods perform better than the “-3CR” methods: of the 5 methods tested, the *CNoPlInt-4CR* method performs best and the *CPlIntInv-3CR* method performs worst. The *CNoPlInt...* methods perform only slightly better than the *CPlInt...* methods for system setups that are favorable to the *CPlInt...* methods (i.e.,  $\theta_{l.s.} = 90^\circ$  –for example, compare Fig. 4.3(b) with Fig. 4.4(b), and Fig. 4.3(d) with Fig. 4.4(d)), and dramatically better than the *CPlInt...* methods for system setups that are unfavorable to the *CPlInt...* methods (e.g.,  $\theta_{l.s.} = 160^\circ$  –for example, compare Fig. 4.3(a) with Fig. 4.4(a), and Fig. 4.3(c) with Fig. 4.4(c)). Note that the two *CPlInt...-4CR* methods perform similarly to each other.
- Subsequent reconstruction of the optic axis of the eye in 3-D space:
  - In general terms, the *PCPlInt* method exhibits the highest sensitivity to noise whereas the *PB* and *PB-EF* methods exhibit the lowest (for example, see Fig. 4.5 and Fig. 4.6). In most cases, the *PCPlInt* method is significantly worse than all the other methods. Furthermore, the sensitivity to noise of the *PCPlInt* method increases towards the bottom of the screen (for example, see Fig. 4.5(a) and Fig. 4.6(a)). This is an expected result since from the discussion in Section 3.3.1 it follows that the sensitivity of the *PCPlInt* method should increase as the optic axis of the eye gets closer to the line defined by the nodal points of the cameras (recall that the cameras are under the screen).
  - The *PB* and *PB-EF* methods perform slightly better than the *PCTriang* and *PCTriangRefr* methods for favorable system setups (i.e., larger  $\theta_{opt.ax.cam}$ 's –for example, see Fig. 4.6(b)-(e)), and significantly better than the *PCTriang* and *PCTriangRefr* methods for unfavorable

system setups (i.e., smaller  $\theta_{\text{opt.ax.cam's}}$  –for example, see Fig. 4.5(b)-(e)). The *PCTriang* method tends to perform marginally better than the *PCTriangRefr* method. The *PB* and *PB-EF* methods perform fairly similarly to each other.

- In summary, the *CNoPlInt-4CR + PB/PB-EF* method with a large angle between the optic axes of the cameras provides the lowest sensitivity to noise, virtually independently of the position of the light sources.

Table 4.6: Noise sensitivity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – *MAX(RMS Dispersion ea. fix. pt.)* [mm]

<i>Method</i>	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	70.89	48.69	49.44	22.38	23.07
<i>CPlIntPsInv-4CR</i>	58.39	28.73	30.01	16.35	16.38
<i>CPlIntTriang-4CR</i>	58.27	28.76	30.04	16.35	16.38
<i>CNoPlInt-3CR</i>	50.96	11.62	12.12	7.12	7.03
<i>CNoPlInt-4CR</i>	50.69	10.24	10.69	6.51	6.46

Table 4.7: Noise sensitivity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – *RMS Dispersion* [mm]

<i>Method</i>	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	33.25	29.41	29.92	18.14	18.11
<i>CPlIntPsInv-4CR</i>	23.39	19.43	19.81	12.93	12.95
<i>CPlIntTriang-4CR</i>	23.39	19.44	19.82	12.95	12.96
<i>CNoPlInt-3CR</i>	13.88	7.39	7.52	5.54	5.48
<i>CNoPlInt-4CR</i>	13.40	6.66	6.78	5.18	5.12

Table 4.8: Noise sensitivity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	51.41	13.26	13.84	8.26	8.23
<i>CPlIntPsInv-4CR</i>	50.58	11.38	11.88	6.89	6.84
<i>CPlIntTriang-4CR</i>	50.58	11.38	11.88	6.89	6.83
<i>CNoPlInt-3CR</i>	51.18	11.38	11.89	7.11	7.03
<i>CNoPlInt-4CR</i>	50.72	10.16	10.60	6.51	6.46

Table 4.9: Noise sensitivity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Dispersion [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	14.72	8.69	8.85	6.19	6.13
<i>CPlIntPsInv-4CR</i>	13.85	7.51	7.64	5.60	5.56
<i>CPlIntTriang-4CR</i>	13.85	7.51	7.64	5.60	5.56
<i>CNoPlInt-3CR</i>	13.83	7.23	7.36	5.47	5.41
<i>CNoPlInt-4CR</i>	13.40	6.64	6.76	5.18	5.13

Table 4.10: Noise sensitivity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	28.84	28.73	28.92	23.20	23.61
<i>CPlIntPsInv-4CR</i>	13.22	10.52	10.96	10.00	10.07
<i>CPlIntTriang-4CR</i>	13.29	10.75	11.20	10.11	10.18
<i>CNoPlInt-3CR</i>	10.40	4.78	4.95	4.08	4.14
<i>CNoPlInt-4CR</i>	10.29	4.07	4.22	3.65	3.71

Table 4.11: Noise sensitivity – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Dispersion [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	18.69	18.28	18.39	16.42	16.49
<i>CPlIntPsInv-4CR</i>	8.39	7.76	7.88	7.36	7.39
<i>CPlIntTriang-4CR</i>	8.61	7.99	8.11	7.58	7.61
<i>CNoPlInt-3CR</i>	4.34	3.31	3.35	3.12	3.12
<i>CNoPlInt-4CR</i>	4.06	2.96	2.99	2.83	2.83

Table 4.12: Noise sensitivity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX(RMS Dispersion ea. fix. pt.) [mm]

Method	<i>PCPI</i> <i>nt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPI</i> <i>nt</i> <i>Inv-3CR</i>	10.60	5.41	5.60	4.81	4.93
<i>CPI</i> <i>nt</i> <i>PsInv-4CR</i>	10.43	4.28	4.45	3.99	4.05
<i>CPI</i> <i>nt</i> <i>Triang-4CR</i>	10.43	4.28	4.45	3.99	4.05
<i>CNoP</i> <i>nt</i> <i>-3CR</i>	10.44	4.48	4.65	3.90	3.96
<i>CNoP</i> <i>nt</i> <i>-4CR</i>	10.31	3.95	4.10	3.59	3.64

Table 4.13: Noise sensitivity – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Dispersion [mm]

Method	<i>PCPI</i> <i>nt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPI</i> <i>nt</i> <i>Inv-3CR</i>	4.80	3.84	3.88	3.58	3.59
<i>CPI</i> <i>nt</i> <i>PsInv-4CR</i>	4.30	3.25	3.29	3.09	3.10
<i>CPI</i> <i>nt</i> <i>Triang-4CR</i>	4.30	3.25	3.29	3.09	3.10
<i>CNoP</i> <i>nt</i> <i>-3CR</i>	4.22	3.14	3.18	2.98	2.99
<i>CNoP</i> <i>nt</i> <i>-4CR</i>	4.03	2.91	2.95	2.80	2.80

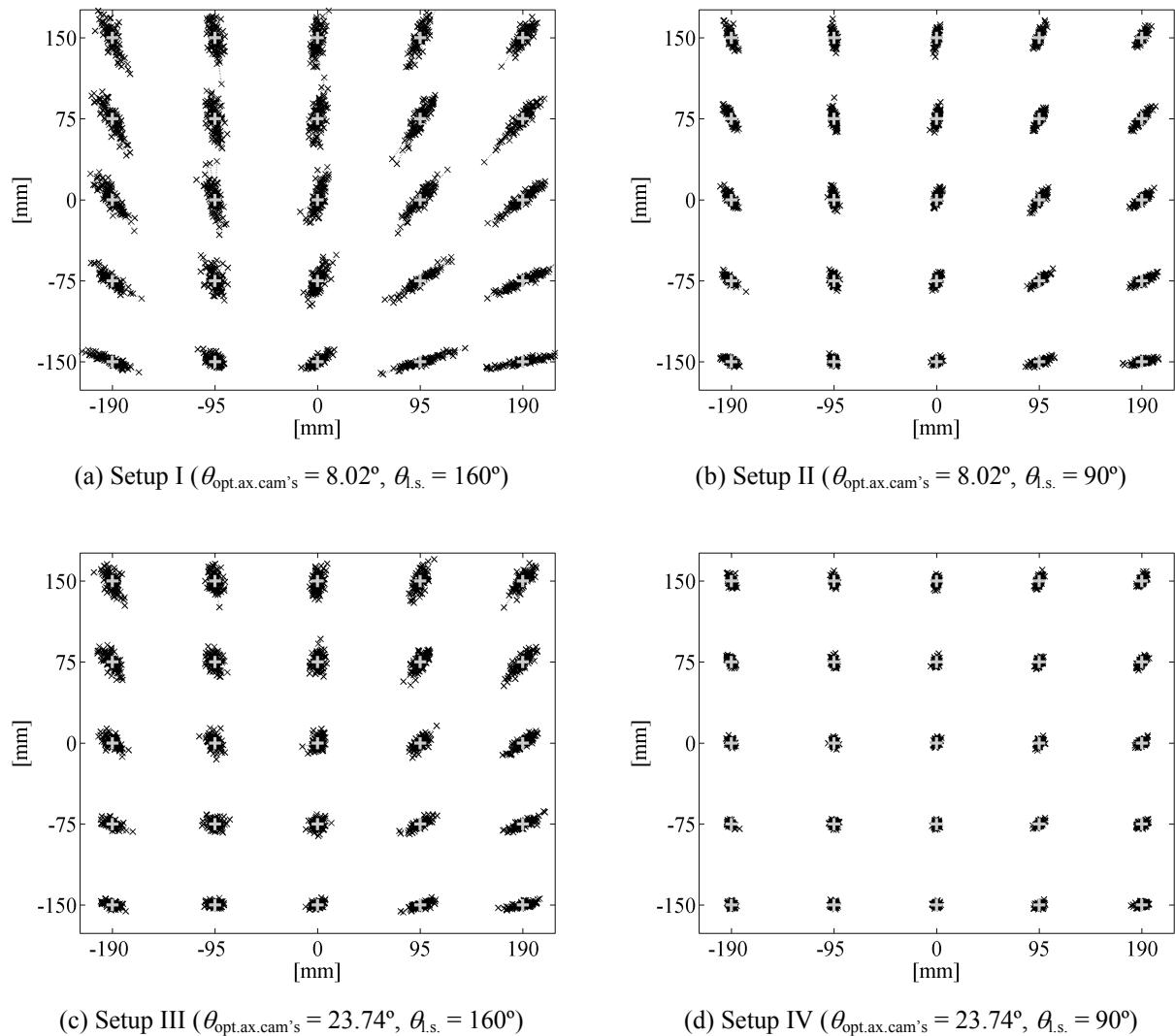


Fig. 4.3: Noise sensitivity of the *CPLIntTriang-4CR + PB* method  
(+ : actual point-of-gaze,  $\times$  : estimated point-of-gaze).

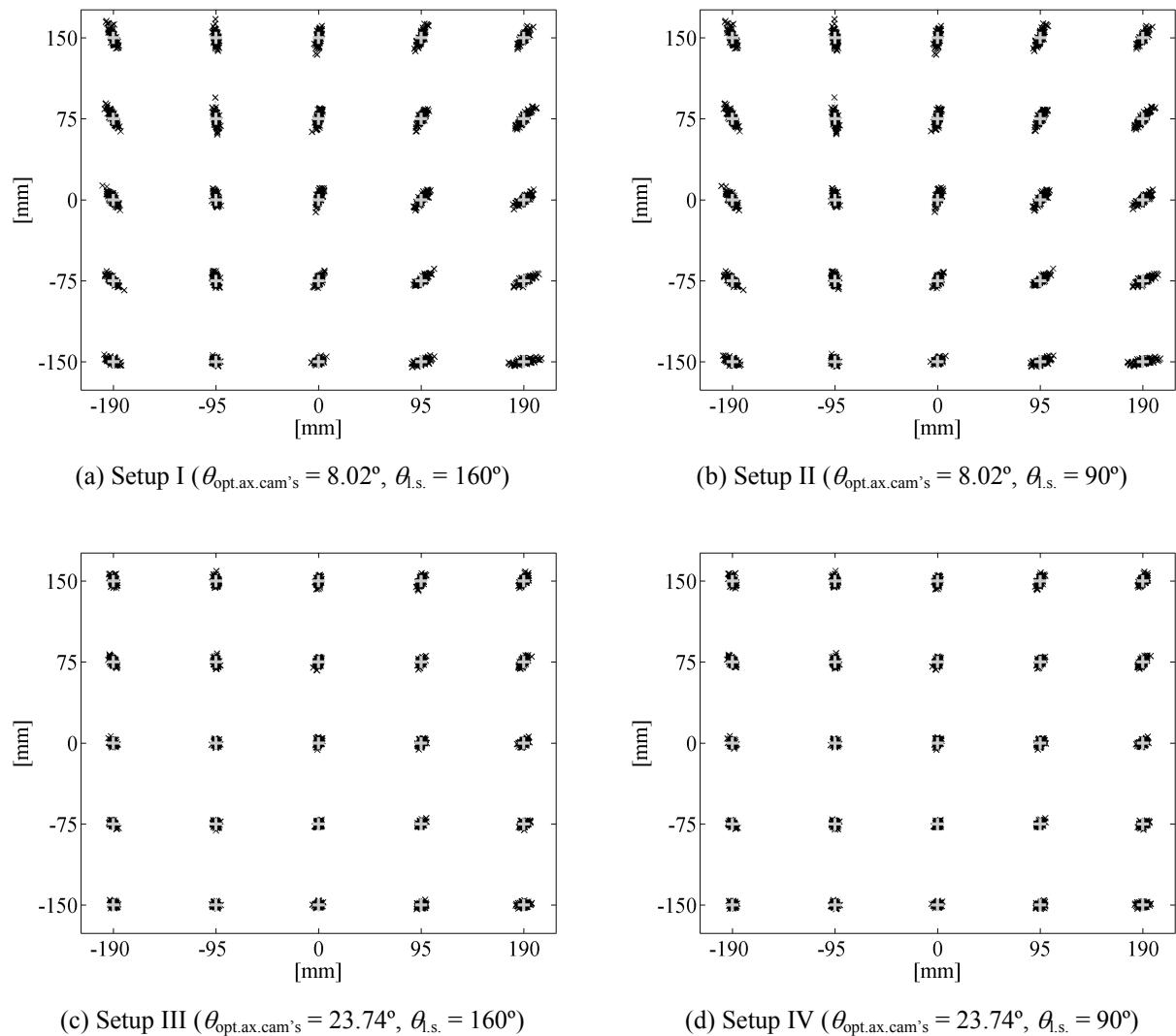


Fig. 4.4: Noise sensitivity of the *CNoPPIInt-4CR + PB* method.  
 (+ : actual point-of-gaze,  $\times$  : estimated point-of-gaze)

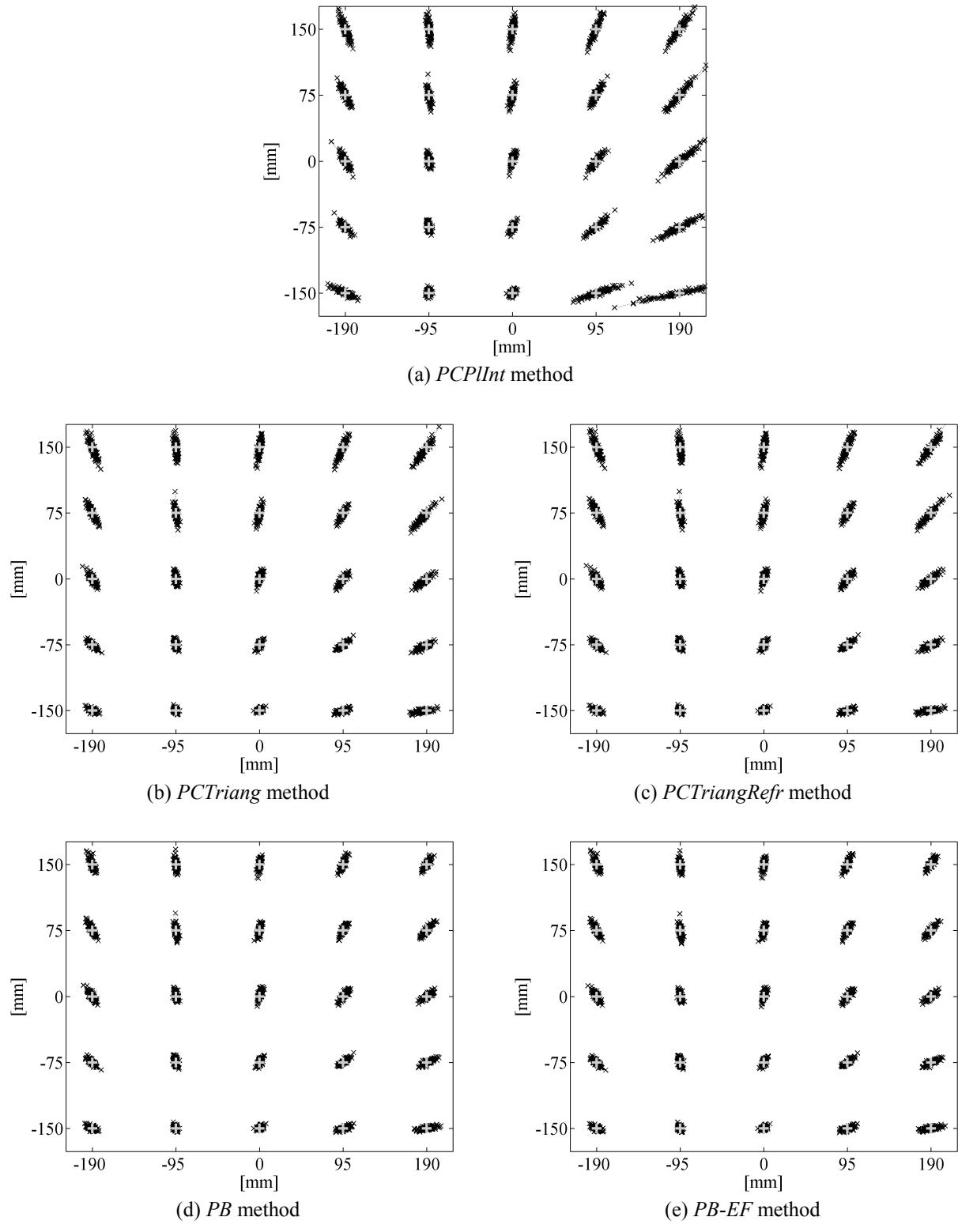


Fig. 4.5: Noise sensitivity of the different methods for the subsequent reconstruction of the optic axis of the eye with system setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) when the 3-D coordinates of the center of curvature of the cornea are calculated with the *CNoPInt-4CR* method.  
 (+ : actual point-of-gaze,  $\times$  : estimated point-of-gaze)

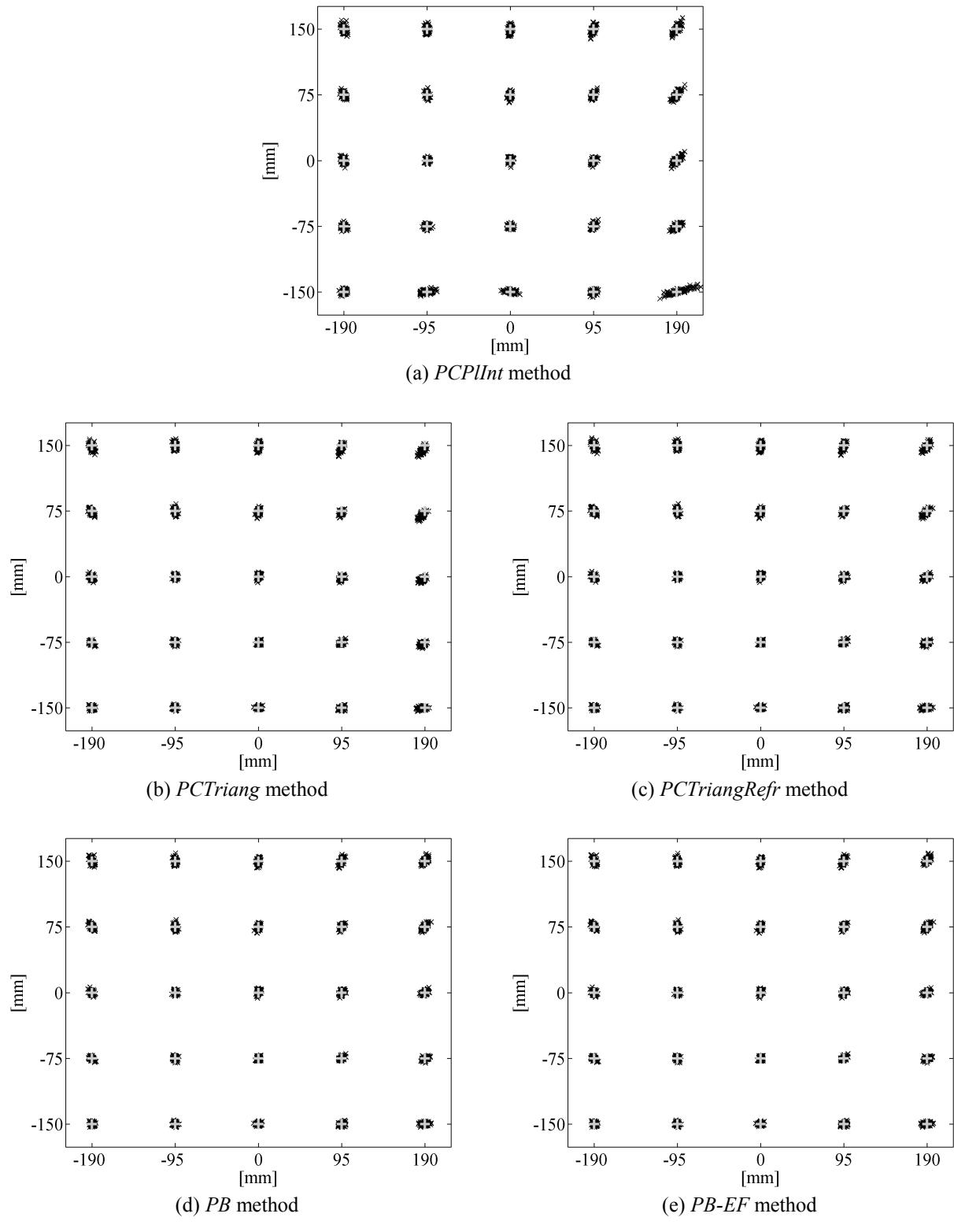


Fig. 4.6: Noise sensitivity of the different methods for the subsequent reconstruction of the optic axis of the eye with system setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) when the 3-D coordinates of the center of curvature of the cornea are calculated with the *CNoPInt-4CR* method.  
 (+ : actual point-of-gaze,  $\times$  : estimated point-of-gaze)

### 4.6.2 Effect of corneal asphericity

In these simulations, the synthetic eye features are generated with the ellipsoidal corneal model and no noise is added to the eye features. The point-of-gaze is estimated using the actual values of the system parameters after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center (i.e., when the simulated eye gazes at  $\mathbf{g} = [0 \ 0 \ 0]^T$ ). To quantify the effect of corneal asphericity, the distance between each point-of-gaze estimate and its corresponding actual point-of-gaze (bias) is determined (keep in mind that this bias is a combination of the bias due to corneal asphericity and the inherent bias, if any). The MAX bias for each point-of-gaze estimation method is provided in Table 4.14 (setup I), Table 4.16 (setup II), Table 4.18 (setup III), and Table 4.20 (setup IV), whereas the RMS bias is provided in Table 4.15 (setup I), Table 4.17 (setup II), Table 4.19 (setup III), and Table 4.21 (setup IV). It must be noted that the point-of-gaze is not estimated for 1 fixation point for setups I and IV and 2 fixation points for setup III (for example, see Fig. 4.7 through Fig. 4.10) because, in the calculation of the coordinates of the images of the corneal reflections for those fixation points, one of the points of reflection ( $\mathbf{q}_{ij}$ ) is more than  $\tau_{\text{VALID}} = 6$  mm from the optic axis of the eye (i.e., for one of the cameras the reflection of one of the light sources falls outside the cornea, and therefore one of the cameras sees only one corneal reflection). Consequently, the MAX and RMS bias provided in Table 4.14 through Table 4.21 cannot be compared across the different system setups, even though they can still be compared across the different point-of-gaze estimation methods. The problem of not being able to estimate the point-of-gaze for a certain region of the screen can be solved by using additional light sources as discussed in Section 4.7.

From the results of these simulations, the following observations can be made:

- Increasing the angle between the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ) has a detrimental effect in the sense that the region of the screen for which the point-of-gaze cannot be estimated becomes larger. For example, compare setup I with setup III and setup II with setup IV in Fig. 4.7 through Fig. 4.10.
- The vertical coordinates of the light sources also have a role on the size of the region of the screen for which the point-of-gaze cannot be estimated: it is smaller when  $\theta_{\text{l.s.}} = 90^\circ$  than when  $\theta_{\text{l.s.}} = 160^\circ$ . For example, compare setup I with setup II and setup III with setup IV in

Fig. 4.7 through Fig. 4.10.

- Beyond its effect on the size of the region of the screen for which the point-of-gaze cannot be estimated, the vertical coordinates of the light sources have a much stronger effect on the magnitude of the point-of-gaze estimation bias of the *CPlInt...* methods than on the bias of the *CNoPliInt...* methods. For example, compare the differences between setups I/III ( $\theta_{l.s.} = 160^\circ$ ) and setups II/IV ( $\theta_{l.s.} = 90^\circ$ ) in Fig. 4.7 (*CPlInt...*) vs. Fig. 4.9 (*CNoPliInt...*), and in Fig. 4.8 (*CPlInt...*) vs. Fig. 4.10 (*CNoPliInt...*).
- For the calculation of the 3-D coordinates of the center of curvature of the cornea, in general terms, the *CNoPliInt...* methods perform better than the *CPlInt...* methods (for example, compare Fig. 4.7 with Fig. 4.9, and Fig. 4.8 with Fig. 4.10) and the “-3CR” methods perform better than the “-4CR” methods (for example, compare Fig. 4.7 with Fig. 4.8, and Fig. 4.9 with Fig. 4.10). The *CNoPliInt...* methods perform slightly better than the *CPlInt...* methods when  $\theta_{l.s.} = 90^\circ$  (setups II/IV), and dramatically better than the *CPlInt...* methods when  $\theta_{l.s.} = 160^\circ$  (setups I/III). Of the 5 methods tested, the *CNoPliInt-3CR* method performs best and the *CPlInt...-4CR* methods perform worst (the two *CPlInt...-4CR* methods perform similarly to each other).
- The performance of the methods for the subsequent reconstruction of the optic axis of the eye in 3-D space can be rated from best to worst as follows: 1) *PB* and *PB-EF* methods, 2) *PCPliInt* method, 3) *PCTriangRefr* method, and 4) *PCTriang* method (keep in mind the contribution of the inherent bias).
- In summary, the *CNoPliInt-3CR + PB/PB-EF* method provides the smallest point-of-gaze estimation bias.

The simulation results from this section together with the simulation results from the previous section suggest that the attention should be focused on the *CNoPliInt-3CR/4CR + PB/PB-EF* methods. However, before narrowing down the remainder of the simulation study to these methods, it is worth considering the sensitivity of the different methods to errors in a selected system parameter as described in the next section.

Table 4.14: Effect of corneal asphericity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm]

Method	<i>PCPInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	50.45	50.78	50.48	49.06	49.02
<i>CPIIntPsInv-4CR</i>	55.19	55.91	55.49	52.15	52.11
<i>CPIIntTriang-4CR</i>	54.82	55.55	55.12	51.68	51.63
<i>CNoPPInt-3CR</i>	14.16	19.88	16.14	7.27	7.22
<i>CNoPPInt-4CR</i>	35.34	40.24	37.35	20.75	20.46

**Note:** The point-of-gaze was not estimated for one fixation point (top right corner of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

Table 4.15: Effect of corneal asphericity – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm]

Method	<i>PCPInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	24.08	24.04	24.03	21.96	21.83
<i>CPIIntPsInv-4CR</i>	26.80	27.56	27.12	24.81	24.81
<i>CPIIntTriang-4CR</i>	26.71	27.49	27.04	24.67	24.67
<i>CNoPPInt-3CR</i>	4.38	6.23	5.00	3.01	3.00
<i>CNoPPInt-4CR</i>	12.86	14.68	13.59	9.04	8.97

**Note:** The point-of-gaze was not estimated for one fixation point (top right corner of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

Table 4.16: Effect of corneal asphericity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	13.04	14.55	13.31	12.36	12.39
<i>CPlIntPsInv-4CR</i>	37.19	43.15	39.49	26.97	27.14
<i>CPlIntTriang-4CR</i>	37.22	43.18	39.52	27.00	27.17
<i>CNoPlInt-3CR</i>	11.96	15.99	12.11	11.65	11.65
<i>CNoPlInt-4CR</i>	35.72	41.50	37.94	24.23	24.39

Table 4.17: Effect of corneal asphericity – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm]

Method	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	6.07	6.79	6.16	5.41	5.43
<i>CPlIntPsInv-4CR</i>	14.99	16.83	15.65	10.71	10.76
<i>CPlIntTriang-4CR</i>	15.00	16.85	15.67	10.71	10.76
<i>CNoPlInt-3CR</i>	4.68	5.88	4.95	4.07	4.08
<i>CNoPlInt-4CR</i>	13.65	15.57	14.35	9.38	9.33

Table 4.18: Effect of corneal asphericity – Setup III ( $\theta_{\text{opt,ax,cam}^*} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm]

Method	<i>PCPI</i> <i>nt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPI</i> <i>ntInv-3CR</i>	47.37	48.49	47.59	45.34	45.44
<i>CPI</i> <i>ntPsInv-4CR</i>	62.54	63.75	63.02	61.83 *	62.35 *
<i>CPI</i> <i>ntTriang-4CR</i>	60.54	61.80	61.09	60.08 *	60.68 *
<i>CNoP</i> <i>nt-3CR</i>	6.60	10.94	8.18	6.27	6.25
<i>CNoP</i> <i>nt-4CR</i>	28.71	32.14	30.12	27.68	27.49

**Note:** The point-of-gaze was not estimated for two fixation points (2 topmost on the right edge of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

\* : The point-of-gaze was not estimated for one additional fixation point because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

Table 4.19: Effect of corneal asphericity – Setup III ( $\theta_{\text{opt,ax,cam}^*} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm]

Method	<i>PCPI</i> <i>nt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPI</i> <i>ntInv-3CR</i>	23.29	23.45	23.30	22.37	22.43
<i>CPI</i> <i>ntPsInv-4CR</i>	29.26	29.87	29.53	29.31 *	29.37 *
<i>CPI</i> <i>ntTriang-4CR</i>	28.50	29.16	28.81	28.27 *	28.42 *
<i>CNoP</i> <i>nt-3CR</i>	1.95	3.27	2.39	1.88	1.87
<i>CNoP</i> <i>nt-4CR</i>	10.79	12.05	11.28	10.34	10.31

**Note:** The point-of-gaze was not estimated for two fixation points (2 topmost on the right edge of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

\* : The point-of-gaze was not estimated for one additional fixation point because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

Table 4.20: Effect of corneal asphericity – Setup IV ( $\theta_{\text{opt,ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – MAX Bias [mm]

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	7.87	8.23	7.89	6.95	7.00
<i>CPIIntPsInv-4CR</i>	26.14	30.80	27.90	25.20	25.08
<i>CPIIntTriang-4CR</i>	26.37	31.01	28.13	25.40	25.26
<i>CNoPPIInt-3CR</i>	5.50	9.99	6.35	5.36	5.34
<i>CNoPPIInt-4CR</i>	30.32	34.49	31.91	28.25	27.76

**Note:** The point-of-gaze was not estimated for one fixation point (top right corner of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

Table 4.21: Effect of corneal asphericity – Setup IV ( $\theta_{\text{opt,ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – RMS Bias [mm]

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	4.10	4.56	4.16	3.87	3.85
<i>CPIIntPsInv-4CR</i>	12.93	14.25	13.38	12.45	12.45
<i>CPIIntTriang-4CR</i>	12.94	14.27	13.40	12.46	12.46
<i>CNoPPIInt-3CR</i>	2.24	3.27	2.47	2.12	2.09
<i>CNoPPIInt-4CR</i>	11.81	13.14	12.28	11.22	11.14

**Note:** The point-of-gaze was not estimated for one fixation point (top right corner of the screen) because one corneal reflection was missing (i.e., it did not pass the validation stage in the generation of the synthetic eye features).

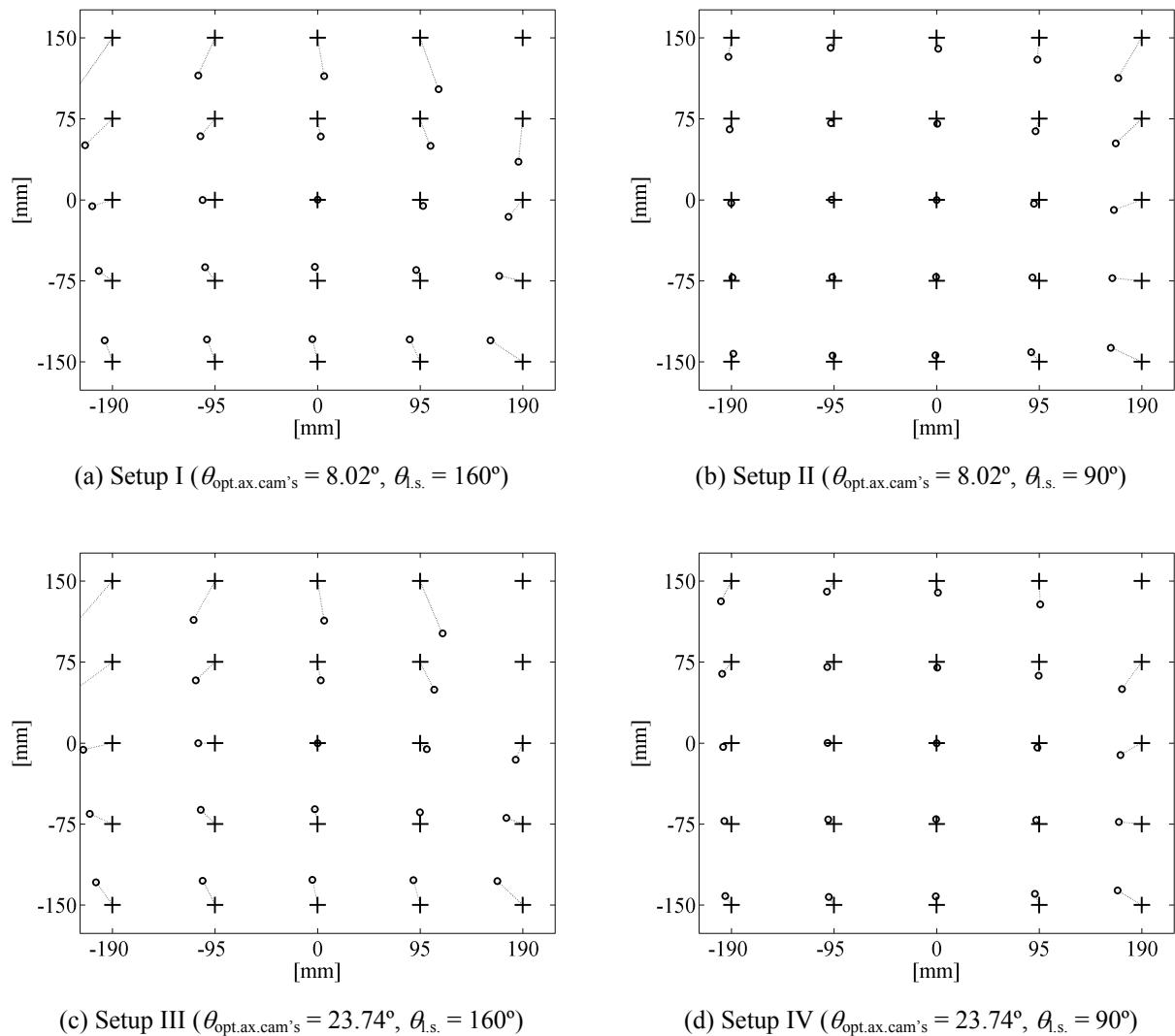


Fig. 4.7: Effect of corneal asphericity – Bias of the *CPLIntTriang-4CR + PCTriang* method.  
 (+ : actual point-of-gaze, ○ : estimated point-of-gaze)

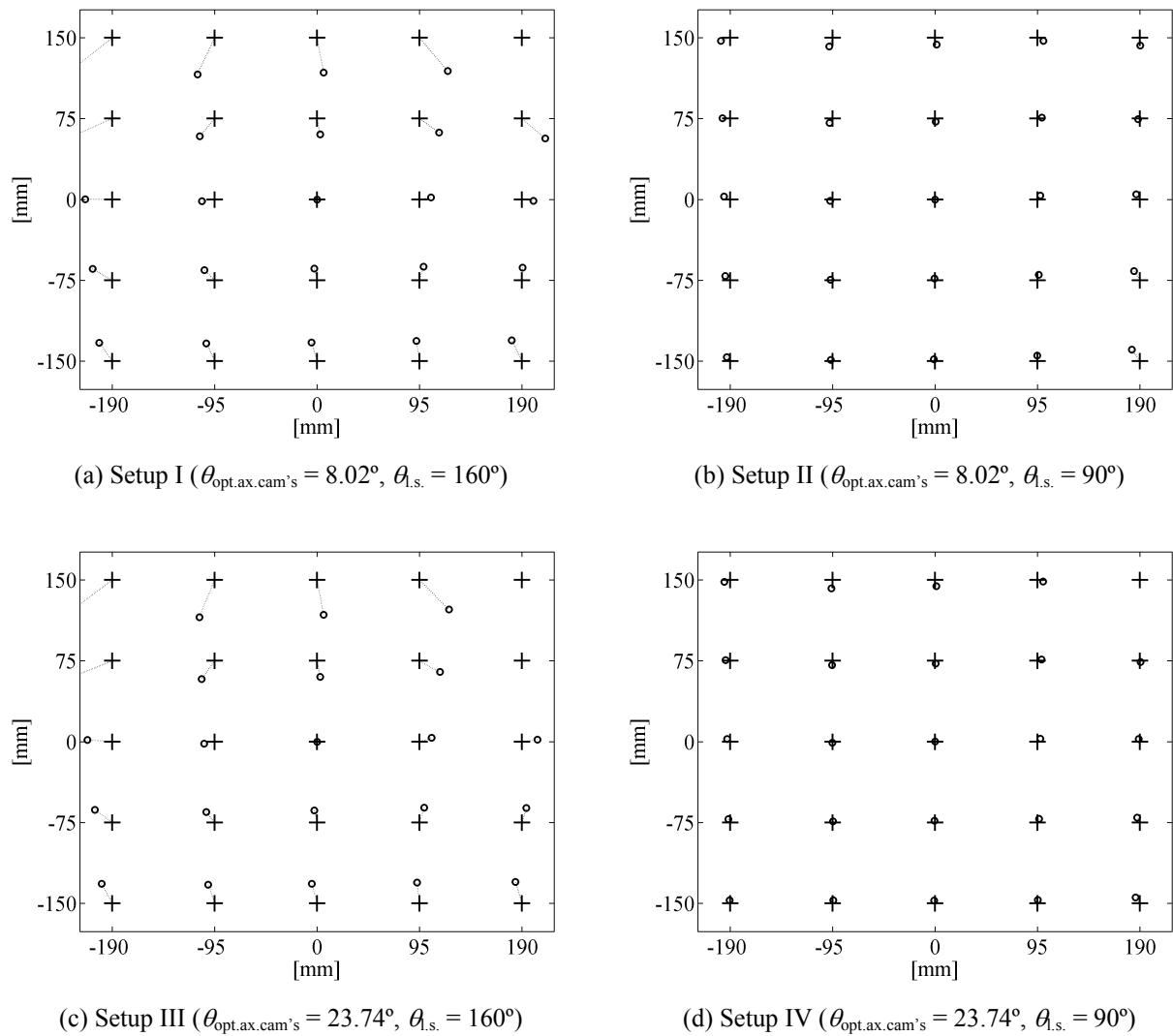
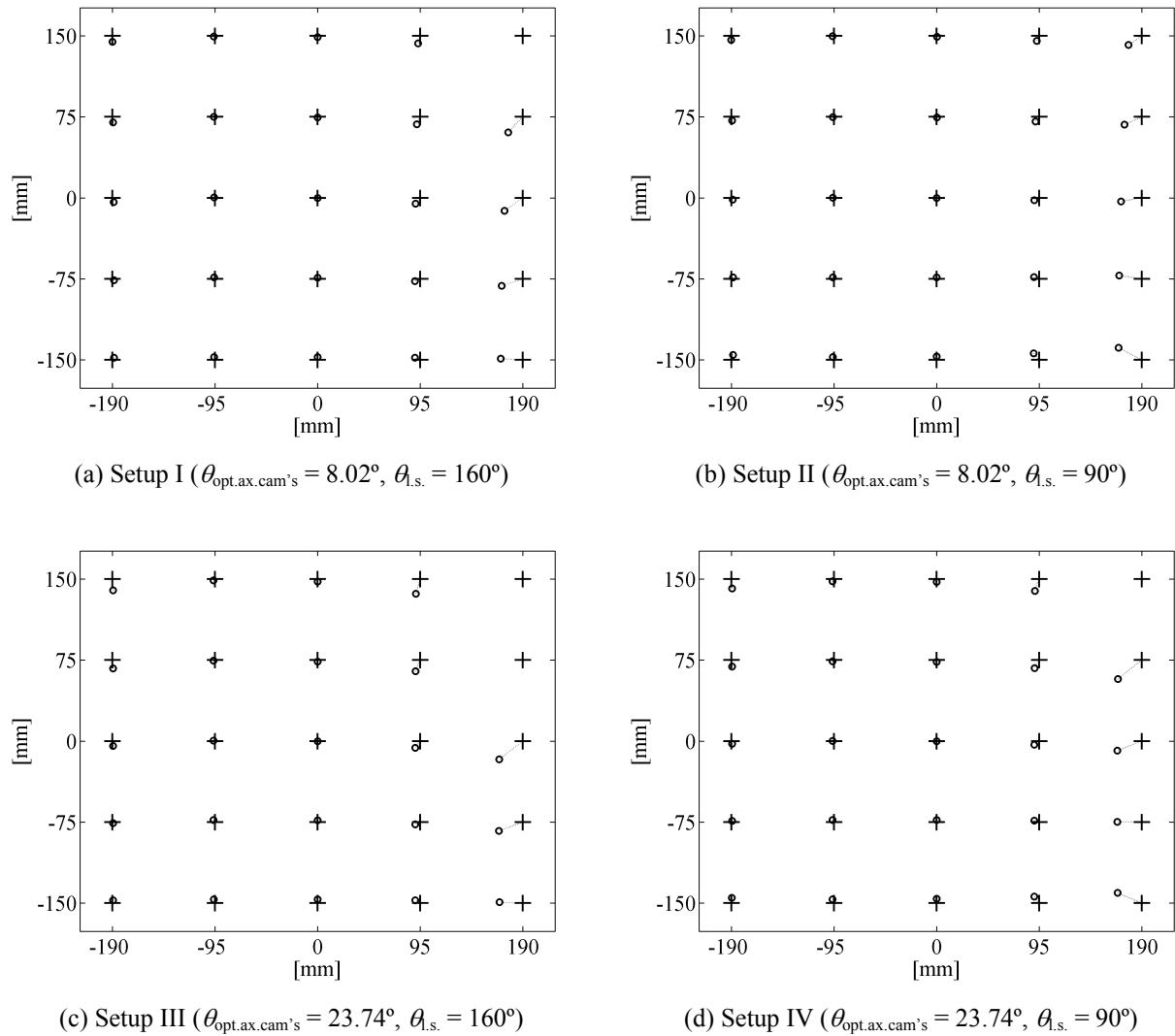


Fig. 4.8: Effect of corneal asphericity – Bias of the *CPLIntInv-3CR + PCPInt* method.  
 (+ : actual point-of-gaze, ○ : estimated point-of-gaze)

Fig. 4.9: Effect of corneal asphericity – Bias of the *CNoPInt-4CR + PB* method.(+ : actual point-of-gaze,  $\circ$  : estimated point-of-gaze)

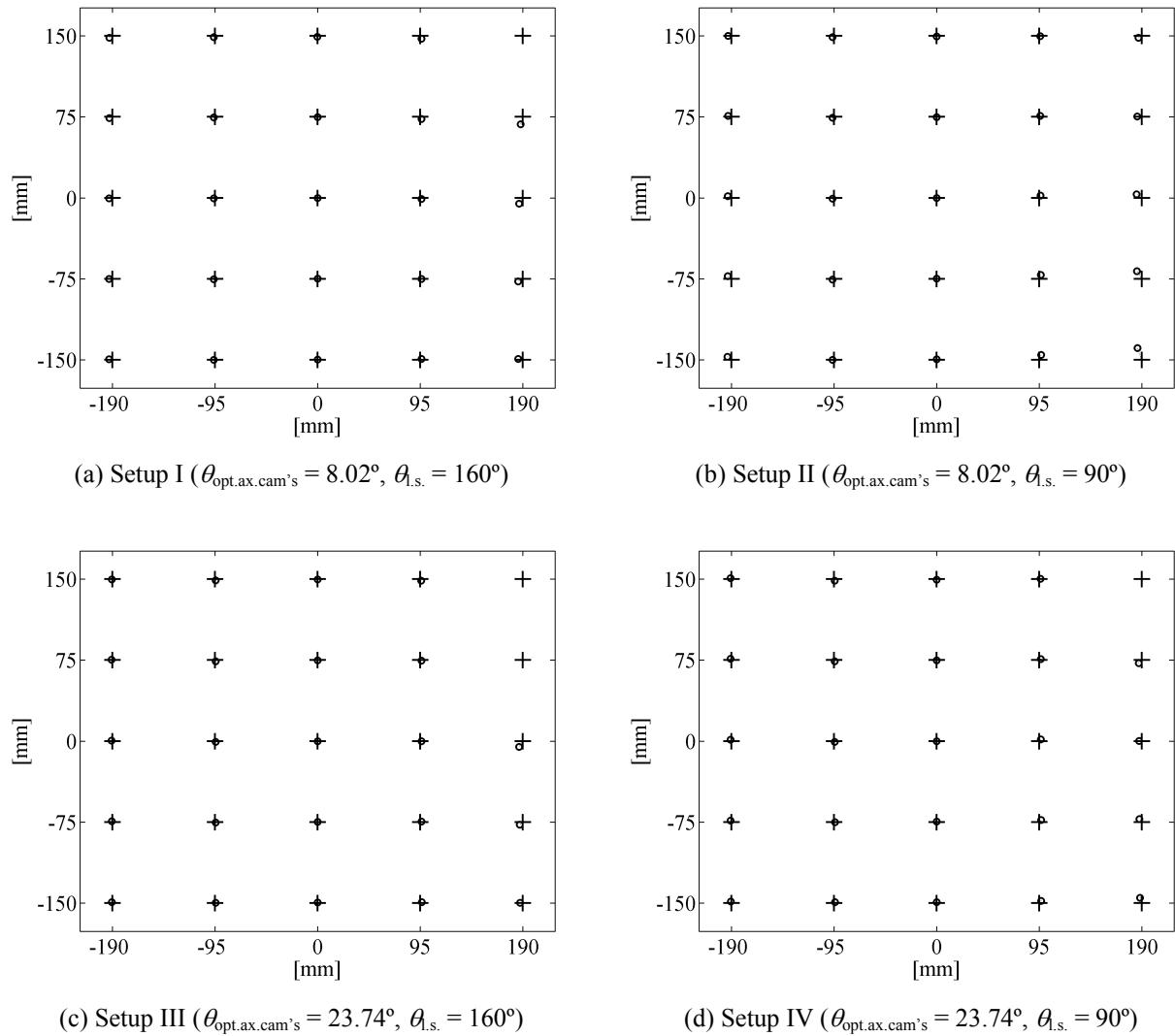


Fig. 4.10: Effect of corneal asphericity – Bias of the *CNoPInt-3CR + PB* method.  
(+ : actual point-of-gaze,  $\circ$  : estimated point-of-gaze)

### 4.6.3 Sensitivity to error in a selected system parameter

In this section, the sensitivity of the point-of-gaze estimates to an error in the tilt angle of one of the cameras is studied through simulations. The tilt angle of a camera is the angle between the optic axis of the camera and a horizontal plane (it is defined analytically in Appendix C). The tilt angle of one of the cameras is selected here because most of the point-of-gaze estimation methods are highly sensitive to small errors in this parameter. An extensive study of the sensitivity to errors in several system parameters is presented later.

In these simulations, the synthetic eye features are generated with the spherical corneal model and no noise is added to the eye features. The point-of-gaze is estimated using the actual values of the system parameters except for the tilt angle of camera 1, which is  $0.1^\circ$  larger than the actual value, after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center. To quantify the effect of this system parameter error, the distance between each point-of-gaze estimate and its corresponding actual point-of-gaze (bias) is determined (keep in mind that this bias is a combination of the bias due to the system parameter error and the inherent bias, if any). The MAX bias for each point-of-gaze estimation method is provided in Table 4.22 (setup I), Table 4.24 (setup II), Table 4.26 (setup III), and Table 4.28 (setup IV), whereas the RMS bias is provided in Table 4.23 (setup I), Table 4.25 (setup II), Table 4.27 (setup III), and Table 4.29 (setup IV).

The results of these simulations show that the only methods that perform reasonably in the presence of a  $0.1^\circ$  error in the tilt angle of camera 1 for all four system setups are the *CPIIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods. Furthermore, the bias exhibited by these two methods is very close to the inherent bias of the *PCTriang* method (see Table 4.3 and Table 4.4), and the calibrated values of  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  are very close to their actual values (within  $0.03^\circ$ ). The *CPIIntPsInv-4CR + PCTriang* method with setups II and IV ( $\theta_{\text{l.s.}} = 90^\circ$ ) performs similarly to the *CPIIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods. The performance of the *CPIIntTriang-4CR + PCTriangRefr* and the *CNoPInt-4CR + PCTriangRefr* methods with setups III and IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) and the *CPIIntPsInv-4CR + PCTriangRefr* method with setup IV could be possibly acceptable in some applications although their MAX Bias is about 2.6-2.9 times larger and their RMS Bias is about 5-5.5 times larger than those of the *CPIIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods.

Another observation from these simulations is that increasing the camera separation, and consequently the angle of convergence of the optic axis of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ), has a beneficial effect in reducing the effect of error in the parameter studied.

In any case, the results of these simulations suggest that the *CPlIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods should be considered further. The following section provides a summary of the simulation results obtained so far.

Table 4.22: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – *MAX Bias* [mm]

<i>Method</i>	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	316.92	793.75	791.22	576.10	576.10
<i>CPlIntPsInv-4CR</i>	22214.76	708.52	132099508.31	428.04	431.90
<i>CPlIntTriang-4CR</i>	10716.14	8.44	475.07	135.60	136.47
<i>CNoPInt-3CR</i>	7157.40	70.16	833.27	198.12	198.10
<i>CNoPInt-4CR</i>	10969.49	7.99	490.09	137.08	137.96

Table 4.23: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup I ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – *RMS Bias* [mm]

<i>Method</i>	<i>PCPlInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntInv-3CR</i>	149.63	437.86	436.41	305.63	305.63
<i>CPlIntPsInv-4CR</i>	4495.29	299.63	26419911.41	211.93	212.67
<i>CPlIntTriang-4CR</i>	2304.27	2.96	248.89	64.72	64.97
<i>CNoPInt-3CR</i>	1645.92	37.14	324.06	83.00	82.62
<i>CNoPInt-4CR</i>	2352.72	2.76	255.15	65.36	65.61

Table 4.24: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – *MAX Bias* [mm]

<i>Method</i>	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	250701.44	806.18	717.82	152.24 *	151.18 *
<i>CPIIntPsInv-4CR</i>	11344.69	7.32	491.22	136.77	137.65
<i>CPIIntTriang-4CR</i>	10888.12	8.23	482.99	136.43	137.31
<i>CNoPIInt-3CR</i>	7126.17	69.29	816.15	199.46	199.43
<i>CNoPIInt-4CR</i>	11173.00	8.01	489.94	137.23	138.11

\* : The point-of-gaze was not estimated for 10 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

Table 4.25: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup II ( $\theta_{\text{opt.ax.cam's}} = 8.02^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – *RMS Bias* [mm]

<i>Method</i>	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	50140.87	487.84	434.61	93.71 *	93.66 *
<i>CPIIntPsInv-4CR</i>	2406.85	2.54	256.21	65.62	65.86
<i>CPIIntTriang-4CR</i>	2336.47	2.86	252.23	65.09	65.34
<i>CNoPIInt-3CR</i>	1634.33	36.25	317.12	83.11	82.76
<i>CNoPIInt-4CR</i>	2390.92	2.77	255.08	65.43	65.68

\* : The point-of-gaze was not estimated for 10 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

Table 4.26: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – MAX Bias [mm]

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	315.37	769.32	769.24	561.56	561.56
<i>CPIIntPsInv-4CR</i>	700.99	33.22	42.89	76.53	78.02
<i>CPIIntTriang-4CR</i>	130837.16	7.76	20.36	58.17	58.91
<i>CNoPPIInt-3CR</i>	1358.62	68.54	91.37	119.88	118.38
<i>CNoPPIInt-4CR</i>	148333.57	7.31	21.19	60.07	60.80

Table 4.27: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup III ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 160^\circ$ ) – RMS Bias [mm]

Method	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	165.37	425.04	425.35	298.03	298.03
<i>CPIIntPsInv-4CR</i>	172.96	14.98	16.5	31.21	32.61
<i>CPIIntTriang-4CR</i>	26170.82	2.59	12.92	27.45	28.35
<i>CNoPPIInt-3CR</i>	366.46	37.21	42.81	50.28	48.76
<i>CNoPPIInt-4CR</i>	29669.79	2.44	13.46	28.40	29.27

Table 4.28: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – *MAX Bias* [mm]

<i>Method</i>	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	1075.13	2345.5	1726.12	632.29 *	1086.31 **
<i>CPIIntPsInv-4CR</i>	4402.57	7.05	20.5	59.84	60.61
<i>CPIIntTriang-4CR</i>	116940.94	7.54	20.76	59.27	60.01
<i>CNoPIInt-3CR</i>	1415.24	68.49	91.18	120.07	118.56
<i>CNoPIInt-4CR</i>	40077.75	7.32	21.16	60.06	60.79

\* : The point-of-gaze was not estimated for 8 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

\*\* : The point-of-gaze was not estimated for 7 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

Table 4.29: Effect of a  $0.1^\circ$  error in the tilt angle of camera 1 – Setup IV ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ,  $\theta_{\text{l.s.}} = 90^\circ$ ) – *RMS Bias* [mm]

<i>Method</i>	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>	397.08	869.38	700.62	211.46 *	328.75 **
<i>CPIIntPsInv-4CR</i>	955.72	2.38	13.47	28.67	29.55
<i>CPIIntTriang-4CR</i>	23392.04	2.51	13.19	28.01	28.89
<i>CNoPIInt-3CR</i>	374.41	36.72	41.97	49.54	48.09
<i>CNoPIInt-4CR</i>	8027.19	2.44	13.44	28.40	29.27

\* : The point-of-gaze was not estimated for 8 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

\*\* : The point-of-gaze was not estimated for 7 of the 25 fixation points because at least one back-projected ray from a pupil boundary image point did not intersect the corneal surface (i.e.,  $k_{r_b,jk}$  in (3.48) is not a real number).

#### 4.6.4 Summary of simulation results

The most important general observations from the simulation results presented so far can be summarized as follows:

- 1) In the absence of error in the system parameters the following is observed: the *CNoPlInt...* methods for the calculation of the 3-D coordinates of the center of curvature of the cornea perform consistently better than the *CPlInt...* methods in the presence of noise in the image coordinates of the eye features and corneal asphericity. The “-4CR” methods perform better in the presence of noise in the image coordinates of the eye features while the “-3CR” methods perform better in the presence of corneal asphericity. The *PB* and *PB-EF* methods for the subsequent reconstruction of the optic axis of the eye in 3-D space perform consistently better than the other methods in the presence of noise in the image coordinates of the eye features and corneal asphericity. For “favorable” system setups the *PCTriangRefr* method can provide a second choice. Note that the *PCTriangRefr* method is much less computationally expensive than the *PB* and *PB-EF* methods since the former provides a closed-form solution whereas the latter rely on numerical optimization.
- 2) In the presence of error in the system parameter tested in the previous section (tilt angle of one of the cameras), the *CPlIntTriang-4CR + PCTriang* and the *CNoPlInt-4CR + PCTriang* methods are the only methods that have a consistently reasonable performance for all the system setups considered.
- 3) Increasing the angle between the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ) reduces the sensitivity to noise in the image coordinates of the eye features and the effect of error in the system parameter tested in the previous section (tilt angle of one of the cameras) but can have a detrimental effect in the presence of corneal asphericity.
- 4) The vertical coordinates of the light sources play a role in system sensitivities, particularly on the ability to estimate the point-of-gaze in certain regions of the screen. Beyond this, the effect of the vertical coordinates of the light sources is much stronger on the *CPlInt...* methods for the calculation of the 3-D coordinates of the center of curvature of the cornea than on the *CNoPlInt...* methods.

Based on 1) and 2), the remainder of this chapter concentrates on the methods indicated in Table 4.30. Based on 3) and 4), the effect of the system setup on system performance is studied further, now using 4 light sources in order to improve the system robustness, particularly to corneal asphericity.

Table 4.30: Point-of-gaze estimation methods subject to further consideration

<i>Method</i>	<i>PCPIInt</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntInv-3CR</i>					
<i>CPIIntPsInv-4CR</i>					
<i>CPIIntTriang-4CR</i>		✓			
<i>CNoPIInt-3CR</i>		✓	✓	✓	✓
<i>CNoPIInt-4CR</i>		✓	✓	✓	✓

## 4.7 Sensitivities of the point-of-gaze estimation with two cameras and four light sources

The use of more than the minimum of two light sources can help to improve the system robustness by increasing the likelihood that at least two corneal reflections are available for each camera regardless of the head position and the point-of-gaze on the screen, and in the presence of eyelid and eyelash interferences. Moreover, the use of more than two light sources can help to reduce the effect of corneal asphericity. Typically, the radius of curvature of the front corneal surface increases towards the boundary with the sclera and only the central part is approximately spherical [27]. By using the two light sources that produce corneal reflections that are closest to the pupil center in the eye images, it is therefore possible to reduce the point-of-gaze estimation bias due to corneal asphericity. In fact, as it is shown through simulations in the remainder of this chapter, the reduction of the effect of corneal asphericity resulting from this approach can be significant, particularly when the geometrical arrangement of the light sources is optimized. Consequently, this approach is followed in all the simulations and experiments described in the

remainder of this thesis. Note that the light sources are selected in such a way that the same pair of light sources is used with both cameras<sup>10</sup>. Formally, the selection of the light sources for each gaze estimate is carried out as follows. Let  $DPCCR_{ij}$  be the distance between the center of the pupil and the corneal reflection corresponding to light source  $i$  in the image from camera  $j$ . Then, after finding  $M_i = \max(DPCCR_{i1}, DPCCR_{i2})$  for each of the light sources, the two light sources corresponding to the two smallest  $M_i$  are selected. After selecting the pair of light sources with its corresponding pair of corneal reflections, the point-of-gaze is estimated using the methods for two cameras and two light sources (summarized in Section 4.2, restricted to the subset indicated in Table 4.30 for the remainder of this chapter).

To demonstrate the improved performance using 4 light sources, a simulation is carried out with  $o_X = 150$  mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) and  $l_{1-2,Y} = 0$  mm (the coordinates of the 4 light sources are provided in Section 4.4.2.2 and  $l_{1-2,Y}$  is the vertical coordinate of the light sources that are located at the sides of the screen;  $\pm o_X$  is the  $X$  coordinate of the nodal points of the cameras, which are provided in Section 4.4.2.1.1). In this simulation, which is also carried out for a single fixed position of the center of the eyeball ( $\mathbf{d} = \mathbf{d}^* = [0 \ 125 \ 650]^T$  mm) and 25 points-of-gaze arranged in a 5-by-5 rectangular grid spanning the entire screen (including the center and the four corners), the synthetic eye features are generated with the ellipsoidal corneal model and no noise is added to the eye features. The point-of-gaze is estimated using the actual values of the system parameters after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center. To quantify the effect of corneal asphericity, the distance between each point-of-gaze estimate and its corresponding actual point-of-gaze (bias) is determined (keep in mind that this bias is a combination of the bias due to corneal asphericity and the inherent bias, if any). The MAX bias for each point-of-gaze estimation method tested is provided in Table 4.31, whereas the RMS bias is provided in Table 4.32. Fig. 4.11 shows the point-of-gaze estimates for 4 of the 9 methods tested. Not only can the point-of-gaze be estimated for any point on the screen but also the point-of-gaze estimation bias is smaller than for any of the setups with 2 light sources studied before (compare Fig. 4.11(a) with Fig. 4.7, Fig. 4.11(c) with Fig. 4.9, and Fig. 4.11(d) with Fig. 4.10).

---

<sup>10</sup> Alternatively and for increased flexibility, the two light sources could be selected independently for each camera, which means that occasionally one camera would be associated with one pair of light sources and the other camera would be associated with a different pair. However, this approach is less robust to errors in system parameters (especially, errors in the extrinsic camera parameters) than the one used in this thesis.

Table 4.31: Effect of corneal asphericity – System with 4 light sources with  $l_{1-2,Y} = 0$  mm and  $o_X = 150$  mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	28.92	N/A	N/A	N/A
CNoPIInt-3CR	16.77	12.31	8.74	8.06
CNoPIInt-4CR	21.73	17.47	13.44	12.85

Table 4.32: Effect of corneal asphericity – System with 4 light sources with  $l_{1-2,Y} = 0$  mm and  $o_X = 150$  mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.54	N/A	N/A	N/A
CNoPIInt-3CR	5.00	3.57	2.54	2.41
CNoPIInt-4CR	7.35	5.95	4.82	4.72

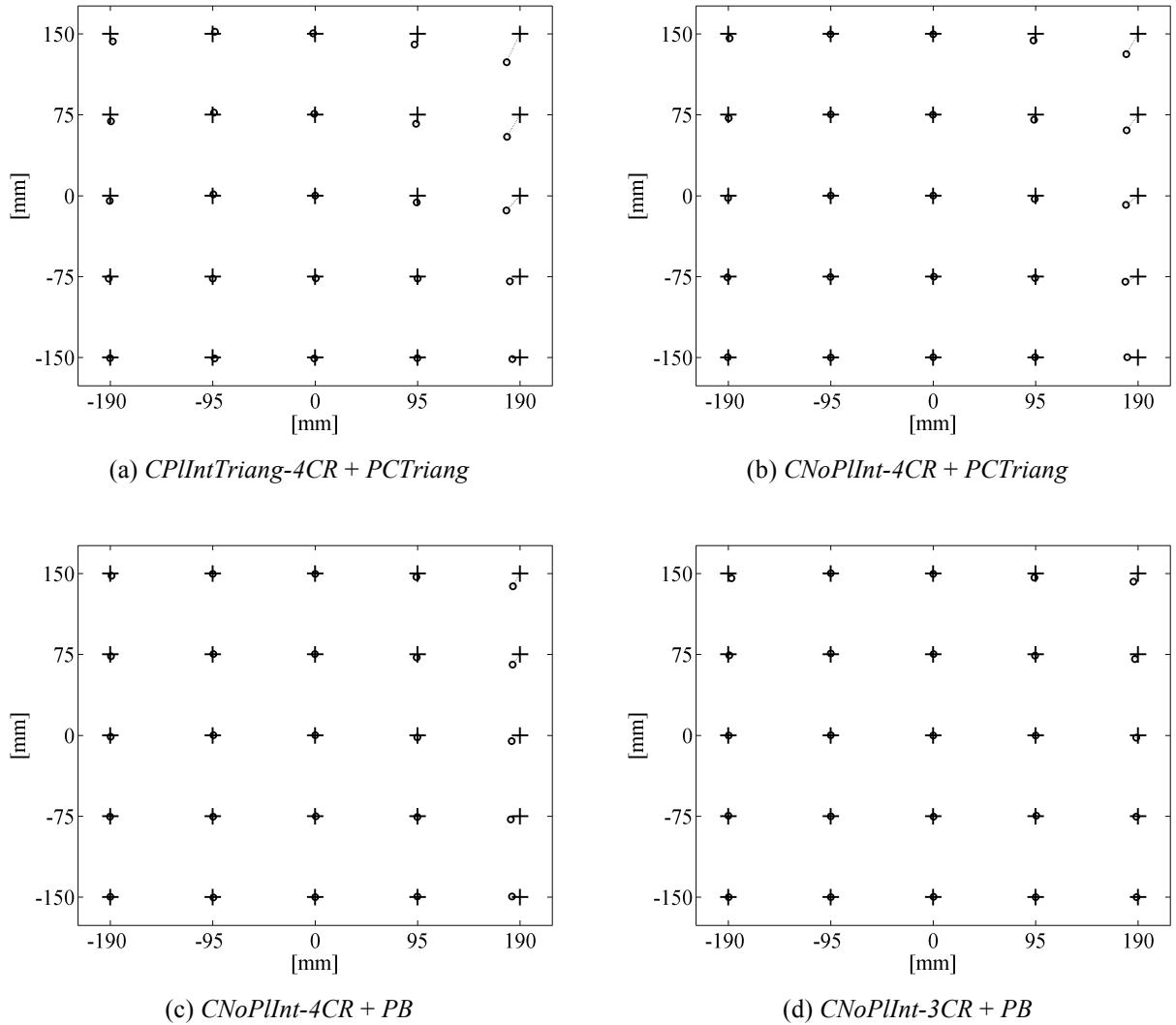


Fig. 4.11: Effect of corneal asphericity – System with 4 light sources with  $l_{1-2,y} = 0$  mm and  $o_x = 150$  mm ( $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ).

(+ : actual point-of-gaze, o : estimated point-of-gaze)

#### 4.7.1 Effects of noise in the image coordinates of the eye features and corneal asphericity as a function of system setup for a fixed position of the center of the eyeball

Having demonstrated the performance improvement using 4 light sources, the effects of noise in the image coordinates of the eye features and corneal asphericity are evaluated through simulations as a function of the camera separation (and consequently the angle of convergence of the cameras,  $\theta_{\text{opt.ax.cam's}}$ ), and as a function of the vertical coordinate of the light sources that are located at the sides of the screen. In these simulations, which are also carried out for a single

fixed position of the center of the eyeball ( $\mathbf{d} = \mathbf{d}^* = [0 \ 125 \ 650]^T$  mm) and 25 points-of-gaze arranged in a 5-by-5 rectangular grid spanning the entire screen, the synthetic eye features are generated with the ellipsoidal corneal model and noise is added to the eye features as described in Section 4.3. The point-of-gaze is estimated using the actual values of the system parameters after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center in noise-free conditions.

To evaluate the effects of noise in the eye features and corneal asphericity as a function of the camera separation, the vertical coordinate of the light sources that are located at the sides of the screen,  $l_{1-2,Y}$ , is kept constant at 0 mm, while  $o_X$  is varied from 50 mm to 200 mm in steps of 25 mm (the distance between the nodal points of the cameras varies between 100 mm and 400 mm in steps of 50 mm). The angle of convergence of the cameras,  $\theta_{\text{opt.ax.cam's}}$ , for each tested value of  $o_X$  is provided in Table 4.33.

Table 4.33: Angle of convergence of the cameras,  $\theta_{\text{opt.ax.cam's}}$ , for each tested value of  $o_X$

$o_X$ [mm]	$\theta_{\text{opt.ax.cam's}}$ [ $^\circ$ ]
50	8.02
75	12.00
100	15.96
125	19.87
150	23.74
175	27.56
200	31.32

The RMS Dispersion as a function of  $o_X$  is shown in Fig. 4.12 (recall that the RMS Dispersion is a measure of the dispersion or scattering of the noisy point-of-gaze estimates with respect to the corresponding noise-free point-of-gaze estimates, i.e., it is a measure of the effect of noise in the image coordinates of the eye features by itself, or, in other words, a measure of noise sensitivity). Note that since the *PB* and *PB-EF* methods exhibit a very similar performance, only the *PB* method is included in Fig. 4.12 (through Fig. 4.17) for the sake of clarity. The results shown in

Fig. 4.12 are consistent with the observations from Section 4.6.1:

- Increasing the angle between the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ) reduces the sensitivity to noise of all methods.
- The *CNoPlInt...* methods for the calculation of the 3-D coordinates of the center of curvature of the cornea exhibit a lower noise sensitivity than the *CPlInt...* method and the “-4CR” methods exhibit a lower noise sensitivity than the corresponding “-3CR” methods.
- The *PB* (and *PB-EF*) method for the subsequent reconstruction of the optic axis of the eye in 3-D space exhibit the lowest noise sensitivity. The *PCTriang* method exhibits a marginally lower noise sensitivity than the *PCTriangRefr* method.

The RMS Bias as a function of  $o_X$  is shown in Fig. 4.13 (recall that the bias refers to the distance of the noise-free point-of-gaze estimates from the corresponding actual points-of-gaze, and that this bias is a combination of the bias due to corneal asphericity and the inherent bias, if any). The results from this figure are consistent with observations from Sections 4.5 and 4.6.2:

- The *CNoPlInt...* methods for the calculation of the 3-D coordinates of the center of curvature of the cornea exhibit a lower bias than the *CPlInt...* method and the “-3CR” methods exhibit a lower bias than the corresponding “-4CR” methods.
- For a given method for the calculation of the 3-D coordinates of the center of curvature of the cornea, the *PB* (and *PB-EF*) method for the subsequent reconstruction of the optic axis of the eye in 3-D space exhibit the lowest bias, followed by the *PCTriangRefr* method in second place and the *PCTriang* method in third place.
- Increasing the angle between the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ) increases the bias due to corneal asphericity, as can be observed with the *CNoPlInt... + PB* methods, which have no inherent bias. On the other hand, as shown in Section 4.5, increasing the angle between the optic axes of the cameras decreases the inherent bias of the ... + *PCTriangRefr/PCTriang* methods. The decrease of the inherent bias of these methods tends to overpower the increase of the bias due to corneal asphericity. As a result, the total bias of the ... + *PCTriangRefr/PCTriang* methods decreases as  $o_X$  increases for most values of  $o_X$  tested (for the entire range of values of  $o_X$  tested for the *CPlIntTriang-4CR + PCTriang* method).

The RMS Error as a function of  $o_X$  is shown in Fig. 4.14 (here “error” refers to the distance of the noisy point-of-gaze estimates from the corresponding actual points-of-gaze, i.e., it is the total error, including the combined effects of noise in the eye features, corneal asphericity and inherent bias, if any). This figure shows that the RMS Error decreases for all methods as  $o_X$  increases. This is an expected result for the ... + *PCTriangRefr/PCTriang* methods for which the RMS Dispersion decreases as  $o_X$  increases for the entire range of values of  $o_X$  tested, and the RMS Bias decreases as  $o_X$  increases for most values of  $o_X$  tested (for the entire range of values of  $o_X$  tested for the *CPIIntTriang-4CR + PCTriang* method). In the case of the *CNoPInt-4CR + PB* method, the reduction of the noise sensitivity is almost cancelled out by the increase of the bias due to corneal asphericity. In the case of the *CNoPInt-3CR + PB* method, which exhibits the lowest RMS Error, the reduction of the noise sensitivity overpowers the increase of the bias due to corneal asphericity.

Having shown that increasing the camera separation, and hence the angle of convergence of the optic axes of the cameras ( $\theta_{\text{opt.ax.cam's}}$ ), reduces the RMS Error of the point-of-gaze estimates, it is time to make a practical consideration about  $\theta_{\text{opt.ax.cam's}}$ . If for a single camera it is possible to track the eye features over an angular range of head rotation about a vertical axis (yaw)  $\theta_{\text{yaw-single cam}}$ , then for two cameras with an angle of convergence of their optic axes  $\theta_{\text{opt.ax.cam's}}$ , the yaw angular range for which the eye features can be tracked for both cameras simultaneously is reduced to roughly  $\theta_{\text{yaw-single cam}} - \theta_{\text{opt.ax.cam's}}$ . Experience has shown that  $\theta_{\text{opt.ax.cam's}}$  greater than about 25° results in an unnatural restriction to the yaw angular range tolerated by the system. For most practical cases,  $\theta_{\text{opt.ax.cam's}}$  in the range 15-25° seems to provide a reasonable trade-off between point-of-gaze estimation error and the yaw angular range tolerated by the system. Based on these observations,  $o_X$  is set to 150 mm (300 mm between the nodal points of the cameras,  $\theta_{\text{opt.ax.cam's}} = 23.74^\circ$ ) for the remainder of the simulation study in this chapter.

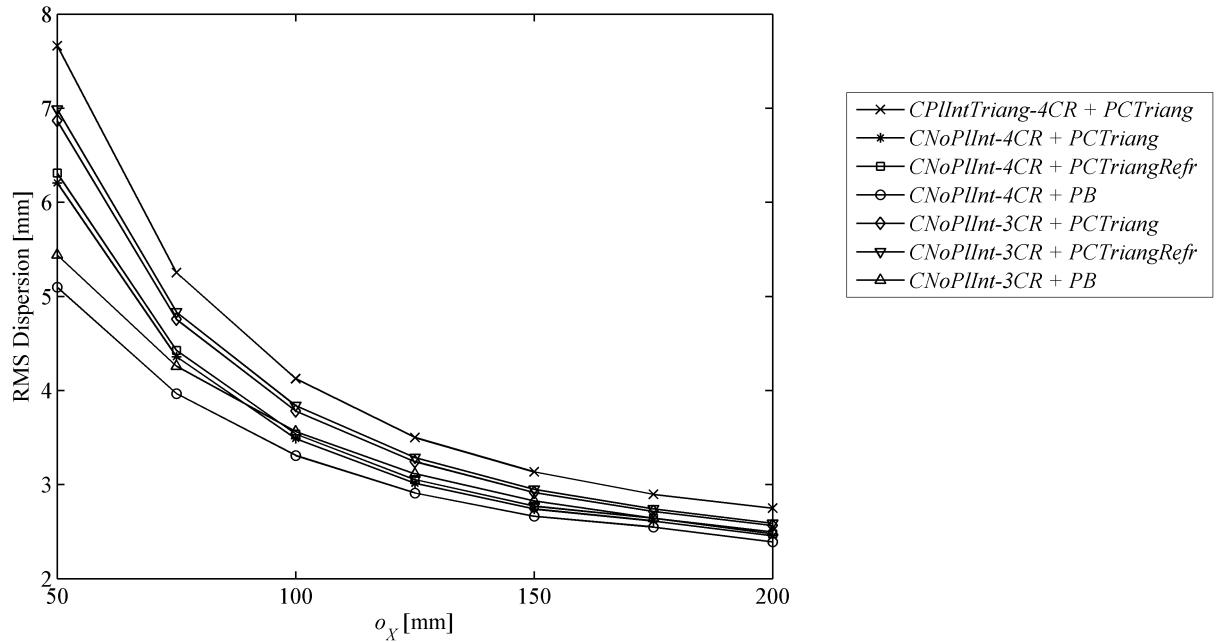


Fig. 4.12: Noise sensitivity as a function of  $o_X$  for the system with 4 light sources with  $l_{1-2,Y} = 0$  mm.

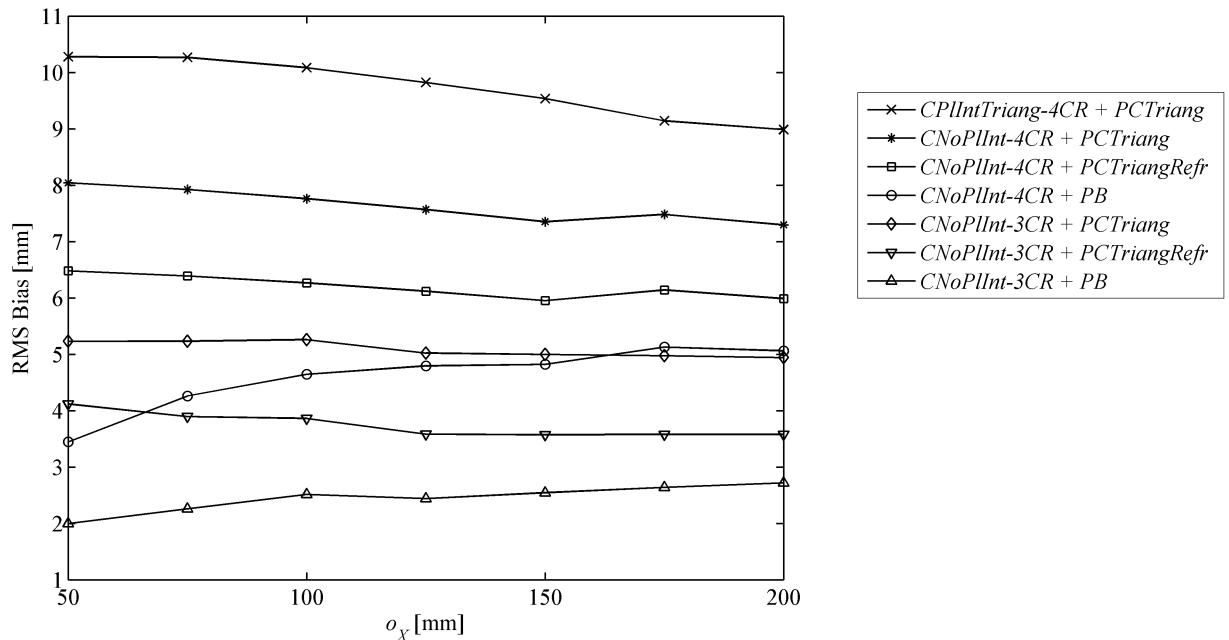


Fig. 4.13: Effect of corneal asphericity – point-of-gaze estimation bias as a function of  $o_X$  for the system with 4 light sources with  $l_{1-2,Y} = 0$  mm.

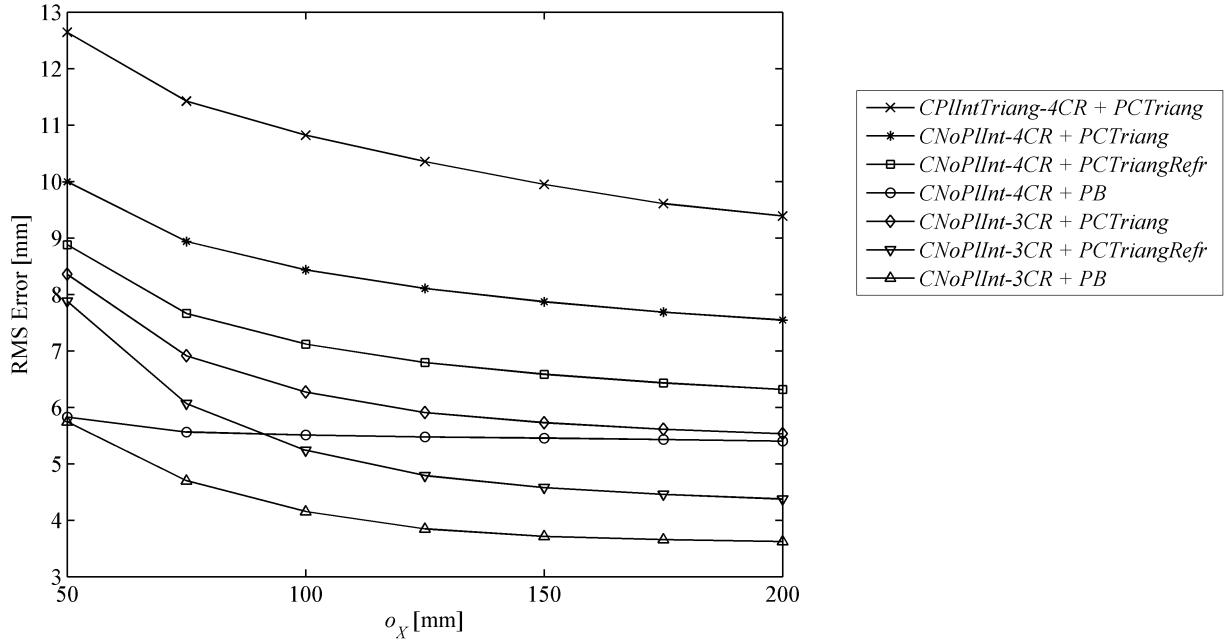


Fig. 4.14: Total point-of-gaze estimation error as a function of  $o_X$  for the system with 4 light sources with  $l_{1-2,Y} = 0$  mm.

To evaluate the effects of noise in the eye features and corneal asphericity as a function of the vertical coordinate of the light sources that are located at the sides of the screen,  $o_X$  is set to 150 mm and  $l_{1-2,Y}$  is varied from -150 mm to 150 mm in steps of 25 mm. The RMS Dispersion as a function of  $l_{1-2,Y}$  is shown in Fig. 4.15. Consistently with the observations from Section 4.6.1, the noise sensitivity of the *CPlInt...* method for the calculation of the 3-D coordinates of the center of curvature of the cornea is strongly dependent on the position of the light sources whereas the noise sensitivity of the *CNoPInt...* methods is weakly dependent on  $l_{1-2,Y}$  but it is lowest in the central range of values of  $l_{1-2,Y}$  tested. The variation of the noise sensitivity of the *CNoPInt...* methods with  $l_{1-2,Y}$  is not very significant for most practical purposes. The relative performance comparison between the different methods follows similar observations as for Fig. 4.12.

The RMS Bias as a function of  $l_{1-2,Y}$  is shown in Fig. 4.16. The RMS Bias decreases with  $l_{1-2,Y}$  for the entire range of  $l_{1-2,Y}$  tested, with the rate of decrease becoming lower for larger values of  $l_{1-2,Y}$ . The RMS Bias of the “-3CR” methods decreases with  $l_{1-2,Y}$  at a faster rate than the “-4CR” methods. The relative performance comparison between the different methods follows similar observations as for Fig. 4.13.

The RMS Error as a function of  $l_{1-2,Y}$  is shown in Fig. 4.17. The RMS Error decreases with  $l_{1-2,Y}$  for the entire range of  $l_{1-2,Y}$  tested, with the rate of decrease becoming lower for larger values of  $l_{1-2,Y}$ . The RMS Error of the “-3CR” methods decreases with  $l_{1-2,Y}$  at a faster rate than the “-4CR” methods until it tends to plateau between  $l_{1-2,Y} = 50$  mm and  $l_{1-2,Y} = 150$  mm. The relative performance comparison between the different methods follows similar observations as for Fig. 4.14. Note that the reduction of the total point-of-gaze estimation error with  $l_{1-2,Y}$  is due to the corresponding reduction of the point-of-gaze estimation bias (Fig. 4.16).

From the analysis of the RMS Error shown in Fig. 4.17 it seems reasonable to set  $l_{1-2,Y} = 50$  mm. In any case, there is another practical consideration: in order to prevent the upper eyelid from casting a shadow on the eye, which can interfere significantly with the extraction of the pupil from the images captured by the cameras, most (or all) of the illumination sources should be below the eye. For the system setup studied, the upper limit for  $l_{1-2,Y}$  arising from such practical consideration is about 50 mm. Based on this,  $l_{1-2,Y}$  is set to 50 mm for the remainder of this chapter.

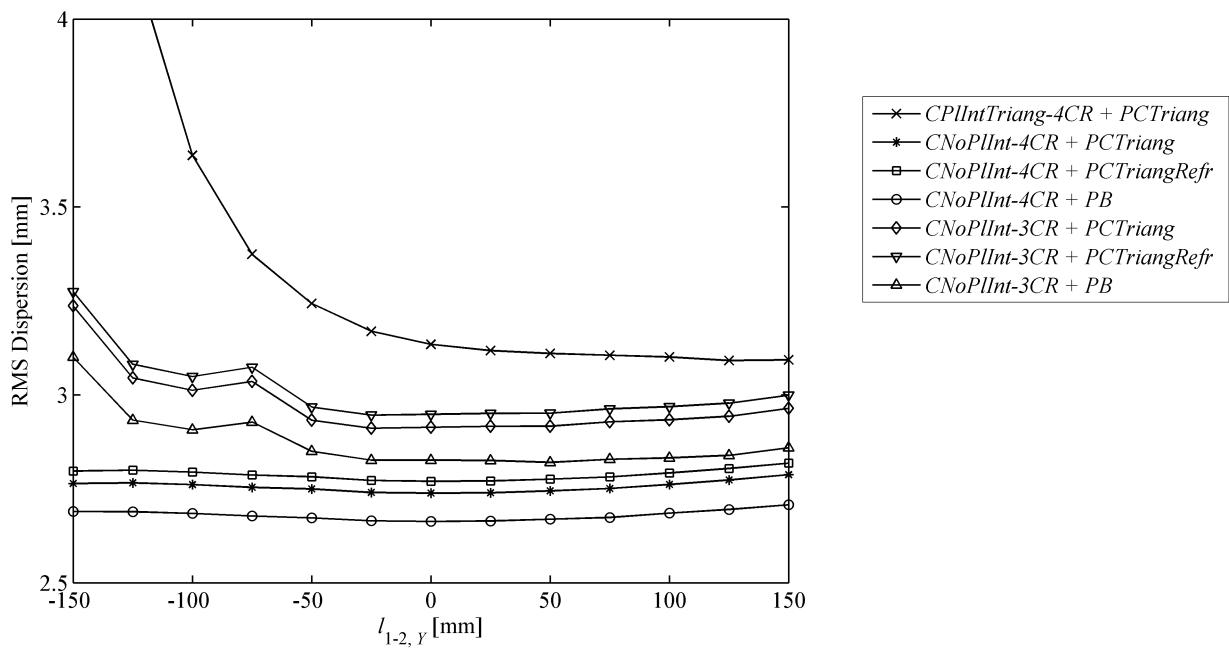


Fig. 4.15: Noise sensitivity of the system with 4 light sources as a function of  $l_{1-2,Y}$  when  $o_X = 150$  mm.

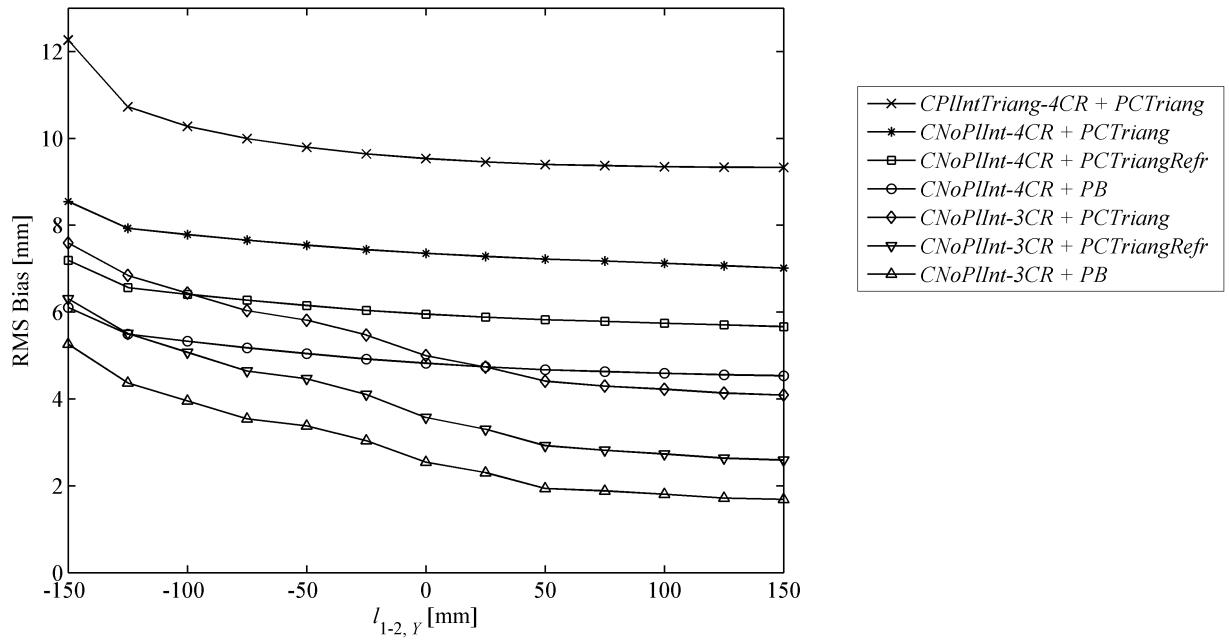


Fig. 4.16: Effect of corneal asphericity – point-of-gaze estimation bias of the system with 4 light sources as a function of  $l_{1-2,y}$  when  $o_X = 150$  mm.

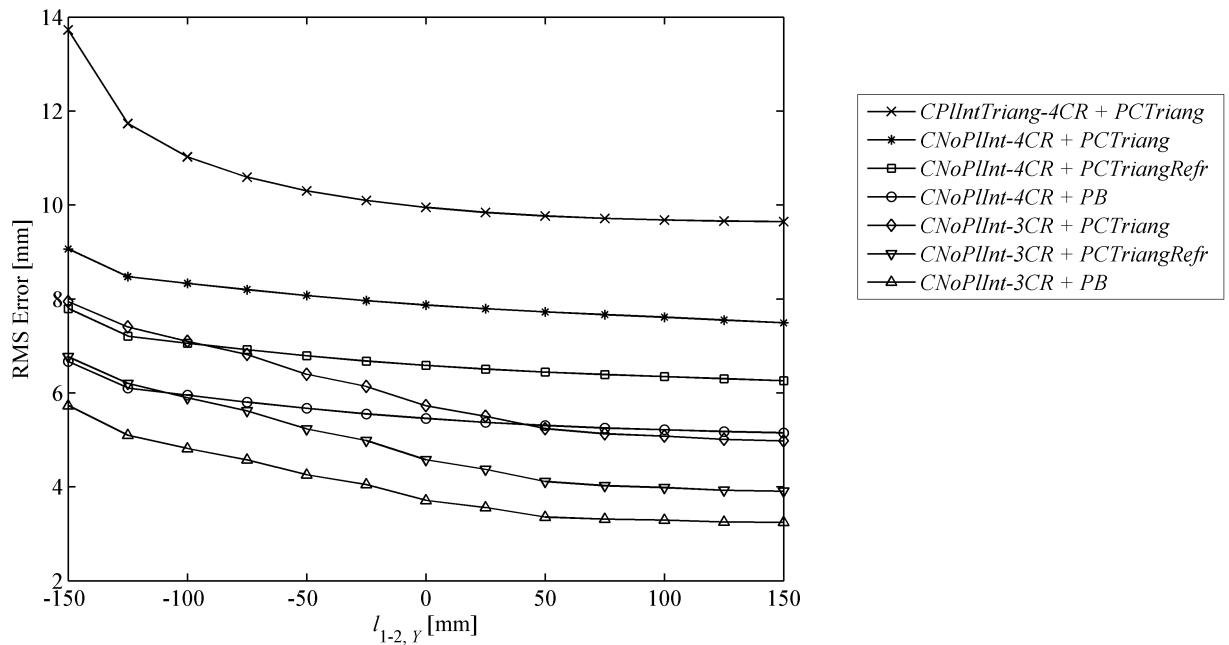


Fig. 4.17: Total point-of-gaze estimation error of the system with 4 light sources as a function of  $l_{1-2,y}$  when  $o_X = 150$  mm.

#### 4.7.2 Effects of noise in the image coordinates of the eye features and corneal asphericity for the selected system setup under head movements

Having selected a setup ( $o_X = 150$  mm,  $l_{1-2,Y} = 50$  mm) for the system with 4 light sources, its performance in the presence of noise in the image coordinates of the eye features and corneal asphericity is evaluated under head movements. In these simulations, the synthetic eye features are generated with the ellipsoidal corneal model and noise is added to the eye features. The center of the eyeball (**d**) of the simulated right eye adopts 27 different positions (Table 4.34) arranged in a 3-by-3 grid spanning, approximately, the entire volume of operation of the system (the volume of operation is defined by the intersection of the fields of view of both cameras and their depths of field –the depth of field is based on the prototype system described in Section 5.2). For each position of **d**, the simulated eye gazes at 25 points arranged in a 5-by-5 rectangular grid spanning the entire screen (same points-of-gaze as in all previous simulations). The point-of-gaze is estimated using the actual values of the system parameters after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center in noise-free conditions.

The results of these simulations for 4 of the 9 methods studied are shown in Fig. 4.18. The RMS Dispersion is presented in Table 4.35 (mm) and Table 4.36 (degrees of visual angle), the RMS Bias is presented in Table 4.37 (mm) and Table 4.38 (degrees), and the RMS Error is presented in Table 4.39 (mm) and Table 4.40 (degrees). The point-of-gaze estimation method that exhibits the lowest RMS Error is the *CNoPlInt-3CR + PB/PB-EF* method (Fig. 4.18(d)).

In order to provide a better visualization of the point-of-gaze estimation bias, the above simulations are repeated for 99 points-of-gaze arranged in a 9-by-11 rectangular grid spanning the entire screen under noise-free conditions. The results for 4 of the 9 methods studied are shown in Fig. 4.19 (same 4 methods shown in Fig. 4.18).

It should be noted, however, that in the preceding simulations there are no errors in the values of the system parameters. A study of the effect of errors in the values of several system parameters for this system setup, under the simulation conditions described in the first paragraph of this section, is presented in Appendix E. One of the findings from that study is that the only method that can have a consistently reasonable performance for all the errors in camera parameters studied is the *CPIIntTriang-4CR + PCTriang* method (Fig. 4.18(a)).

The next chapter presents preliminary experimental results obtained with adults and infants.

Table 4.34: Positions of the center of the eyeball (**d**) for the simulations with head movements [mm]

Position #	$d_X$	$d_Y$	$d_Z$
1	0	125	650
2	-50	125	650
3	50	125	650
4	-50	150	650
5	0	150	650
6	50	150	650
7	-50	95	650
8	0	95	650
9	50	95	650
10	-35	98	600
11	0	98	600
12	35	98	600
13	-35	123	600
14	0	123	600
15	35	123	600
16	-35	71	600
17	0	71	600
18	35	71	600
19	-45	152	700
20	0	152	700
21	45	152	700
22	-45	182	700
23	0	182	700
24	45	182	700
25	-45	117	700
26	0	117	700
27	45	117	700

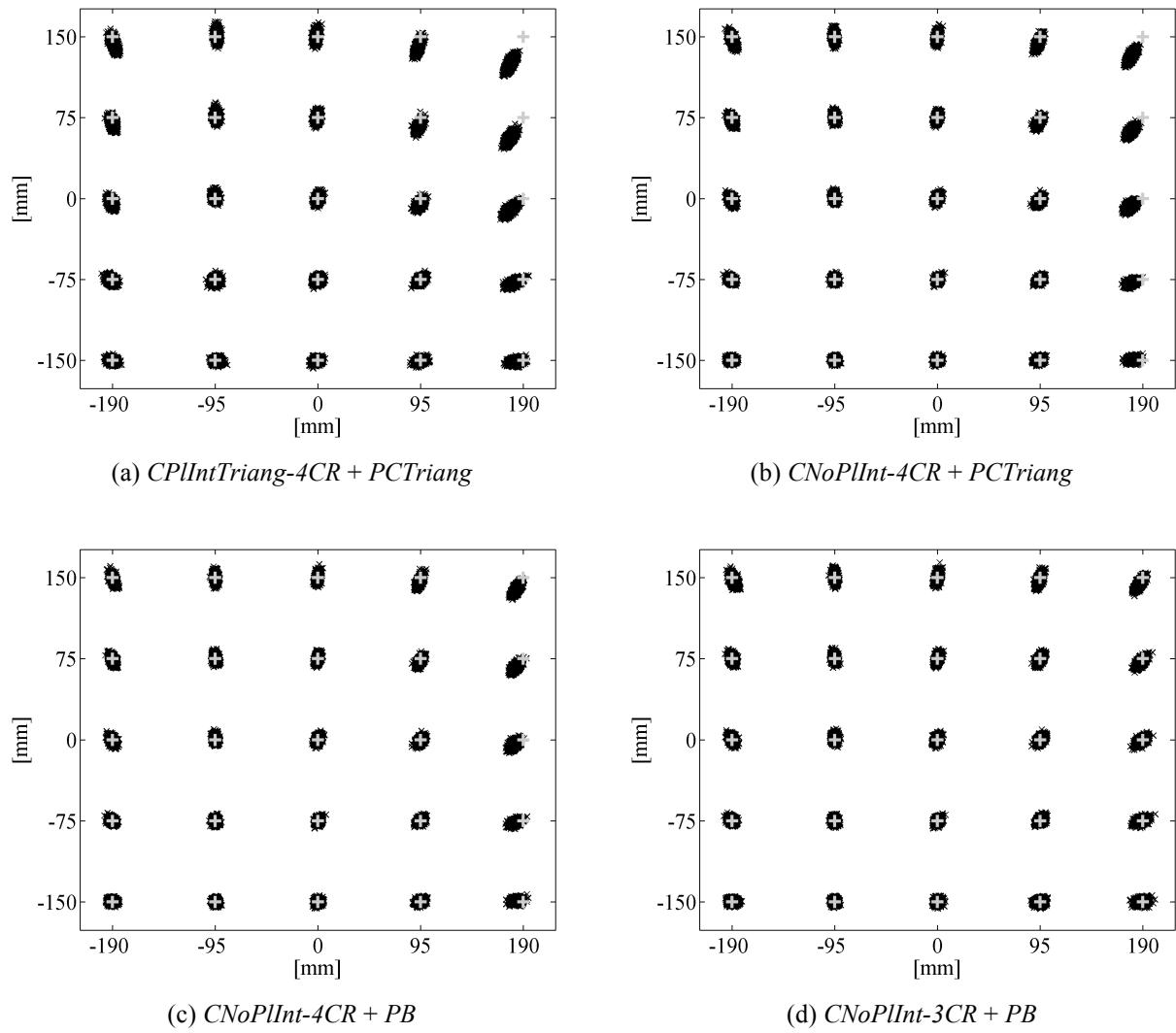


Fig. 4.18: Point-of-gaze estimation results under head movements for the system with 4 light sources with  $l_{1-2,y} = 50$  mm and  $o_x = 150$  mm.  
 (+ : actual point-of-gaze,  $\times$  : estimated point-of-gaze)

Table 4.35: Noise sensitivity – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm – *RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	3.14	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	2.97	3.01	2.89	2.89
<i>CNoPlInt-4CR</i>	2.75	2.78	2.70	2.70

Table 4.36: Noise sensitivity – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm – *RMS Dispersion [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	0.26	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	0.25	0.25	0.24	0.24
<i>CNoPlInt-4CR</i>	0.23	0.23	0.22	0.22

Table 4.37: Effect of corneal asphericity – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm – *RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	9.41	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	4.52	3.03	2.03	1.88
<i>CNoPlInt-4CR</i>	7.36	5.95	4.77	4.65

Table 4.38: Effect of corneal asphericity – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm – *RMS Bias [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	0.81	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	0.38	0.26	0.17	0.16
<i>CNoPlInt-4CR</i>	0.62	0.50	0.40	0.39

Table 4.39: Total point-of-gaze estimation error – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm  
– *RMS Error* [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	9.86	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	5.37	4.23	3.47	3.41
<i>CNoPlInt-4CR</i>	7.82	6.53	5.38	5.32

Table 4.40: Total point-of-gaze estimation error – System with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm  
– *RMS Error* [ $^{\circ}$ ]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPlIntTriang-4CR</i>	0.84	N/A	N/A	N/A
<i>CNoPlInt-3CR</i>	0.45	0.35	0.29	0.28
<i>CNoPlInt-4CR</i>	0.66	0.55	0.45	0.44

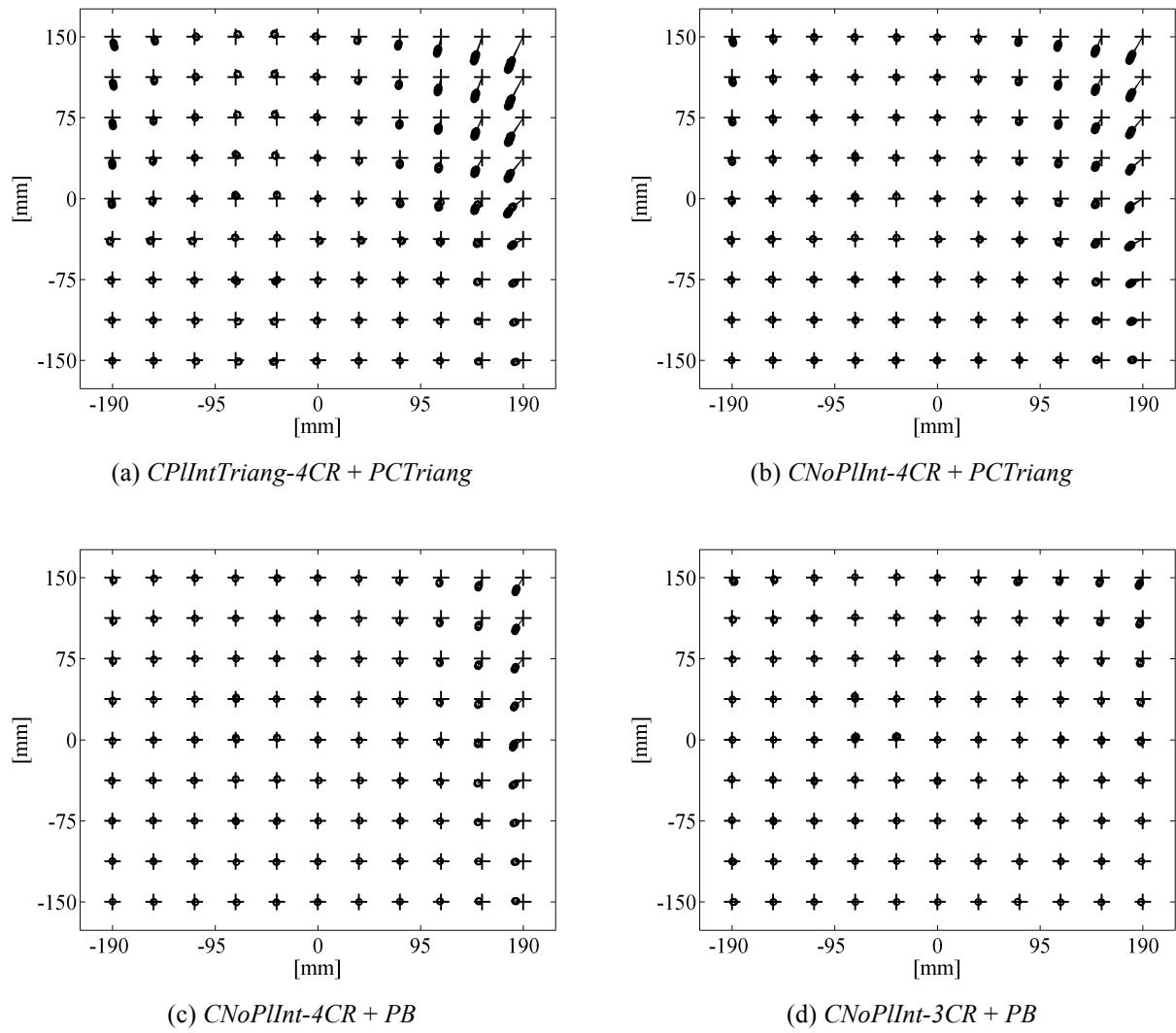


Fig. 4.19: Point-of-gaze estimation bias under head movements for the system with 4 light sources with  $l_{1-2,Y} = 50$  mm and  $o_X = 150$  mm.  
 (+ : actual point-of-gaze, ○ : estimated point-of-gaze)

# Chapter 5

## Experimental Results

### 5.1 Introduction

This chapter presents proof-of-concept experimental results obtained with a prototype system that uses two cameras and four light sources to estimate the point-of-gaze on a computer screen. The performance of the prototype system was evaluated with adults and infants<sup>11</sup>. The results obtained with the adults provide a measurement of the expected point-of-gaze estimation accuracy. The results obtained with the infants demonstrate the ability of the system to monitor infants' point-of-gaze.

Section 5.2 describes the two setups of the prototype system that were tested. Section 5.3 describes briefly the system calibration procedure (measurement / estimation of system parameters). Section 5.4 presents the experimental results with adults and Section 5.5 presents the experimental results with infants. Section 5.6 summarizes the experimental results.

### 5.2 System setups

The prototype system uses two synchronized monochrome CCD cameras (Scorpion SCOR-14SOM, Point Grey Research, Richmond, BC, Canada) with 35 mm lenses (aperture set to  $f/5.6$ ) and four near-infrared light sources (each light source consists of 20 infrared LEDs with nominal wavelength of 850 nm) attached to a 19" LCD monitor (5:4 aspect ratio) by a custom aluminum frame. The prototype system can tolerate moderate head movements before the eye features are no longer in the field of view of the cameras or are out of focus.

---

<sup>11</sup> With approval from the research ethics board from the Hospital for Sick Children ( SickKids), Toronto, Ontario, Canada.

Two setups of the prototype system were tested. The system setup used for the initial proof-of-concept experiments, including the experiments with infants presented in Section 5.5, is described in Section 5.2.1 (the initial proof-of-concept experiments are presented in [86] and [87]; the experiments with infants are included in [87]). A new system setup that is based on the guidelines derived from the simulation study from Chapter 4 is described in Section 5.2.2 (the positions of the light sources, the nodal points of the cameras and the point of intersection of the optic axes of the cameras in the new system setup are within 1 cm of those of the system setup simulated in Section 4.7.2). This new system setup was used for all the experiments with adults described in Section 5.4.

### 5.2.1 Initial system setup

The initial system setup is shown in Fig. 5.1. In this setup, the nodal points of the cameras were approximately 20 cm apart and the cameras were oriented such that their optic axes intersected at a distance of approximately 65 cm from the screen (typical viewing distance). Both cameras were configured to operate in a standard imaging mode with a resolution of 1280 pixels by 960 pixels and a frame rate of 15 Hz. In these conditions, the range of lateral head movements was about 10 cm when the point-of-gaze was estimated for a specific eye, about 16 cm when the point-of-gaze was estimated for either eye and about 4 cm when the point-of-gaze was estimated for both eyes simultaneously (binocular mode). The range of vertical head movements was about 8 cm and the range of backwards/forward head movements was about 10 cm.

### 5.2.2 New system setup

The new system setup, which is based on the guidelines derived from the simulation study from Chapter 4, is shown in Fig. 5.2. In this setup, the nodal points of the cameras are approximately 30 cm apart and, similarly to the initial system setup, the cameras are oriented such that their optic axes intersect at a distance of approximately 65 cm from the screen. Both cameras are configured to operate in a custom imaging mode with a resolution of 1392 pixels by 900 pixels and a frame rate of 20 Hz (1392 pixels is the maximum horizontal resolution of the image sensor whereas 900 pixels is the maximum vertical resolution that allows the cameras to operate at a frame rate of 20 Hz –the maximum vertical resolution of the image sensor is 1040 pixels). In these conditions, the range of lateral head movements is about 1 cm larger while the range of vertical head movements is about 1 cm smaller than in the initial setup. The range of

backwards/forward head movements remains the same. Note that the light sources at the sides of the monitor are about 5 cm higher than in the initial setup.

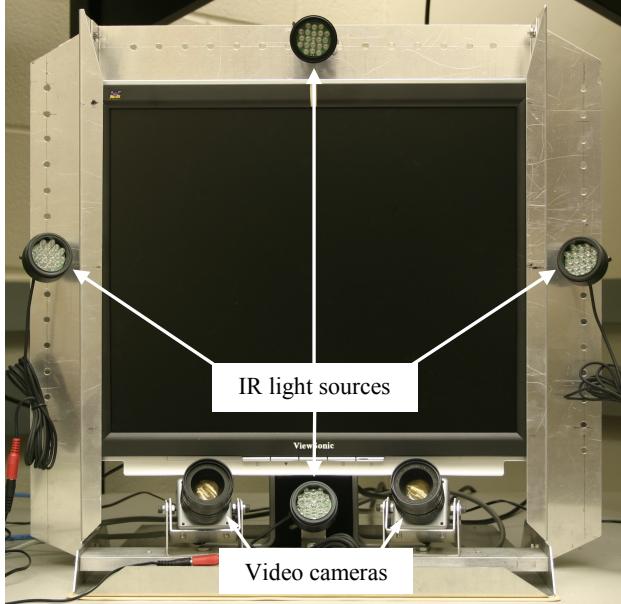


Fig. 5.1: Initial system setup.



Fig. 5.2: New system setup.

### 5.3 System calibration

In order to be able to estimate the point-of-gaze, the positions of the light sources and the intrinsic and extrinsic camera parameters must be known accurately. The positions of the light sources,  $\mathbf{I}_i$ , with respect to the WCS (Section 2.5.1), which is attached to the LCD monitor, are measured directly using rulers and calipers. The intrinsic camera parameters and the position and orientation of the cameras with respect to the WCS (extrinsic camera parameters) are determined through the camera calibration procedure outlined in the next paragraph.

Since the cameras that capture images of the eye(s) (eye cameras) are positioned under the screen (Fig. 5.1, Fig. 5.2), and therefore cannot observe the monitor, an auxiliary camera that views both the computer screen and the region of operation of the system (Fig. 5.3) is used together with a double-sided planar checkerboard pattern. The calibration is based on a camera calibration toolbox for MATLAB® [104] and the entire calibration procedure can be summarized as follows:

(1) Images of the calibration pattern, at several different orientations within the region of operation of the system, are captured simultaneously by the two eye cameras and the auxiliary camera. (2) By using the corners of the checkerboard pattern from all the different views, the intrinsic parameters of the three cameras, and the relative position and orientation of the two eye cameras with respect to the auxiliary camera are calculated. (3) A checkerboard pattern is then displayed on the screen, and the position and orientation of the auxiliary camera with respect to the WCS is calculated. (4) Using this information, the position and orientation of the eye cameras with respect to the WCS are determined.

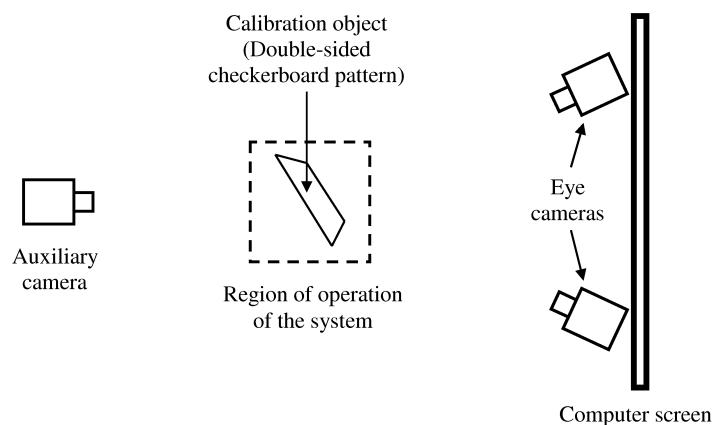


Fig. 5.3: Top-view schematic representation of the camera calibration setup.

Since all the system components are fixed with respect to each other, the system calibration procedure needs to be performed only once during system setup and is simpler than the one described in [80].

## 5.4 Experimental results with adults

The experiments with adults were carried out with three subjects without eyeglasses or contact lenses using the new system setup (Section 5.2.2) in binocular mode (both eyes were tracked simultaneously). The head of each subject was placed at 65 cm (nominal distance), 60 cm (minimum distance) and 70 cm (maximum distance) from the screen. For each head position,

each subject completed a number of trials (2 for Subjects 1 and 2, and 7 for Subject 3). In each trial, the subject was asked to sequentially fixate 25 target points (arranged in a 5-by-5 rectangular grid) on the computer screen. For each target point, 50 estimates (2.5 seconds @ 20 estimates/second) of the image coordinates of the centers of the corneal reflections and the parameters of the ellipse (center, major and minor axes, orientation of the major axis) fitted to the pupil-iris boundary points extracted from each eye image were recorded. The point-of-gaze was then estimated with seven different point-of-gaze estimation methods (the methods from Table 4.30 except the ... + *PB* methods). The subject-specific angular deviation of the visual axis from the optic axis,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , was determined when the head was at the nominal distance from the screen and the subject fixated the target point at the center of the screen.

For each of the trials, the mean point-of-gaze estimate was calculated for each target fixation point. It is assumed that, due to the averaging, the effect of noise in the image coordinates of the eye features on the mean point-of-gaze estimates is, in general, negligible and, therefore, the mean point-of-gaze estimates can be used to evaluate the point-of-gaze estimation bias. The RMS Bias (RMS distance between the mean point-of-gaze estimates and the corresponding target fixation points) for the left and right eyes corresponding to the seven point-of-gaze estimation methods tested is summarized, respectively, in Table 5.1 and Table 5.2 for Subject 1, Table 5.9 and Table 5.10 for Subject 2, and Table 5.17 and Table 5.18 for Subject 3. For Subject 1, the relative performance of the three different methods for the calculation of the 3-D coordinates of the center of curvature of the cornea is qualitatively similar to the simulation results from Section 4.7.2 (Table 4.37): the *CPIIntTriang-4CR* method exhibits the largest RMS Bias whereas the *CNoPInt-3CR* method exhibits the smallest RMS Bias. In contrast, for Subjects 2 and 3, the smallest bias corresponds to the *CNoPInt-4CR* method. The relative performance of the three different methods for the subsequent reconstruction of the optic axis of the eye in 3-D space is not very consistent with the simulation results from Section 4.7.2 as the method that exhibits the lowest RMS Bias varies from case to case. It should be noted, however, that the performance differences between the different methods for the subsequent reconstruction of the optic axis of the eye are, in general, not significant for the experimental results. Overall, the range of the RMS Bias exhibited by the seven different methods for the estimation of the point-of-gaze is smaller for the experiments (Table 5.1, Table 5.2, Table 5.9, Table 5.10, Table 5.17 and Table 5.18) than for the simulations (Table 4.37).

The differences between the experimental RMS Bias and the RMS Bias observed in the simulations can be explained by taking into account that (a) for at least two of the subjects (Subjects 1 and 3) the degree of corneal asphericity is lower (known from corneal topography results) than in the simulations, resulting in a smaller contribution of corneal asphericity to the total bias; (b) for all three subjects the pupil diameter was smaller (for Subjects 1 and 2, significantly smaller) than in the simulations, resulting in a smaller inherent bias of the ... +  $PCTriang$  and ... +  $PCTriangRefr$  methods; (c) there are fixation inaccuracies that arise from the fact that the extent of the fovea is not infinitesimal (it corresponds to 0.6–1° of visual angle [27]) and from small eye movements that occur normally during fixation [112], [113] (these fixation inaccuracies, which were not taken into account in the simulations, are most critical during the single-point personal calibration procedure); and (d) there are occasional errors in the extracted image coordinates of the eye features that cannot be modeled as a zero-mean Gaussian process (these occasional errors from the image processing algorithms are, in fact, very difficult to model). The bias due to corneal asphericity and the inherent bias depend on the point-of-gaze estimation method used whereas the bias due to the fixation inaccuracies and the occasional non-zero-mean errors from the image processing algorithms do not. As the contribution of the method-dependent bias becomes smaller ((a) and (b)), the relative contribution of the method-independent bias ((c) and (d)) becomes more significant and the differences in the total bias exhibited by the different methods become less pronounced, to the extent that point-of-gaze estimation methods that exhibit the smallest bias in the simulations do not necessarily exhibit the smallest bias in the experimental results. Nevertheless, the RMS Bias is less than 7 mm (equivalent to about 0.6° of visual angle at a distance of 65 cm) for all three subjects and all seven point-of-gaze estimation methods tested.

The mean point-of-gaze estimates corresponding to Subject 1, obtained with three of the seven point-of-gaze estimation methods tested, are presented in Fig. 5.4 for the left and right eyes in separate plots and in Fig. 5.5 for both eyes in a single plot to show the consistency between the point-of-gaze estimates from both eyes. These figures illustrate how the point-of-gaze estimation bias is reduced from the  $CPIIntTriang-4CR + PCTriang$  method to the  $CNoPPIInt-4CR + PB-EF$  method and from the  $CNoPPIInt-4CR + PB-EF$  method to the  $CNoPPIInt-3CR + PB-EF$  method, particularly for the top half of the screen (note that the bias reduction is more pronounced for the left eye). The reduction of the point-of-gaze estimation bias results in a better consistency

between the point-of-gaze estimates from the left eye and the point-of-gaze estimates from the right eye, as it can be seen in Fig. 5.5. The mean point-of-gaze estimates corresponding to Subjects 2 and 3 are presented, respectively, in Fig. 5.7 and Fig. 5.8, and Fig. 5.10 and Fig. 5.11 (recall that Subject 3 completed 7 trials for each of the 3 head positions, that is, 21 trials in total). Even though the differences in the bias exhibited by the three point-of-gaze estimation methods shown is less pronounced than for Subject 1, overall, the *CNoPlInt-... + PB-EF* methods exhibit smaller bias than the *CPIIntTriang-4CR + PCTriang* method, which in turn results in better consistency between the point-of-gaze estimates from the left and right eyes. Note that the scattering of the mean point-of-gaze estimates is due to fixation inaccuracies (see (c) in the previous paragraph) and errors from the image processing algorithms (see (d) in the previous paragraph).

The RMS Dispersion (RMS distance between the point-of-gaze estimates and the corresponding mean point-of-gaze estimates) for the left and right eyes corresponding to the seven point-of-gaze estimation methods tested is summarized, respectively, in Table 5.3 and Table 5.4 for Subject 1, Table 5.11 and Table 5.12 for Subject 2, and Table 5.19 and Table 5.20 for Subject 3. For both eyes of all three subjects the relative performance of the three different methods for the calculation of the 3-D coordinates of the center of curvature of the cornea is similar to the simulation results from Section 4.7.2 (Table 4.35): the *CPIIntTriang-4CR* method exhibits the largest dispersion whereas the *CNoPlInt-4CR* method exhibits the smallest dispersion. The three different methods for the subsequent reconstruction of the optic axis of the eye in 3-D space exhibit very similar performances. Overall, the experimental RMS Dispersion (Table 5.3, Table 5.4, Table 5.11, Table 5.12, Table 5.19, and Table 5.20) is close to RMS Dispersion observed in the simulations from Section 4.7.2 (Table 4.35).

The RMS Error (RMS distance between the point-of-gaze estimates and the corresponding target fixation points) in mm for the left and right eyes corresponding to the seven point-of-gaze estimation methods tested is summarized, respectively, in Table 5.5 and Table 5.6 for Subject 1, Table 5.13 and Table 5.14 for Subject 2, and Table 5.21 and Table 5.22 for Subject 3. The RMS Error in degrees of visual angle is summarized in Table 5.7 and Table 5.8 for Subject 1, Table 5.15 and Table 5.16 for Subject 2, and Table 5.23 and Table 5.24 for Subject 3. For Subject 1, the relative performance of the three different methods for the calculation of the 3-D coordinates of the center of curvature of the cornea is qualitatively similar to the simulation results from

Section 4.7.2 (Table 4.39 (mm) and Table 4.40 (degrees)) for the left eye: the *CPlIntTriang-4CR* method exhibits the largest RMS Error whereas the *CNoPlInt-3CR* method exhibits the smallest RMS Error. For his right eye, however, the *CNoPlInt-4CR + PCTriangRefr* and *CNoPlInt-4CR + PB-EF* methods exhibit a slightly smaller RMS Error than the *CNoPlInt-3CR + PCTriangRefr* and *CNoPlInt-3CR + PB-EF* methods, opposite to the simulation results. A plausible explanation for this observation is that while the RMS Bias of the *CNoPlInt-4CR + PCTriangRefr* and *CNoPlInt-4CR + PB-EF* methods is only about 2-3 % larger than that of the *CNoPlInt-3CR + PCTriangRefr* and *CNoPlInt-3CR + PB-EF* methods, the RMS Dispersion of the *CNoPlInt-4CR + PCTriangRefr* and *CNoPlInt-4CR + PB-EF* methods is about 7 % smaller than that of the *CNoPlInt-3CR + PCTriangRefr* and *CNoPlInt-3CR + PB-EF* methods, suggesting that the reduction of the dispersion by the *CNoPlInt-4CR + PCTriangRefr* and *CNoPlInt-4CR + PB-EF* methods overpowers the reduction of the bias by the *CNoPlInt-3CR + PCTriangRefr* and *CNoPlInt-3CR + PB-EF* methods, thereby resulting in a lower RMS Error for the *CNoPlInt-4CR + PCTriangRefr* and *CNoPlInt-4CR + PB-EF* methods. Regarding the relative performance of the three different methods for the calculation of the 3-D coordinates of the center of curvature of the cornea for Subjects 2 and 3, the smallest RMS Error corresponds to the *CNoPlInt-4CR* method for both eyes. This is an expected result since for these two subjects the *CNoPlInt-4CR* method exhibits both the smallest RMS Bias and the smallest RMS Dispersion. The relative performance of the three different methods for the subsequent reconstruction of the optic axis of the eye in 3-D space is not very consistent with the simulation results from Section 4.7.2 as the method that exhibits the lowest RMS Error varies from case to case. It should be noted, however, that the performance differences between the different methods for the subsequent reconstruction of the optic axis of the eye are, in general, not significant for the experimental results. Overall, similarly to the RMS Bias, the range of the RMS Error exhibited by the seven different methods for the estimation of the point-of-gaze is smaller for the experiments (Table 5.5, Table 5.6, Table 5.13, Table 5.14, Table 5.21 and Table 5.22 (mm); Table 5.7, Table 5.8, Table 5.15, Table 5.16, Table 5.23 and Table 5.24 (degrees)) than for the simulations (Table 4.39 (mm) and Table 4.40 (degrees)). The reduced range of the experimental RMS Error is due to the behavior of the bias described earlier. Nevertheless, despite the small discrepancies between experimental and simulation results, the experimental results presented in this section are very satisfactory as they show that it is possible to obtain an RMS Error of 0.4-0.6° of visual angle (Table 5.7, Table 5.8, Table 5.15, Table 5.16, Table 5.23 and Table 5.24). Such point-of-gaze estimation accuracy is

comparable to the best commercially available systems, which use personal calibration procedures that require the subject to sequentially fixate multiple target points.

All the point-of-gaze estimates obtained with the three selected point-of-gaze estimation methods (*CPIIntTriang-4CR + PCTriang* method, *CNoPPIInt-4CR + PB-EF* method and *CNoPPIInt-3CR + PB-EF* method) are presented in Fig. 5.6 for Subject 1, Fig. 5.9 for Subject 2 and Fig. 5.12 for Subject 3 (note that since Subject 3 completed 21 trials and, nominally, 50 point-of-gaze estimates were obtained for each target fixation point in each trial, there are nominally 1050 point-of-gaze estimates per target fixation point in Fig. 5.12). Although, in general, it is somewhat difficult to appreciate the differences in the dispersion of the point-of-gaze estimates exhibited by the different point-of-gaze estimation methods, it is fairly easy to see the differences in dispersion between Fig. 5.12(a) (*CPIIntTriang-4CR + PCTriang* method) and Fig. 5.12(b) (*CNoPPIInt-4CR + PB-EF* method). The differences in the bias exhibited by the different methods can be also observed in Fig. 5.6, Fig. 5.9 and Fig. 5.12.

Table 5.1: Experimental point-of-gaze estimation error for Subject 1 – *Left eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	5.53	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.86	3.57	3.61
<i>CNoPPIInt-4CR</i>	4.95	4.41	4.59

Table 5.2: Experimental point-of-gaze estimation error for Subject 1 – *Right eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	4.51	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.80	3.83	3.69
<i>CNoPPIInt-4CR</i>	4.07	3.93	3.78

Table 5.3: Experimental point-of-gaze estimation error for Subject 1 – *Left eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	3.32	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.19	3.20	3.19
<i>CNoPPIInt-4CR</i>	2.93	2.94	2.94

Table 5.4: Experimental point-of-gaze estimation error for Subject 1 – *Right eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	3.34	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.17	3.18	3.13
<i>CNoPPIInt-4CR</i>	2.95	2.96	2.92

Table 5.5: Experimental point-of-gaze estimation error for Subject 1 – *Left eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	6.45	N/A	N/A
<i>CNoPPIInt-3CR</i>	5.01	4.80	4.82
<i>CNoPPIInt-4CR</i>	5.75	5.30	5.44

Table 5.6: Experimental point-of-gaze estimation error for Subject 1 – *Right eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	5.60	N/A	N/A
<i>CNoPPIInt-3CR</i>	4.95	4.99	4.84
<i>CNoPPIInt-4CR</i>	5.02	4.92	4.78

Table 5.7: Experimental point-of-gaze estimation error for Subject 1 – *Left eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.54	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.41	0.39	0.40
<i>CNoPPIInt-4CR</i>	0.48	0.44	0.45

Table 5.8: Experimental point-of-gaze estimation error for Subject 1 – *Right eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.46	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.41	0.41	0.40
<i>CNoPPIInt-4CR</i>	0.41	0.40	0.39

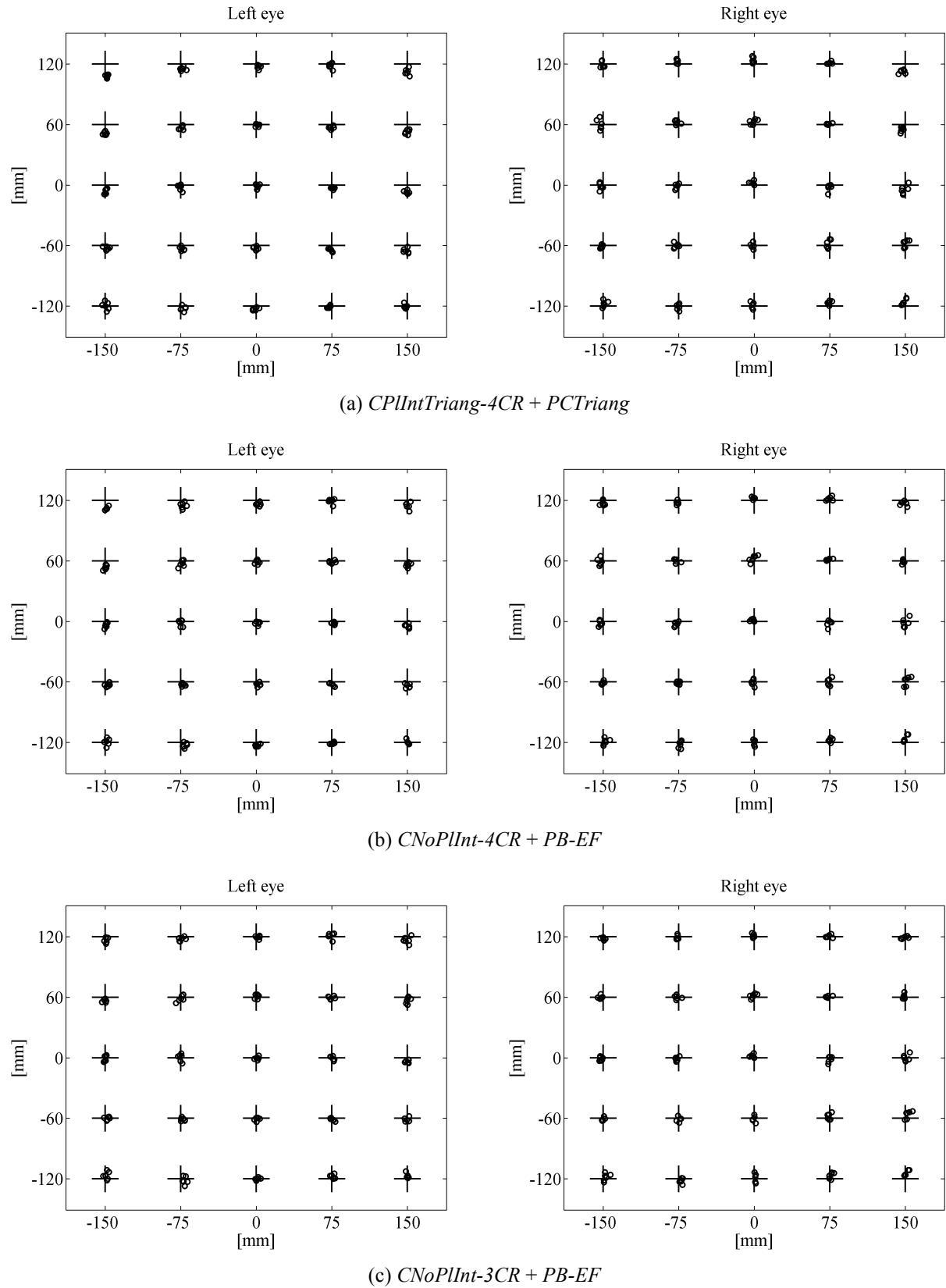


Fig. 5.4: Experimental mean point-of-gaze estimates for Subject 1.  
(+ : target fixation point, ○ : estimated point-of-gaze)

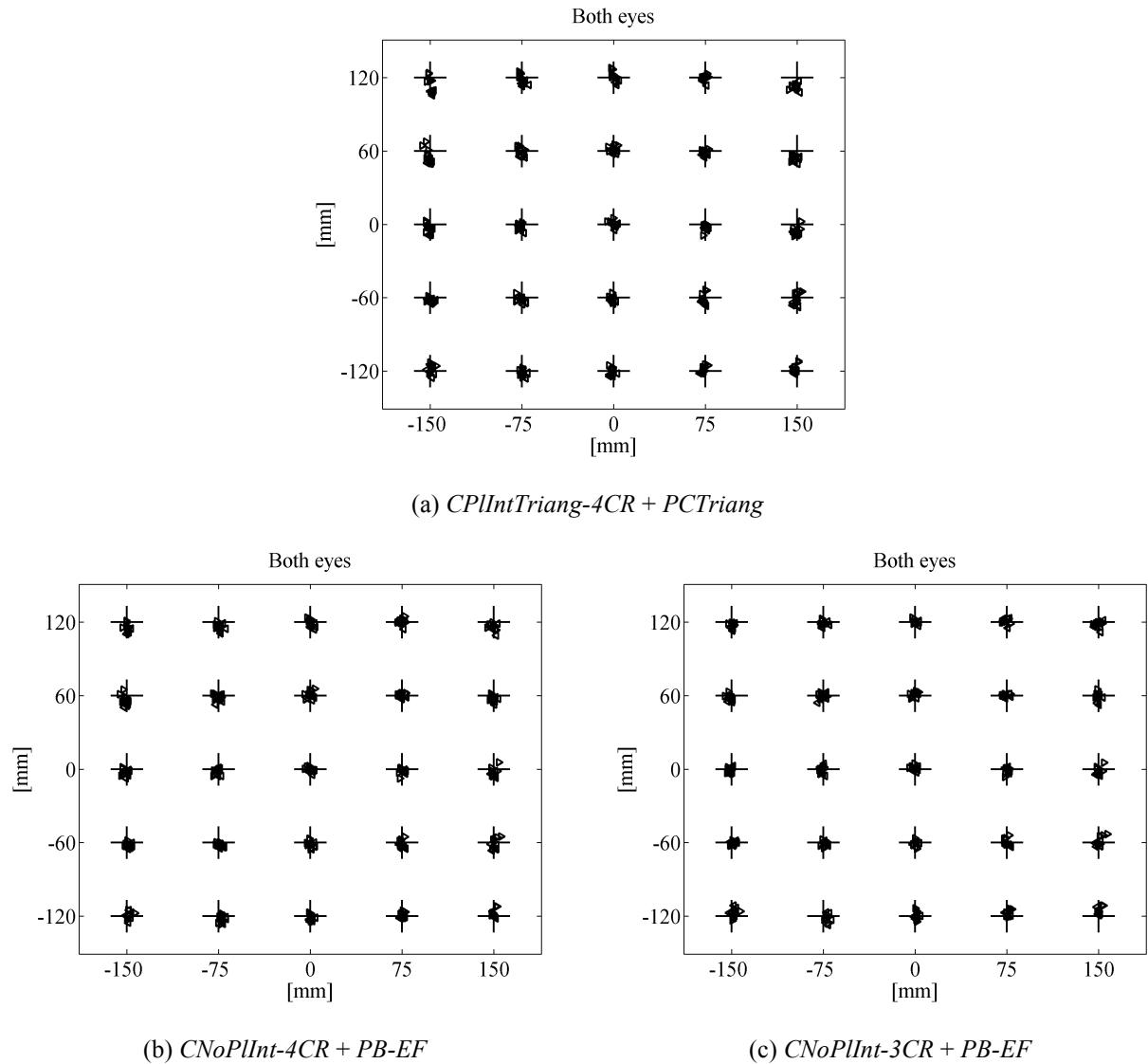


Fig. 5.5: Experimental mean point-of-gaze estimates for Subject 1.  
 (+ : target fixation point,  $\triangleleft$  : estimated point-of-gaze for the left eye,  $\triangleright$  : estimated point-of-gaze for the right eye)

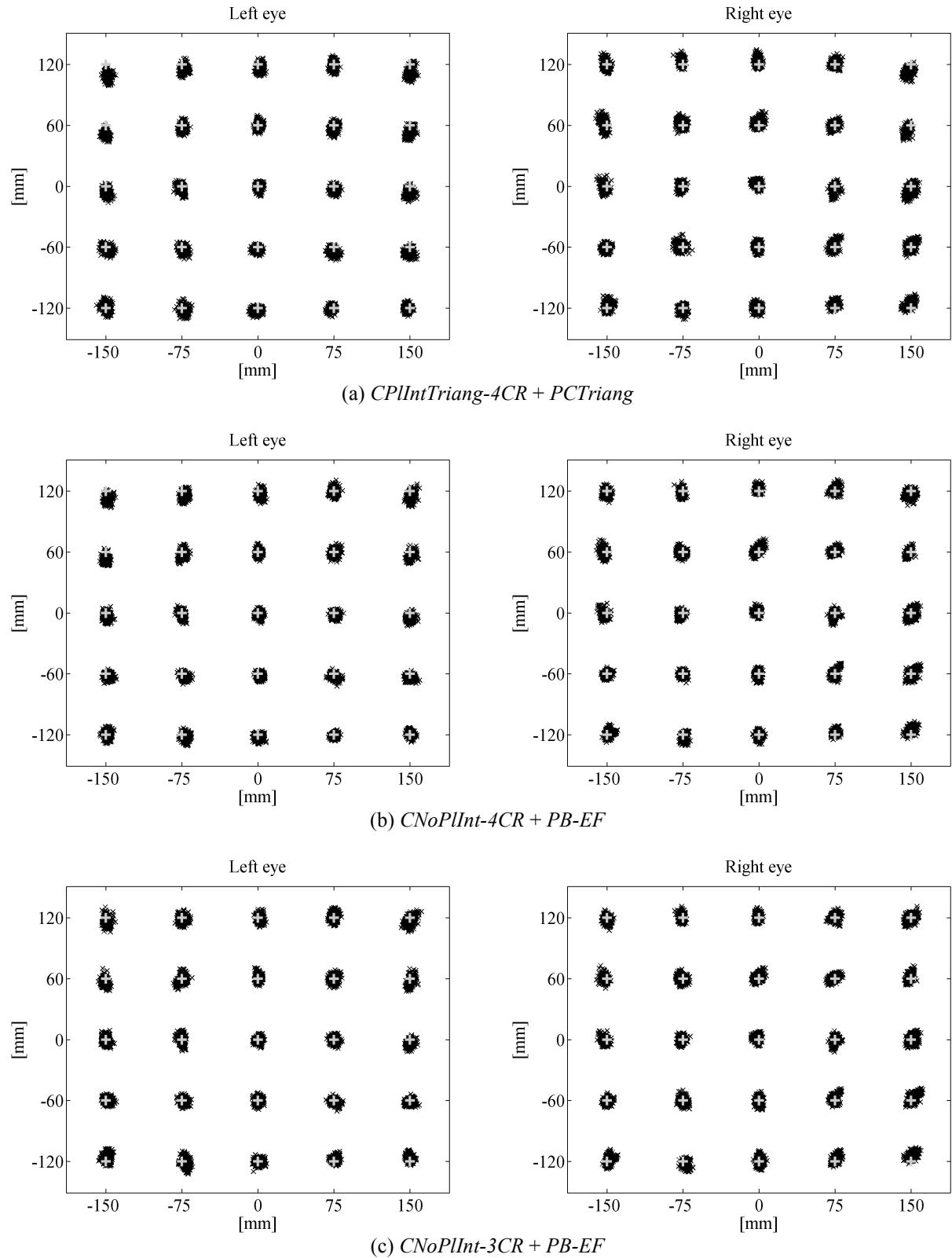


Fig. 5.6: Experimental point-of-gaze estimates for Subject 1.  
(+ : target fixation point,  $\times$  : estimated point-of-gaze)

Table 5.9: Experimental point-of-gaze estimation error for Subject 2 – *Left eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	6.71	N/A	N/A
<i>CNoPPIInt-3CR</i>	6.68	6.70	6.72
<i>CNoPPIInt-4CR</i>	6.41	6.30	6.33

Table 5.10: Experimental point-of-gaze estimation error for Subject 2 – *Right eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	5.77	N/A	N/A
<i>CNoPPIInt-3CR</i>	5.84	6.05	5.90
<i>CNoPPIInt-4CR</i>	5.40	5.35	5.36

Table 5.11: Experimental point-of-gaze estimation error for Subject 2 – *Left eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	2.84	N/A	N/A
<i>CNoPPIInt-3CR</i>	2.71	2.72	2.67
<i>CNoPPIInt-4CR</i>	2.50	2.51	2.47

Table 5.12: Experimental point-of-gaze estimation error for Subject 2 – *Right eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	3.18	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.02	3.04	2.99
<i>CNoPPIInt-4CR</i>	2.78	2.80	2.76

Table 5.13: Experimental point-of-gaze estimation error for Subject 2 – *Left eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	7.28	N/A	N/A
<i>CNoPPIInt-3CR</i>	7.23	7.25	7.26
<i>CNoPPIInt-4CR</i>	6.88	6.79	6.81

Table 5.14: Experimental point-of-gaze estimation error for Subject 2 – *Right eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	6.58	N/A	N/A
<i>CNoPPIInt-3CR</i>	6.58	6.77	6.62
<i>CNoPPIInt-4CR</i>	6.08	6.05	6.03

Table 5.15: Experimental point-of-gaze estimation error for Subject 2 – *Left eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.59	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.58	0.58	0.58
<i>CNoPPIInt-4CR</i>	0.56	0.55	0.55

Table 5.16: Experimental point-of-gaze estimation error for Subject 2 – *Right eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.54	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.54	0.56	0.55
<i>CNoPPIInt-4CR</i>	0.50	0.50	0.50

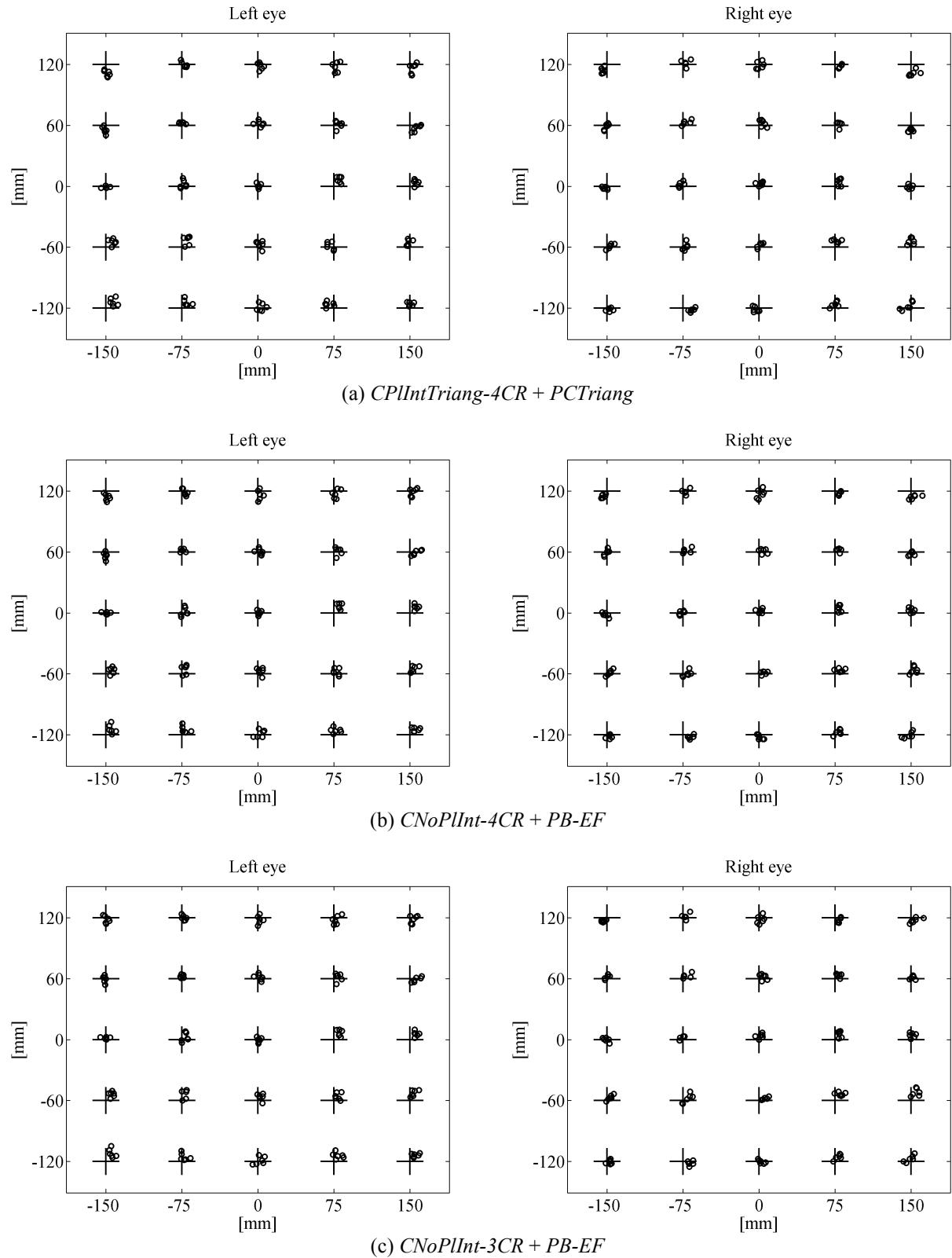


Fig. 5.7: Experimental mean point-of-gaze estimates for Subject 2.  
 (+ : target fixation point, ○ : estimated point-of-gaze)

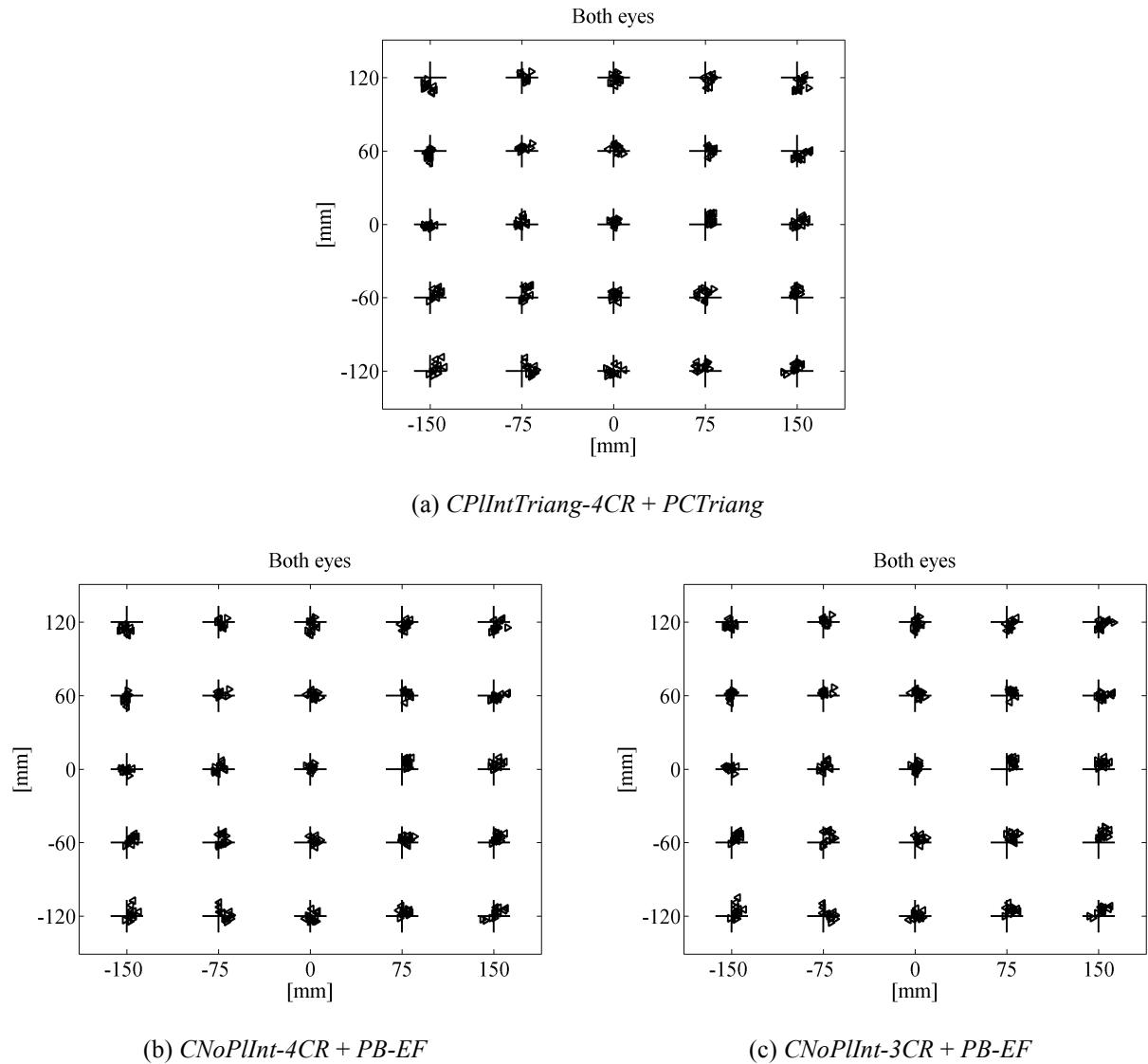


Fig. 5.8: Experimental mean point-of-gaze estimates for Subject 2.  
 (+ : target fixation point,  $\triangleleft$  : estimated point-of-gaze for the left eye,  $\triangleright$  : estimated point-of-gaze for the right eye)

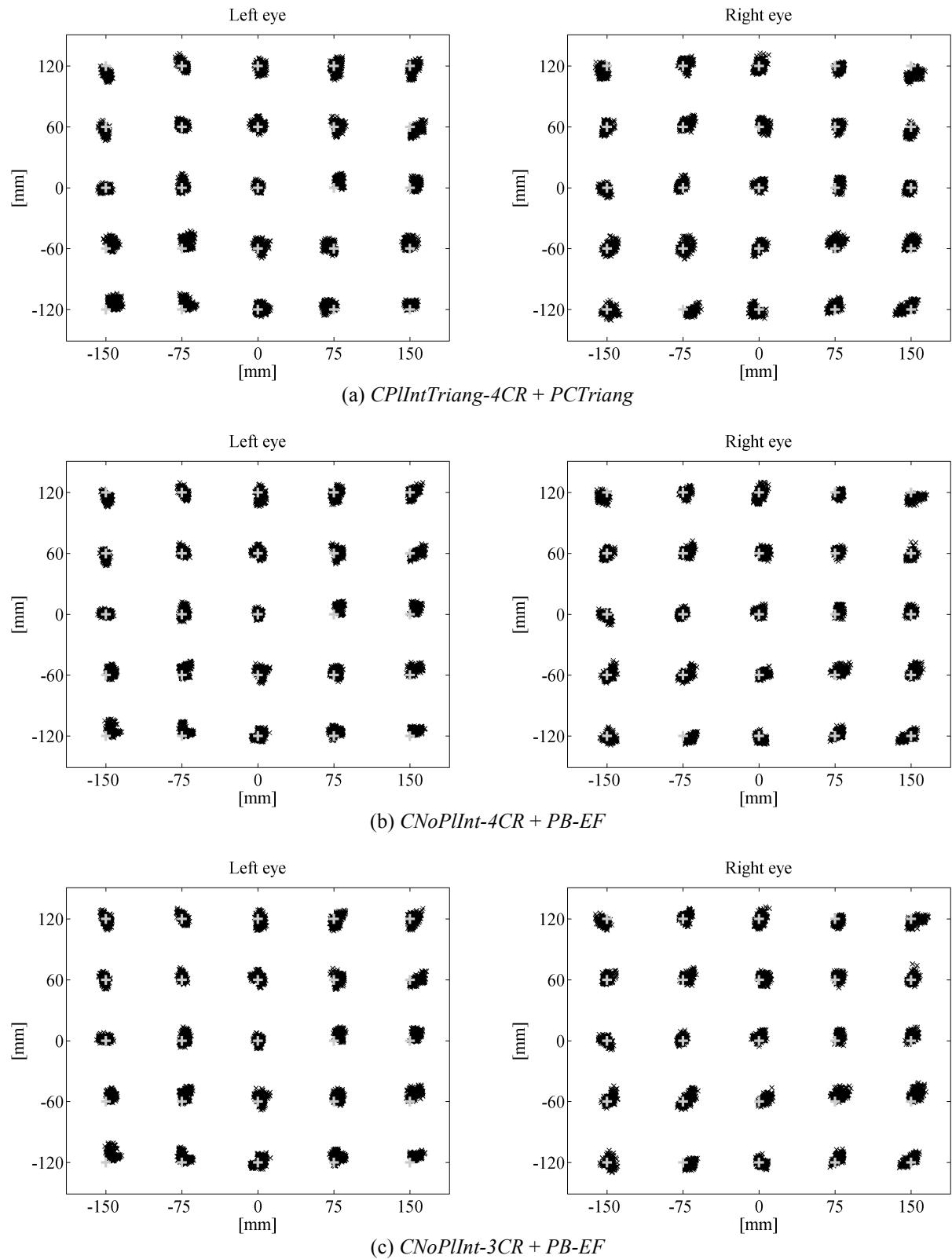


Fig. 5.9: Experimental point-of-gaze estimates for Subject 2.  
(+ : target fixation point,  $\times$  : estimated point-of-gaze)

Table 5.17: Experimental point-of-gaze estimation error for Subject 3 – *Left eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	6.46	N/A	N/A
<i>CNoPPIInt-3CR</i>	6.32	6.14	6.09
<i>CNoPPIInt-4CR</i>	5.86	5.63	5.60

Table 5.18: Experimental point-of-gaze estimation error for Subject 3 – *Right eye – RMS Bias [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	5.63	N/A	N/A
<i>CNoPPIInt-3CR</i>	5.29	5.10	5.14
<i>CNoPPIInt-4CR</i>	4.80	4.70	4.94

Table 5.19: Experimental point-of-gaze estimation error for Subject 3 – *Left eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	3.34	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.07	3.08	3.02
<i>CNoPPIInt-4CR</i>	2.86	2.87	2.83

Table 5.20: Experimental point-of-gaze estimation error for Subject 3 – *Right eye – RMS Dispersion [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	3.30	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.11	3.12	3.07
<i>CNoPPIInt-4CR</i>	2.86	2.87	2.83

Table 5.21: Experimental point-of-gaze estimation error for Subject 3 – *Left eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	7.27	N/A	N/A
<i>CNoPPIInt-3CR</i>	7.02	6.86	6.79
<i>CNoPPIInt-4CR</i>	6.51	6.30	6.26

Table 5.22: Experimental point-of-gaze estimation error for Subject 3 – *Right eye – RMS Error [mm]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	6.52	N/A	N/A
<i>CNoPPIInt-3CR</i>	6.14	5.98	5.99
<i>CNoPPIInt-4CR</i>	5.59	5.51	5.69

Table 5.23: Experimental point-of-gaze estimation error for Subject 3 – *Left eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.60	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.58	0.57	0.56
<i>CNoPPIInt-4CR</i>	0.54	0.52	0.52

Table 5.24: Experimental point-of-gaze estimation error for Subject 3 – *Right eye – RMS Error [°]*

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	0.54	N/A	N/A
<i>CNoPPIInt-3CR</i>	0.50	0.49	0.49
<i>CNoPPIInt-4CR</i>	0.46	0.45	0.47

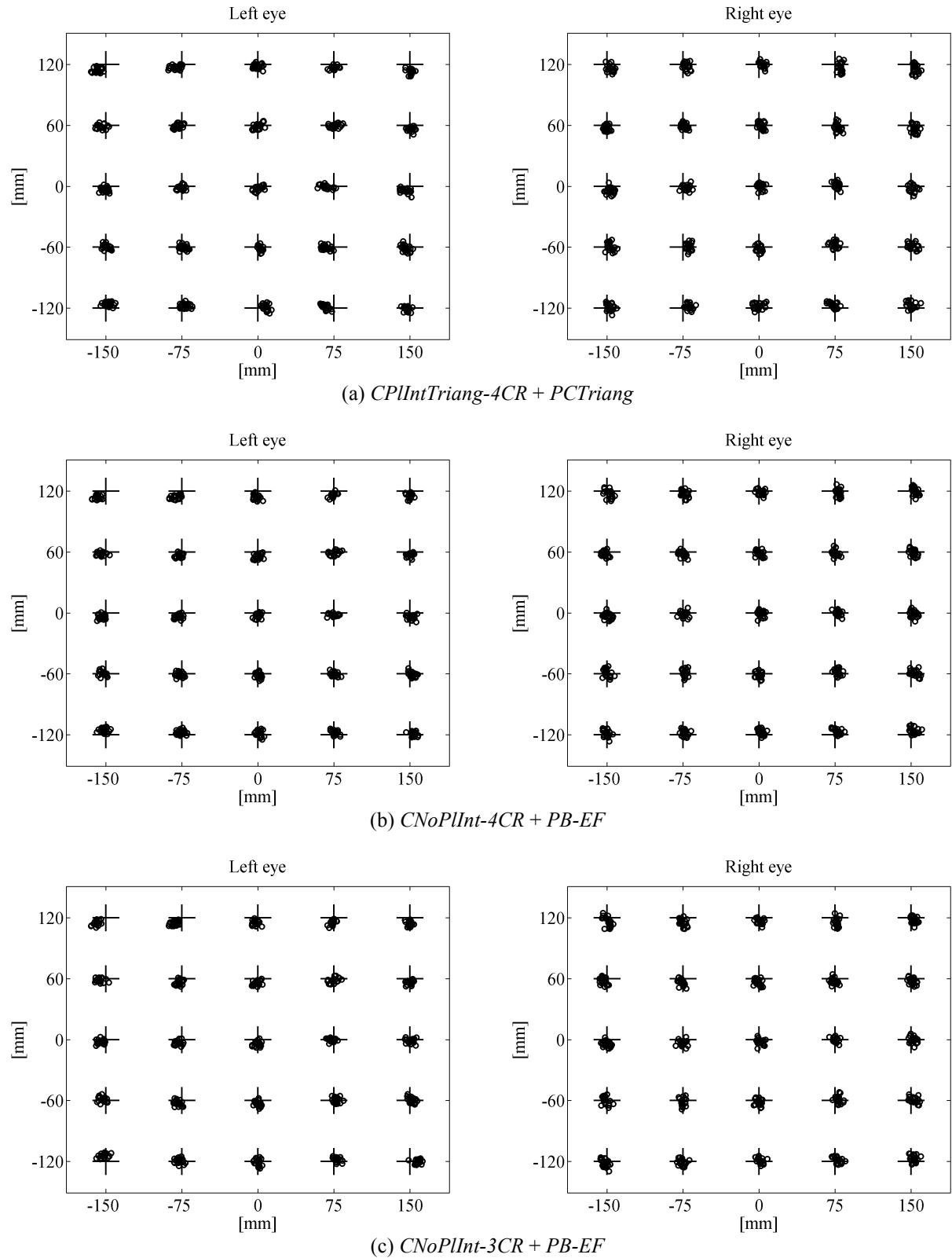


Fig. 5.10: Experimental mean point-of-gaze estimates for Subject 3.  
(+ : target fixation point, ○ : estimated point-of-gaze)

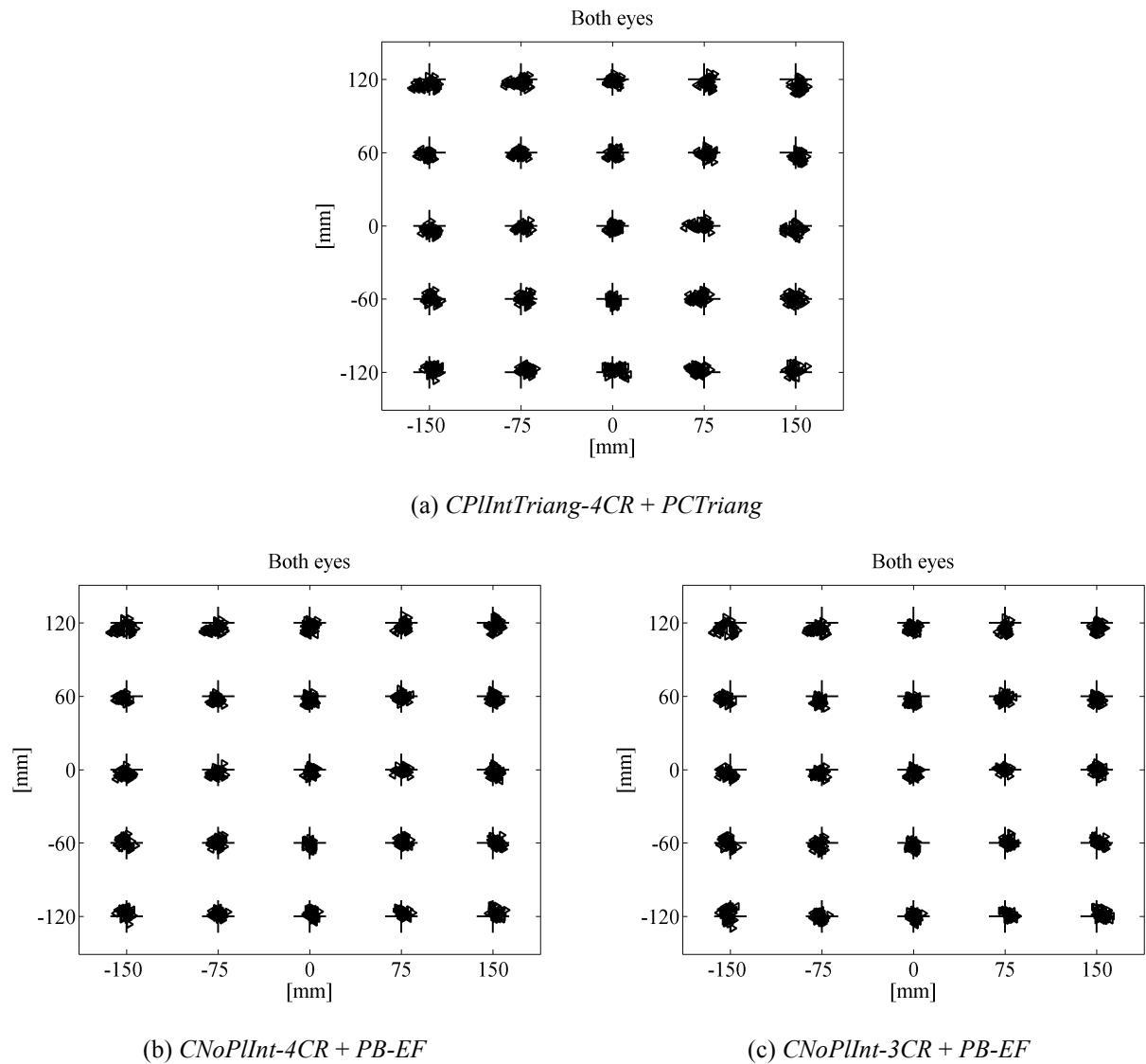


Fig. 5.11: Experimental mean point-of-gaze estimates for Subject 3.  
 (+ : target fixation point,  $\triangleleft$  : estimated point-of-gaze for the left eye,  $\triangleright$  : estimated point-of-gaze for the right eye)

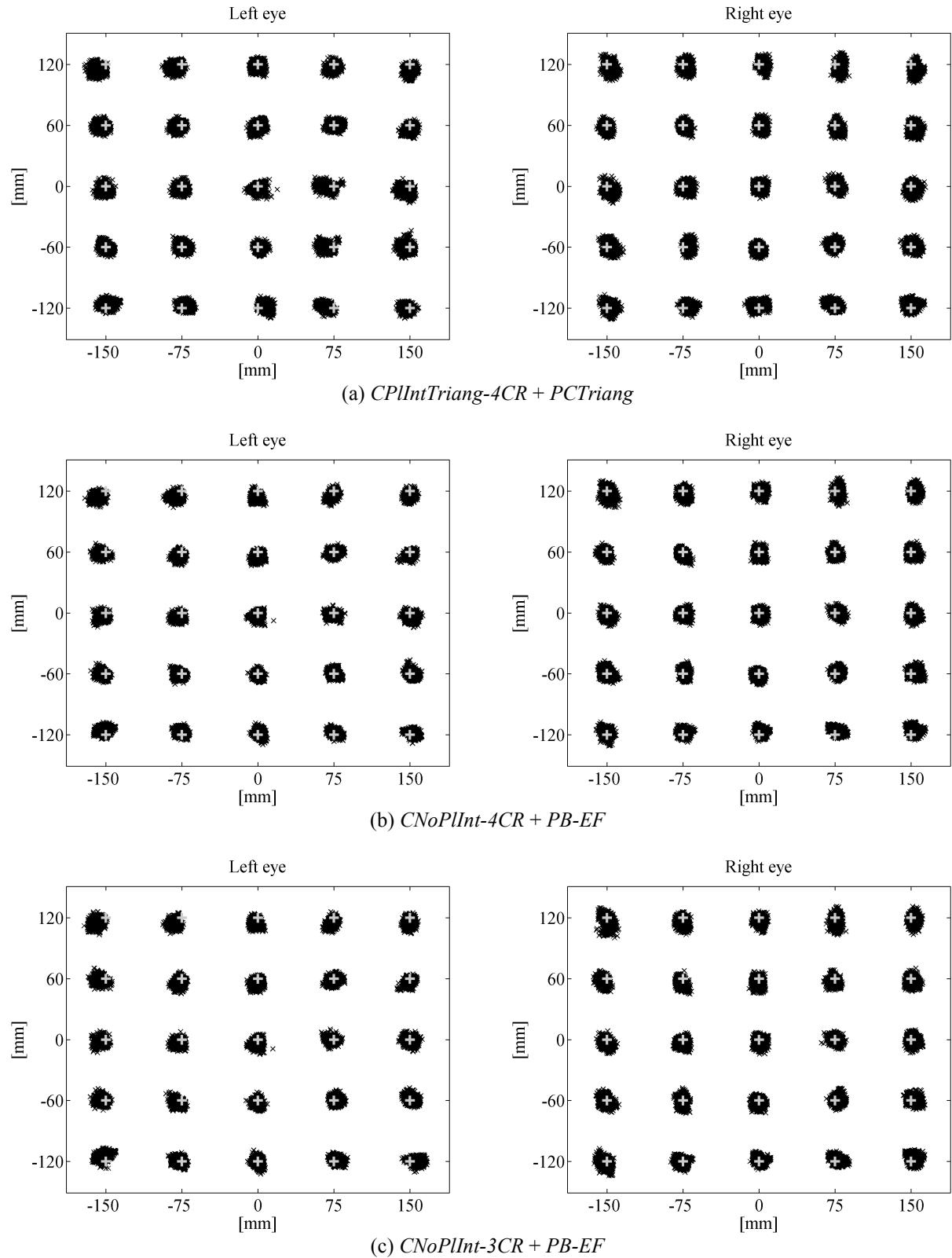


Fig. 5.12: Experimental point-of-gaze estimates for Subject 3.  
(+ : target fixation point,  $\times$  : estimated point-of-gaze)

## 5.5 Experimental results with infants

The goal of the experiments with infants was to demonstrate the ability to detect whether infants fixate a visual stimulus and record the eye movements elicited by a moving grating (optokinetic nystagmus) using the proposed point-of-gaze estimation system. In these experiments, which were carried out with the initial system setup (Section 5.2.1), the infants sat on their parent's lap with their face at about 65 cm from the screen and the point-of-gaze was estimated for both eyes simultaneously (with the *CPIIntTriang-4CR + PCTriang* method). In order to calibrate for the subject-specific angular deviation of the visual axis from the optic axis,  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , a prominent looming stimulus was presented at the center of the screen for about 3 seconds. A 500 ms tone was produced at the onset of the stimulus. After the calibration procedure, two different experiments were carried out.

In the first experiment (fixation experiment), the same looming stimulus was presented sequentially at the center of the screen and at several pre-selected random positions on the screen. The looming stimulus was presented at each position for 3 seconds and a 500 ms tone was produced at the onset of the stimulus at each new position. The recordings of the  $X$  and  $Y$ -coordinates (with respect to the WCS) of the stimulus position and the point-of-gaze estimates as a function of time, corresponding to one of the infants, are shown in Fig. 5.13. Fig. 5.13 demonstrates that it is relatively easy to determine from the recordings that the visual stimulus was seen by the infant at all seven positions.

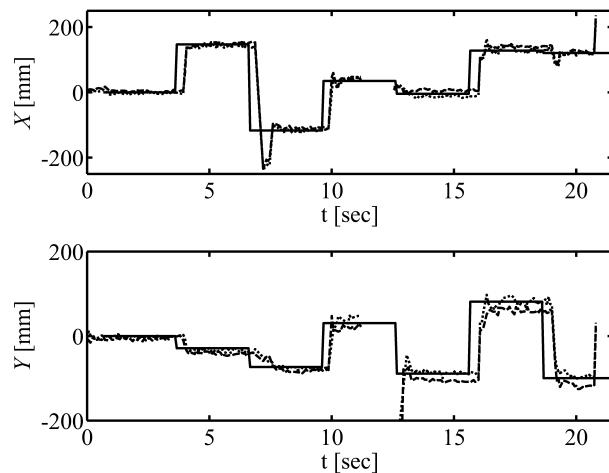


Fig. 5.13: Recording of the stimulus position (solid line) and the estimates of the infant's point-of-gaze (dotted line: right eye, dashed line: left eye) as a function of time for the fixation experiment.  
The gap in the traces corresponds to a period without recording when the infant turned the head away.

In the second experiment (optokinetic nystagmus experiment), a full-screen black and white vertical grating (square wave, 50 % duty cycle, full contrast) with a spatial frequency of 0.8 cycles/cm (equivalent to about 0.7 cycles/degree) was moving from left to right at a speed of 25 mm/second (equivalent to about 2.2 degrees/second). A recording of the horizontal eye movements elicited by this stimulus is shown in Fig. 5.14. This recording exhibits four clear segments of the slow phase (slow smooth eye movements that track the motion of the grating) of the optokinetic nystagmus. One fast phase (corrective saccade in the opposite direction of the motion) is also indicated in this figure. The infant's eye movements demonstrate that she followed the moving grating pattern. Note that since this experiment consisted of measuring eye movements rather than the absolute coordinates of the point-of-gaze, it would have been enough to measure the movements of the optic axis of the eye. Since the system can reconstruct the optic axis of the eye in 3-D space without personal calibration, the personal calibration was therefore not essential for this experiment.

The results shown in Fig. 5.13 and Fig. 5.14 demonstrate that by monitoring the point-of-gaze it is relatively easy to determine if and when infants see the visual stimuli.

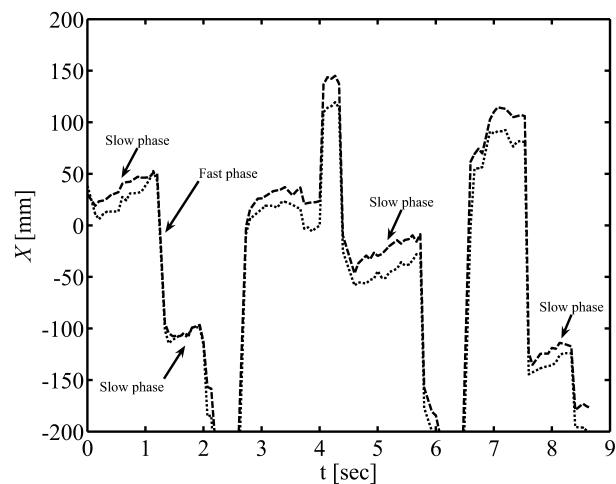


Fig. 5.14: Optokinetic nystagmus in a 7-month old infant.

The horizontal component of the infant's point-of-gaze estimates is shown (dotted line: right eye, dashed line: left eye). The slow phase eye movements track the motion of the grating and the fast phase eye movements correspond to corrective saccades in the opposite direction.

## 5.6 Summary of experimental results

The experimental results with adults show that using the proposed methodologies it is possible to estimate the point-of-gaze with an accuracy of 0.4-0.6° of visual angle after completing a single-point personal calibration procedure. Such point-of-gaze estimation accuracy is comparable to the best commercially available systems, which require multiple-point personal calibration procedures. Furthermore, to the best of my knowledge, the point-of-gaze estimation accuracy obtained with this prototype system is significantly better than that of any other working systems described in the literature that use a single-point personal calibration procedure (e.g., [80], [94]).

The experimental results with infants demonstrate the ability of a system that requires a single-point personal calibration procedure to estimate infants' point-of-gaze. The ability to monitor infants' visual scanning behavior can enable many applications that currently rely on subjective measurements performed by trained human observers and are very difficult or impossible to carry out with systems that require multiple-point personal calibration procedures. Some examples of applications with infants are the study of the development and the objective assessment of functions of the visual and oculomotor systems of preverbal infants, the objective determination of attention allocation, and the study of cognitive development.

# Chapter 6

## Summary of Contributions and Future Work

### 6.1 Summary of contributions

All the goals set for this thesis (Section 1.2) were attained and the ability to estimate the point-of-gaze using the methodologies developed in this thesis was successfully demonstrated with both adults and infants using a prototype system. The following paragraphs summarize the contributions of this thesis.

A detailed mathematical model for video-based gaze estimation systems that use the pupil and one or more corneal reflections extracted from images captured by one or more video cameras was presented. The corneal reflections (first Purkinje images, glints) are virtual images of light sources (usually infrared) that illuminate the eye. These virtual images are created by reflection on the front surface of the cornea, which acts as a convex mirror. Given the geometrical setup of a gaze estimation system (position of light sources, camera parameters), the mathematical model relates the position and orientation of the eye (or the position of the eye and the point-of-gaze) in 3-D space with the images of the pupil and corneal reflections captured by the camera(s).

The mathematical model is used not only to estimate gaze in 3-D space but also as the basis of a simulation framework to generate synthetic eye feature data. The generation of synthetic eye feature data consists of determining the pupil boundary, the pupil center and the centers of the corneal reflections (eye features) in the images captured by the camera(s), given a set of eye parameters, the system parameters (position of the light sources, camera parameters) and the 3-D positions of the center of the eyeball and the point-of-gaze. The simulation framework is an invaluable tool that supports studies of the sensitivity of the estimation of the point-of-gaze to several different types of error such as errors in the estimation of the coordinates of the eye features in each image of the eye, the deviation of the actual corneal shape from the ideal spherical shape assumed for the estimation of the point-of-gaze (corneal asphericity), and errors in the system parameters. In order to study the effect of corneal asphericity, an ellipsoidal corneal

model was developed in detail.

The point-of-gaze estimation problem using a spherical corneal model was analyzed in great detail for the full range of possible system configurations, from the simplest, that includes one camera and one light source, to the most complex, that include multiple cameras and multiple light sources. This analysis, which is probably the most complete analysis of the problem of estimating the point-of-gaze from the pupil and corneal reflections available to date, provides fundamental insight into the limitations and potential of each system configuration, including the personal calibration requirements. An earlier version of this analysis was published in June 2006 in [84], which as of August 2009 has been cited at least 30 times. Of the publications referenced in this thesis, [84] is cited in [29], [30], [61], [62], [67], [73], [83], [85]-[87], [90]-[95], [114], [115].

The analysis of the full range of system configurations shows that one camera and two light sources is the simplest system configuration that can be used to estimate the point-of-gaze in 3-D space in the presence of head movements, using only the pupil and the corneal reflections extracted from the eye images, after each subject completes a one-time personal calibration procedure in which (s)he is required to fixate at least two known target calibration points. By using two cameras and two or more light sources it is possible to reconstruct the optic axis of the eye in 3-D space, and therefore measure eye movements, without the need for personal calibration. The visual axis and the point-of-gaze in 3-D space can be estimated after completing a one-time personal calibration procedure in which the subject is required to fixate only one known target calibration point. The purpose of this single-point personal calibration procedure is to determine the angular deviation between the visual axis and the optic axis of the eye.

Because of its theoretical and practical interest, attention was focused on the point-of-gaze estimation problem using a spherical corneal model with two cameras and two or more light sources. Given the geometrical conditions set by the laws of reflection and refraction, there is more than one way to express them analytically, some of which involve certain approximations. Furthermore, for a given system of nonlinear equations there is often more than one way of solving them. Most importantly, each way of expressing the geometrical conditions analytically and each way of solving the corresponding equations can exhibit a significantly different behavior in non-ideal conditions (e.g., noise in the image coordinates of the eye features, corneal

asphericity, errors in system parameters). Consequently, a detailed simulation study of different methods to estimate the point-of-gaze using two cameras and two or more light sources was carried out. The sensitivity of the point-of-gaze estimates to the above sources of errors was evaluated for the different methods as a function of the relative positions of the cameras and light sources (system setup). The understanding of the sensitivities of the different methods and the role of the system setup is essential for a successful system implementation. No other study of this nature and magnitude is currently available in the literature.

A series of preliminary proof-of-concept experiments were carried out with adults and infants using a prototype system with two cameras and four light sources. The prototype system tolerates moderate head movements, which are only limited by the field of view and depth of field of the cameras used. One of the setups of the prototype system is based directly on the guidelines from the simulation study mentioned in the preceding paragraph.

The experimental results with adults show that using the proposed methodologies it is possible to estimate the point-of-gaze with an accuracy of 0.4-0.6° of visual angle after completing a single-point personal calibration procedure. Such point-of-gaze estimation accuracy is comparable to the best commercially available systems, which require multiple-point personal calibration procedures. Furthermore, to the best of my knowledge, the point-of-gaze estimation accuracy obtained with the prototype system is significantly better than that of any other working systems described in the literature that use a single-point personal calibration procedure.

The experimental results with infants demonstrate the ability of a system that requires a single-point personal calibration procedure to estimate infants' point-of-gaze. The ability to monitor infants' visual scanning behavior can enable many applications that currently rely on subjective measurements performed by trained human observers and are very difficult or impossible to carry out with systems that require multiple-point personal calibration procedures. Some examples of applications with infants are the study of the development and the objective assessment of functions of the visual and oculomotor systems of preverbal infants, the objective determination of attention allocation, and the study of cognitive development.

## 6.2 Future work

Future work directions can be divided into those related to the gaze estimation methodologies and those related to their applications.

### 6.2.1 Point-of-gaze estimation methodologies

The number of applications enabled by the system developed in this thesis can be expanded by using higher resolution cameras to increase the range of allowed head movements while maintaining the system accuracy, and incorporating video-based head tracking algorithms. In this way, the system would be able to determine the position and orientation of the eyes and the point-of-gaze in 3-D space, the position and orientation of the head in 3-D space (head pose), and the orientation of the eyes with respect to the head (of interest in applications in ophthalmology and neurology), while requiring a very simple personal calibration procedure.

In order to overcome the problem of corneal asphericity, it may be worth exploring the use of non-spherical (e.g., ellipsoidal, ellipsotoric, polynomial) corneal models<sup>12</sup> for the estimation of the point-of-gaze. Since non-spherical corneal models have more subject-specific parameters than a spherical model, a more elaborate personal calibration procedure will be required. In general terms, the larger number of subject-specific parameters could be estimated by having the subject fixate a large number of target calibration points. However, this goes against the requirement of minimal subject cooperation. Using a system with multiple cameras and multiple light sources it could be possible to develop a personal calibration procedure that is consistent with the minimal subject cooperation requirement. This personal calibration procedure would consist of two stages. In the first stage, which would be used to estimate the corneal model parameters, the subject would be required just to look around without necessarily knowing where (s)he is looking at (e.g., while watching a video-clip). In the second stage, the subject would be required to fixate a single target point to determine the angular deviation between the visual axis and the optic axis of the eye, similarly to the system developed in this thesis.

---

<sup>12</sup> The use of an ellipsoidal corneal model for the estimation of the point-of-gaze was considered in [79]. However, that work presents experimental results obtained with only one subject and the point-of-gaze estimation accuracy reported is 0.6°, which is similar to the accuracy obtained with the prototype system presented in this thesis.

A possible way to completely avoid the problem of corneal asphericity, and even eliminate the need for active illumination, is to reconstruct the optic axis of the eye using the image of the iris-sclera boundary (limbus) extracted from multiple cameras by extending some of the concepts presented in this thesis. This approach can also allow for the reconstruction of the optic axis of the eye in 3-D space without knowing any subject-specific eye parameter. Of course, in order to reconstruct the visual axis and estimate the point-of-gaze in 3-D space, a personal calibration procedure in which the subject has to fixate at least one target point is required. A limitation of this approach is that often the top and bottom of the iris are covered by the eyelids, resulting in reduced system robustness, mainly increased sensitivity to noise in the reconstruction of the optic axis of the eye. As higher resolution cameras become available, the sensitivity to noise will be reduced and this approach may become a widely used practical solution.

Recent work done by our research group suggests that it may be possible to replace the single-point personal calibration procedure that determines the angular deviation between the visual axis and the optic axis of the eye by an automatic calibration procedure in which the subject does not have to fixate a specific point but rather looks naturally at a computer display while, for example, a video clip is played [114], [115]. From the point of view of the subject it would be as if no personal calibration procedure were performed. This calibration procedure requires the simultaneous tracking of both eyes and relies on the fact that, for persons with a normal visual system, the visual axes of both eyes intersect at the point-of-gaze on the plane of the display. After completing the calibration procedure and determining the angular deviation between the visual axis and the optic axis for both eyes, the system would be able to estimate the point-of-gaze even when only one eye is tracked. The results of the work completed so far indicate that if the accuracy of the reconstruction of the optic axis of the eye in 3-D space is improved, the proposed automatic personal calibration procedure can become a practical solution that will represent a significant advancement of the field.

### 6.2.2 Applications

Gaze estimation systems based on the methodologies presented in this thesis can enable a large number of applications that were very difficult or impossible to carry out with previously existing technologies. Important application areas are pediatric ophthalmology (e.g., the study of the development and the assessment of functions of the visual and oculomotor systems of preverbal infants) and pediatric cognitive sciences (e.g., the determination of attention allocation,

the study of cognitive development). Work with infants, especially during the first months of life, has been limited mainly to procedures involving trained human observers, due to the lack of instruments that could measure eye movements and determine visual scanning patterns non-invasively, objectively and accurately with minimal subject cooperation. The methodologies presented in this thesis are a significant step towards changing this situation. For example, they can be used for a novel, non-contact, automated system for the objective assessment of visual function (visual acuity, contrast sensitivity, color sensitivity) of preverbal infants by analyzing the differences in visual scanning behavior between visual stimuli above and below the visual acuity / contrast sensitivity threshold, or between stimuli of different colors. The methodologies developed in this thesis can also be used for the accurate measurement of eye misalignment in cases of strabismus. This not only is important for the prescription of corrective eyeglasses (prisms) and for corrective surgery but can also have a significant role in the understanding of the causes of strabismus by monitoring the development of the oculomotor system of a large number of infants. These applications can have a significant impact on clinical research and practice in pediatric ophthalmology.

## References

- [1] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 455-470, 2002.
- [2] R. S. Allison, M. Eizenman, and B. S. K. Cheung, "Combined head and eye tracking system for dynamic testing of the vestibular system," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 11, pp. 1073-1082, Nov. 1996.
- [3] R. S. Allison, M. Eizenman, R. D. Tomlinson, J. Nedzelski, and J. A. Sharpe, "Vestibulo-ocular reflex deficits to rapid head turns following intratympanic gentamicin instillation," *Journal of Vestibular Research-Equilibrium & Orientation*, vol. 7, no. 5, pp. 369-380, Sep.-Oct. 1997.
- [4] H. R. Lee and M. Eizenman, "Vergence eye movements in strabismus patients," in *Neurobiology of Eye Movements: From Molecules to Behavior, Annals of the New York Academy of Sciences*, vol. 956, H. J. Kaminski and R. J. Leigh, Eds. New York, NY: New York Academy of Sciences, 2002, pp. 499-503.
- [5] M. S. Salman, J. A. Sharpe, M. Eizenman, L. Lillakas, T. To, C. Westall, M. J. Steinbach, and M. Dennis, "Saccades in children with spina bifida and Chiari type II malformation," *Neurology*, vol. 64, no. 12, pp. 2098-2101, Jun. 2005.
- [6] M. S. Salman, J. A. Sharpe, M. Eizenman, L. Lillakas, C. Westall, T. To, M. Dennis, and M. J. Steinbach, "Saccades in children," *Vision Research*, vol. 46, no. 8-9, pp. 1432-1439, Apr. 2006.
- [7] M. Eizenman, R. Sapir-Pichhadze, C. A. Westall, A. M. Wong, H. Lee, and Y. Morad, "Eye-movement responses to disparity vergence stimuli with artificial monocular scotomas," *Current Eye Research*, vol. 31, no. 6, pp. 471-480, Jun. 2006.
- [8] M. S. Salman, J. A. Sharpe, M. Eizenman, L. Lillakas, T. To, C. Westall, M. J. Steinbach, and M. Dennis, "Saccadic adaptation in Chiari type II malformation," *Canadian Journal of Neurological Sciences*, vol. 33, no. 4, pp. 372-378, Nov. 2006.
- [9] M. S. Salman, J. A. Sharpe, M. Eizenman, L. Lillakas, T. To, C. Westall, M. Dennis, and M. J. Steinbach, "Saccadic adaptation in children," *Journal of Child Neurology*, vol. 21, no. 12, pp. 1025-1031, Dec. 2006.
- [10] M. Eizenman, L. H. Yu, L. Grupp, E. Eizenman, M. Ellenbogen, M. Gemar, and R. D. Levitan, "A naturalistic visual scanning approach to assess selective attention in major depressive disorder," *Psychiatry Research*, vol. 118, no. 2, pp. 117-128, May 2003.
- [11] C. Karatekin and R. F. Asarnow, "Exploratory eye movements to pictures in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder (ADHD)," *Journal of Abnormal Child Psychology*, vol. 27, no. 1, pp. 35-49, Feb. 1999.
- [12] M. De Luca, E. Di Pace, A. Judica, D. Spinelli, and P. Zoccolotti, "Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia," *Neuropsychologia*, vol. 37, no. 12, pp. 1407-1420, Nov. 1999.
- [13] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, pp. 372-422, 1998.

- [14] M. Eizenman, T. Jares, and A. Smiley, "A new methodology for the analysis of eye movements and visual scanning in drivers," in *Proc. of the 31<sup>st</sup> Annual Conference on Ergonomics and Safety, Hall, Quebec, Canada*, 1999.
- [15] A. Zelinsky, "Visual human-machine interaction," in *Advanced Topics in Artificial Intelligence, Lecture Notes in Artificial Intelligence 1747*, N. Foo, Ed. Berlin: Springer-Verlag, 1999, pp. 440-452.
- [16] J-C. Popieul, P. Simon, R. Leroux, and J-C. Angué, "Automatic processing of a car driver eye scanning movements on a simulated highway driving context," in *Proc. of the IEEE Intelligent Vehicles Symposium 2000*, 2000, pp. 13-18.
- [17] J. L. Harbluk, Y. I. Noy, P. L. Trbovich, and M. Eizenman "An on-road assessment of cognitive distraction: impacts on drivers' visual behavior and braking performance," *Accident Analysis and Prevention*, vol. 39, no. 2, pp. 372-379, Mar. 2007.
- [18] D. Cleveland, "Unobtrusive eyelid closure and visual point of regard measurement system," in *Proc. of the Technical Conference on Ocular Measures of Driver Alertness, sponsored by The Federal Highway Administration – Office of Motor Carrier and Highway Safety and The National Highway Traffic Safety Administration – Office of Vehicle Safety Research, Herndon, VA, USA*, 1999, pp. 57-74.
- [19] P. A. Wetzel, G. Krueger-Anderson, C. Poprik, and P. Bascom, "An eye tracking system for analysis of pilots' scan paths," United States Air Force Armstrong Laboratory, Tech. Rep. AL/HR-TR-1996-0145, Apr. 1997.
- [20] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631-645, Oct. 1999.
- [21] G. Loshe, "Consumer eye movement patterns of Yellow Pages advertising," *Journal of Advertising*, vol. 26, pp. 61-73, 1997.
- [22] R. Sharma, V. I. Pavlović, and T. S. Huang, "Toward multimodal human-computer interface," *Proc. of the IEEE*, vol. 86, no. 5, pp. 853-869, May 1998.
- [23] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1527-1534, Nov./Dec. 1989.
- [24] L. A. Frey, K. P. White, and T. E. Hutchinson, "Eye-gaze word processing," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 4, pp. 944-950, Jul./Aug. 1990.
- [25] P. Majaranta and K.-J. Räihä, "Twenty years of eye typing: System and design issues," in *Proc. of the Second Symposium on Eye Tracking Research and Applications (ETRA 2002), New Orleans, LA, USA*, Mar. 2002, pp. 15-22.
- [26] O. Tuisku, P. Majaranta, P. Isokoski, and K.-J. Räihä, "Now Dasher! Dash away! Longitudinal study of fast text entry by eye gaze," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008), Savannah, GA, USA*, Mar. 2008, pp. 19-26.

- [27] L. R. Young and D. Sheena, "Methods and designs - Survey of eye movement recording methods," *Behavior Research Methods & Instrumentation*, vol. 7, no. 5, pp. 397-429, 1975.
- [28] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer Vision and Image Understanding*, vol. 98, pp. 4-24, 2005.
- [29] G. Daunys, B. K. Ersbøll, M. Böhme, O. Stepankova, A. Villanueva, E. Barth, M. Vester-Christensen, T. Delbruck, D. Dervinis, D. Droege, J. Fejt, M. Fejtová, D. W. Hansen, L. K. Hansen, D. Leimberg, A. Meyer, T. Martinetz, N. Ramanauskas, and V. Vysniauskas, "D5.2 Report on new approaches to eye tracking," Communication by Gaze Interaction (COGAIN), Tech. Rep. IST-2003-511598: Deliverable 5.2, Oct. 2006. (Available at <http://www.cogain.org/results/reports/COGAIN-D5.2.pdf>)
- [30] M. Böhme, A. Meyer, T. Martinetz, and E. Barth, "Remote eye tracking: State of the art and directions for future development," in *Proc. of the 2<sup>nd</sup> Conference on Communication by Gaze Interaction - COGAIN 2006: Gazing into the Future, Turin, Italy*, Sep. 2006, pp. 1-5.
- [31] (2009, Feb.). [Online]. Available: <http://commons.wikimedia.org/wiki/File:Eyesection.svg>
- [32] S. Baluja and D. Pomerleau, "Non-intrusive gaze tracking using artificial neural networks," School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-CS-94-102, Jan. 1994.
- [33] K.-H. Tan, D. J. Kriegman, and N. Ahuja, "Appearance-based eye gaze estimation," in *Proc. of the Sixth IEEE Workshop on Applications of Computer Vision (WACV'02)*, 2002, pp. 191-195.
- [34] R. Newman, Y. Matsumoto, S. Rougeaux, and A. Zelinsky, "Real-time stereo tracking for head pose and gaze estimation," in *Proc. of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 122-128.
- [35] Y. Matsumoto and A. Zelinsky, "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement," in *Proc. of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 499-504.
- [36] J. Heinzmann and A. Zelinsky, "3-D facial pose and gaze point estimation using a robust real-time tracking paradigm," in *Proc. of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 142-147.
- [37] H. Yamazoe, A. Utsumi, T. Yonezawa, and S. Abe, "Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008), Savannah, GA, USA*, Mar. 2008, pp. 245-250.
- [38] (2009, Jan.). faceLAB 4.6. Seeing Machines, Canberra, Australia. [Online]. Available: <http://www.seeingmachines.com/facelab.htm>
- [39] (2009, Jan.). Smart Eye Pro 5.0 (technical white paper). Smart Eye AB, Göteborg, Sweden. [Online]. Available: [http://www.smarteye.se/files/SmartEyePro50\\_whitepaper.pdf](http://www.smarteye.se/files/SmartEyePro50_whitepaper.pdf)
- [40] J.-G. Wang and E. Sung, "Gaze determination via images of irises," *Image and Vision Computing*, vol. 19, no. 12, pp. 891-911, Oct. 2001.

- [41] J.-G. Wang and E. Sung, "Study on eye gaze estimation," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 32, no. 3, pp. 332-350, Jun. 2002.
- [42] J.-G. Wang, E. Sung, and R. Venkateswarlu, "Estimating the eye gaze from one eye," *Computer Vision and Image Understanding*, vol. 98, pp. 83-103, 2005.
- [43] A. Cowey, "The basis of a method of perimetry with monkeys," *The Quarterly Journal of Experimental Psychology*, vol. 15, part 2, pp. 81-90, 1963.
- [44] (2009, Jan.). LC Technologies, Inc., Fairfax, VA, USA. [Online]. Available: <http://www.eyegaze.com>
- [45] (2009, Jan.). Tobii Technology AB, Danderyd, Sweden. [Online]. Available: <http://www.tobii.com>
- [46] (2009, Jan.). Applied Science Laboratories (ASL), Bedford, MA, USA. [Online]. Available: <http://www.a-s-l.com>
- [47] (2009, Jan.). SensoMotoric Instruments GmbH (SMI), Teltow/Berlin, Germany. [Online]. Available: <http://www.smivision.com>
- [48] K. Nguyen, C. Wagner, D. Koons, and M. Flickner, "Differences in the infrared bright pupil response of human eyes," in *Proc. of the Second Symposium on Eye Tracking Research and Applications (ETRA 2002)*, New Orleans, LA, USA, Mar. 2002, pp. 133-138.
- [49] J. Merchant, R. Morrisette, and J. L. Porterfield, "Remote measurement of eye direction allowing subject motion over one cubic foot of space," *IEEE Transactions on Biomedical Engineering*, vol. BME-21, no. 4, pp. 309-317, Jul. 1974.
- [50] J. San Agustin, A. Villanueva, and R. Cabeza, "Pupil brightness variation as a function of gaze direction," in *Proc. of the Fourth Symposium on Eye Tracking Research and Applications (ETRA 2006)*, San Diego, CA, USA, Mar. 2006, p. 49.
- [51] R. Grace, V. E. Byrne, D. M. Bierman, J. M. Legrand, D. Gricourt, R. K. Davis, J. J. Staszewski, and B. Carnahan, "A drowsy driver detection system for heavy vehicles," in *Proc. of the Technical Conference on Ocular Measures of Driver Alertness, sponsored by The Federal Highway Administration – Office of Motor Carrier and Highway Safety and The National Highway Traffic Safety Administration – Office of Vehicle Safety Research, Herndon, VA, USA*, 1999, pp. 75-86.
- [52] Y. Ebisawa, "Improved video-based eye-gaze detection method," in *Proc. of the 10<sup>th</sup> Anniversary IEEE Instrumentation and Measurement Technology Conference*, 1994, vol. 2, pp. 963-966.
- [53] M. Ohtani and Y. Ebisawa, "Eye-gaze detection based on the pupil detection technique using two light sources and the image difference method," in *Proc. of the 17<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1995, vol. 2, pp. 1623-1624.
- [54] Y. Ebisawa, "Unconstrained pupil detection technique using two light sources and the image difference method," in *Visualization and Intelligent Design in Engineering and Architecture II*, S. Hernández and C. A. Brebbia, Eds. Southampton, UK: Computational Mechanics Publications, 1995, pp. 79-89.

- [55] Y. Ebisawa, "Improved video-based eye-gaze detection method," *IEEE Transactions on Instrumentation and Measurement*, vol. 47, no. 4, pp. 948-955, Aug. 1998.
- [56] A. Tomono, M. Iida, and Y. Kobayashi, "A TV camera system which extracts feature points for non-contact eye movement detection," in *Optics, Illumination, and Image Sensing for Machine Vision IV, Proc. SPIE*, vol. 1194, D. J. Svetkoff, Ed. Bellingham, WA, USA: SPIE-The International Society for Optical Engineering, 1989, pp. 2-12.
- [57] D. H. Yoo, J. H. Kim, B. R. Lee, and M. J. Chung, "Non-contact eye gaze tracking system by mapping of corneal reflections," in *Proc. of the 5th IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 94-99.
- [58] D. H. Yoo and M. J. Chung, "A novel non-intrusive eye gaze estimation using cross-ratio under large head motion," *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 25-51, 2005.
- [59] F. L. Coutinho and C. H. Morimoto, "Free head motion eye gaze tracking using a single camera and multiple light sources," in *Proc. of the XIX Brazilian Symposium on Computer Graphics and Image Processing*, 2006, pp. 171-178.
- [60] J. J. Kang, **E. D. Guestrin**, W. J. MacLean, and M. Eizenman, "Simplifying the cross-ratios method of point-of-gaze estimation," presented at the *30<sup>th</sup> Canadian Medical and Biological Engineering Conference (CMBEC30)*, Toronto, Ontario, Canada, Jun. 2007 (in *Proc. FICCDAT 2007*, Toronto, Ontario, Canada, Jun. 2007).
- [61] **E. D. Guestrin**, M. Eizenman, J. J. Kang, and E. Eizenman, "Analysis of subject-dependent point-of-gaze estimation bias in the cross-ratios method," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008)*, Savannah, GA, USA, Mar. 2008, pp. 237-244.
- [62] J. J. Kang, M. Eizenman, **E. D. Guestrin**, and E. Eizenman, "Investigation of the cross-ratios method for point-of-gaze estimation," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2293-2302, Sep. 2008.
- [63] J. Merchant, "Fixation point measurement by the oculometer technique," *Optical Engineering*, vol. 13, no. 4, pp. 339-342, Jul./Aug. 1974.
- [64] A. G. Gale, "A note on the remote oculometer technique for recording eye movements," *Vision Research*, vol. 22, no. 1, pp. 201-202, 1982.
- [65] K. P. White, T. E. Hutchinson, and J. N. Carley, "Spatially dynamic calibration of an eye-tracking system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 4, pp. 1162-1168, Jul./Aug. 1993.
- [66] C. H. Morimoto, D. Koons, A. Amir, M. Flickner, and S. Zhai, "Keeping an eye for HCI," in *Proc. of the XII Brazilian Symposium on Computer Graphics and Image Processing*, 1999, pp. 171-176.
- [67] J. J. Cerrolaza, A. Villanueva, and R. Cabeza, "Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008)*, Savannah, GA, USA, Mar. 2008, pp. 259-266.

- [68] X. L. C. Brolly and J. B. Mulligan, "Implicit calibration of a remote gaze tracker," in *Proc. of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, 2004, vol. 8, p. 134.
- [69] A. Sugioka, Y. Ebisawa, and M. Ohtani, "Noncontact video-based eye-gaze detection method allowing large head displacements," in *Proc. of the 18<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1996, vol. 2, pp. 526-528.
- [70] Y. Ebisawa, M. Ohtani, A. Sugioka, and S. Esaki, "Single mirror tracking system for free-head video-based eye-gaze detection method," in *Proc. of the 19<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1997, vol. 4, pp. 1448-1451.
- [71] K. Talmi and J. Liu, "Eye and gaze tracking for visually controlled interactive stereoscopic displays," *Signal Processing: Image Communication*, vol. 14, pp. 799-810, 1999.
- [72] J. Liu, "Determination of the point of fixation in a head-fixed coordinate system," in *Proc. of the 14<sup>th</sup> International Conference on Pattern Recognition*, 1998, vol. 1, pp. 501-504.
- [73] Z. Zhu and Q. Ji, "Novel eye gaze tracking techniques under natural head movement," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2246-2260, Dec. 2007.
- [74] Y. Ebisawa, M. Ohtani, and A. Sugioka, "Proposal of a zoom and focus control method using and ultrasonic distance-meter for video-based eye-gaze detection under free-head conditions," in *Proc. of the 18<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1996, vol. 2, pp. 523-525.
- [75] S. Pastoor, J. Liu, and S. Renault, "An experimental multimedia system allowing 3-D visualization and eye-controlled interaction without user-worn devices," *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 41-52, Mar. 1999.
- [76] S.-W. Shih, Y.-T. Wu, and J. Liu, "A calibration-free gaze tracking technique," in *Proc. of the 15<sup>th</sup> International Conference on Pattern Recognition*, Sep. 2000, vol. 4, pp. 201-204.
- [77] T. Ohno, N. Mukawa, and A. Yoshikawa, "FreeGaze: A gaze tracking system for everyday gaze interaction," in *Proc. of the Second Symposium on Eye Tracking Research and Applications (ETRA 2002), New Orleans, LA, USA*, Mar. 2002, pp. 125-132.
- [78] **E. D. Guestrin**, "A novel head-free point-of-gaze estimation system," M.A.Sc. Thesis, Dept. of Electrical and Computer Engineering, University of Toronto, Canada, 2003.
- [79] D. Beymer and M. Flickner, "Eye gaze tracking using an active stereo head," in *Proc. of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)*, Jun. 2003, vol. 2, pp. 451-458.
- [80] S.-W. Shih and J. Liu, "A novel approach to 3-D gaze tracking using stereo cameras," *IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, vol. 34, no. 1, pp. 234-245, Feb. 2004.
- [81] T. Ohno and N. Mukawa, "A free-head, simple calibration, gaze tracking system that enables gaze-based interaction," in *Proc. of the Third Symposium on Eye Tracking Research and Applications (ETRA 2004), San Antonio, TX, USA*, Mar. 2004, pp. 115-122.

- [82] T. Ohno, K. Hara, and H. Inagaki, "Simple-to-calibrate gaze tracking method," in *Passive Eye Monitoring Algorithms, Applications and Experiments*, R. I. Hammoud, Ed. Berlin: Springer, 2008, pp. 111-131.
- [83] **E. D. Guestrin** and M. Eizenman, "Remote point-of-gaze estimation with free head movements," in *Proc. of the 3<sup>rd</sup> European Medical and Biological Engineering Conference (EMBEC'05), Prague, Czech Republic*, Nov. 2005, pp. 3746-3751.
- [84] **E. D. Guestrin** and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1124-1133, Jun. 2006. (Erratum published in the *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 8, p. 1728, Aug. 2006.)
- [85] **E. D. Guestrin** and M. Eizenman, "Remote point-of-gaze estimation for studies of selective attention and mood disorders," presented at the *30<sup>th</sup> Canadian Medical and Biological Engineering Conference (CMBEC30)*, Toronto, Ontario, Canada, Jun. 2007 (in *Proc. FICCDAT 2007*, Toronto, Ontario, Canada, Jun. 2007).
- [86] **E. D. Guestrin** and M. Eizenman, "Remote point-of-gaze estimation with free head movements requiring a single-point calibration," in *Proc. of the 29<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2007), Lyon, France*, Aug. 2007, pp. 4556-4560.
- [87] **E. D. Guestrin** and M. Eizenman, "Remote point-of-gaze estimation requiring a single-point calibration for applications with infants," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008), Savannah, GA, USA*, Mar. 2008, pp. 267-274.
- [88] A. Meyer, M. Böhme, T. Martinetz, and E. Barth, "A single-camera remote eye tracker," in *Proc. of the Tutorial and Research Workshop on Perception and Interactive Technologies (PIT 2006), Kloster Irsee, Germany*, Jun., 2006 / *Lecture Notes in Artificial Intelligence*, vol. 4021, E. André et. al., Eds. Berlin: Springer-Verlag, 2006, pp. 208-211.
- [89] C. Hennessey, B. Noureddin, and P. Lawrence, "A single camera eye-gaze tracking system with free head motion," in *Proc. of the Fourth Symposium on Eye Tracking Research and Applications (ETRA 2006), San Diego, CA, USA*, Mar. 2006, pp. 87-94.
- [90] A. Villanueva and R. Cabeza, "Models for gaze tracking systems," *EURASIP Journal on Image and Video Processing*, vol. 2007, article ID 23570, 2007.
- [91] A. Villanueva, J. J. Cerrolaza, and R. Cabeza, "Geometry issues of a gaze tracking system," in *Universal Access in Human-Computer Interaction, Lecture Notes in Computer Science (LNCS)*, vol. 4555, C. Stephanidis, Ed. Berlin: Springer-Verlag, 2007, pp. 1006-1015.
- [92] A. Villanueva and R. Cabeza, "A novel gaze estimation system with one calibration point," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 38, no. 4, pp. 1123-1138, Aug. 2008.
- [93] A. Villanueva and R. Cabeza, "Evaluation of corneal refraction in a model of a gaze tracking system," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 12, pp. 2812-2822, Dec. 2008.
- [94] T. Nagamatsu, J. Kamahara, T. Iko, and N. Tanaka, "One-point calibration gaze tracking based on eyeball kinematics using stereo cameras," in *Proc. of the Fifth Symposium on*

- Eye Tracking Research and Applications (ETRA 2008), Savannah, GA, USA, Mar. 2008,* pp. 95-98.
- [95] J. Chen, Y. Tong, W. Gray, and Q. Ji, "A robust 3D eye gaze tracking system using noise reduction," in *Proc. of the Fifth Symposium on Eye Tracking Research and Applications (ETRA 2008), Savannah, GA, USA, Mar. 2008*, pp. 189-196.
  - [96] R. H. S. Carpenter, *Movements of the Eyes*. London, UK: Pion, 1977.
  - [97] M. Alpern, "Specification of the direction of regard," in *The Eye*, 2<sup>nd</sup> ed., vol. 3, H. Davson, Ed. London, UK: Academic Press Inc. (London) Ltd, 1969.
  - [98] A. M. Slater and J. M. Findlay, "The measurement of fixation position in the newborn baby," *Journal of Experimental Child Psychology*, vol. 14, pp. 349-364, 1972.
  - [99] A. G. Bennett and R. B. Rabbetts, *Clinical Visual Optics*. London, UK: Butterworths, 1984, pp. 216-217.
  - [100] R. B. Mandell and R. St. Helen, "Mathematical model of the corneal contour," *British Journal of Physiological Optics*, vol. 26, pp. 183-197, 1971.
  - [101] A. I. Tew, "Simulation results for an innovative point-of-regard sensor using neural networks," *Neural Computing & Applications*, vol. 5, no. 4, pp. 230-237, 1997.
  - [102] H. Burek and W. A. Douthwaite, "Mathematical models of the general corneal surface," *Ophthalmic and Physiological Optics*, vol. 13, no. 1, pp. 68-72, Jan. 1993.
  - [103] J. Heikkilä and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1997, vol. 1, pp. 1106-1112.
  - [104] J.-Y. Bouguet. (2009, Apr.) Camera Calibration Toolbox for MATLAB. California Institute of Technology, Department of Electrical Engineering, Pasadena, CA. [Online]. Available: [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc)
  - [105] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. of the International Conference on Computer Vision (ICCV'99), Corfu, Greece, Sep. 1999*, pp. 666-673.
  - [106] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.
  - [107] Z. Zhang, "A flexible new technique for camera calibration," Microsoft Research, Microsoft Corp., Redmond, WA, Tech. Rep. MSR-TR-98-71, Aug. 2008. (Available at <http://research.microsoft.com/en-us/um/people/zhang/Papers/TR98-71.pdf>)
  - [108] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2001.
  - [109] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ: Prentice Hall, 1998.
  - [110] M. C. Corbett, E. S. Rosen, and D. P. S. O'Brart, *Corneal Topography: Principles and Applications*. London, UK: BMJ Books, 1999, p. 6.

- [111] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C – The Art of Scientific Computing, 2<sup>nd</sup> ed.* New York, NY: Cambridge University Press, 1992.
- [112] A. L. Yarbus, *Eye Movements and Vision*. New York: Plenum Press, 1967.
- [113] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, “The role of fixational eye movements in visual perception,” *Nature Reviews Neuroscience*, vol. 5, pp. 229-240, 2004.
- [114] D. Model, **E. D. Guestrin**, and M. Eizenman, “An automatic calibration procedure for remote eye-gaze tracking systems,” in *Proc. of the 31<sup>st</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009), Minneapolis, MN, USA*, Sep. 2009, pp. 4751-4754.
- [115] M. Eizenman, D. Model, and **E. D. Guestrin**, “Covert monitoring of the point-of-gaze,” in *Proc. of the 2009 IEEE Toronto International Conference - Science and Technology for Humanity (IEEE TIC-STH 2009) - Symposium on Biomedical Engineering, Toronto, ON, Canada*, Sep. 2009, pp. 551-556.
- [116] M. Alpern, “Kinematics of the eye,” in *The Eye*, 2<sup>nd</sup> ed., vol. 3, H. Davson, Ed. London, UK: Academic Press Inc. (London) Ltd, 1969.
- [117] L. Ferman, H. Collewijn, and A. V. Van den Berg, “A direct test of Listing’s law – I. Human ocular torsion measured in static tertiary positions,” *Vision Research*, vol. 27, no. 6, pp. 929-938, 1987.
- [118] L. Ferman, H. Collewijn, and A. V. Van den Berg, “A direct test of Listing’s law – II. Human ocular torsion measured under dynamic conditions,” *Vision Research*, vol. 27, no. 6, pp. 939-951, 1987.
- [119] S. Belongie. (2009, Apr.). Rodrigues’ rotation formula. *MathWorld–A Wolfram Web Resource*, created by Eric W. Weisstein. [Online]. Available: <http://mathworld.wolfram.com/RodriguesRotationFormula.html>
- [120] E. W. Weisstein. (2009, Apr.). Radius of curvature. *MathWorld–A Wolfram Web Resource*. [Online]. Available: <http://mathworld.wolfram.com/RadiusofCurvature.html>

## Appendix A

### Relation Between the ECS and the WCS

The relation between the Eye Coordinate System (ECS) and the World Coordinate System (WCS) can be described by the following sequence of elementary transformations:

- 1) Translate the WCS  $XYZ$  to obtain a new coordinate system  $X'YZ'$  whose origin is located at the center of the eyeball,  $\mathbf{d}$ :

$$\underbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}_{\text{WCS}} = \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \mathbf{d} . \quad (\text{A.1})$$

- 2) Rotate the  $X'YZ'$  coordinate system  $180^\circ$  about the  $Y'$ -axis to obtain a new coordinate system  $X''Y''Z''$  whose  $Z''$ -axis points towards the  $XY$ -plane of the WCS (typically, the scene plane):

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{\mathbf{R}_{\text{flip, eye}}} \begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} . \quad (\text{A.2})$$

- 3) Rotate the  $X''Y''Z''$  coordinate system about the  $Y''$ -axis by the pan angle  $\theta_{\text{eye}}$  as shown in Fig. A.1 ( $\theta_{\text{eye}} > 0$  for rotations to the right of the subject) to obtain a new coordinate system  $X'''Y'''Z'''$ :

$$\begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos \theta_{\text{eye}} & 0 & -\sin \theta_{\text{eye}} \\ 0 & 1 & 0 \\ \sin \theta_{\text{eye}} & 0 & \cos \theta_{\text{eye}} \end{bmatrix}}_{\mathbf{R}_{\theta, \text{eye}}} \begin{bmatrix} X''' \\ Y''' \\ Z''' \end{bmatrix} . \quad (\text{A.3})$$

Note that after this transformation, the  $X'''Z'''$ -plane remains horizontal / the  $Y'''$ -axis remains vertical.

4) Rotate the  $X'''Y'''Z'''$  coordinate system about the  $X'''$ -axis by the tilt angle  $\varphi_{\text{eye}}$  as shown in Fig. A.2 ( $\varphi_{\text{eye}} > 0$  for upward rotations) to obtain a new coordinate system  $X^IVY^IVZ^IV$  whose  $Z^IV$ -axis is coincident with the optic axis of the eye:

$$\begin{bmatrix} X''' \\ Y''' \\ Z''' \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_{\text{eye}} & \sin \varphi_{\text{eye}} \\ 0 & -\sin \varphi_{\text{eye}} & \cos \varphi_{\text{eye}} \end{bmatrix}}_{\mathbf{R}_{\varphi, \text{eye}}} \begin{bmatrix} X^IV \\ Y^IV \\ Z^IV \end{bmatrix}. \quad (\text{A.4})$$

Note that after this transformation, the  $Y^IVZ^IV$ -plane remains vertical / the  $X^IV$ -axis remains horizontal.

5) Rotate the  $X^IVY^IVZ^IV$  coordinate system about the  $Z^IV$ -axis by the torsion angle  $\kappa_{\text{eye}}$  as shown in Fig. A.3 ( $\kappa_{\text{eye}} > 0$  for counterclockwise rotations) to finally obtain the ECS  $x_{\text{eye}}y_{\text{eye}}z_{\text{eye}}$ :

$$\begin{bmatrix} X^IV \\ Y^IV \\ Z^IV \end{bmatrix} = \underbrace{\begin{bmatrix} \cos \kappa_{\text{eye}} & -\sin \kappa_{\text{eye}} & 0 \\ \sin \kappa_{\text{eye}} & \cos \kappa_{\text{eye}} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_{\kappa, \text{eye}}} \underbrace{\begin{bmatrix} x_{\text{eye}} \\ y_{\text{eye}} \\ z_{\text{eye}} \end{bmatrix}}_{\text{ECS}}. \quad (\text{A.5})$$

Combining (A.1)-(A.5), (2.24) is obtained with

$$\mathbf{R}_{\text{eye}} \equiv \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}) = \mathbf{R}_{\text{flip}, \text{eye}} \mathbf{R}_{\theta, \text{eye}} \mathbf{R}_{\varphi, \text{eye}} \mathbf{R}_{\kappa, \text{eye}}. \quad (\text{A.6})$$

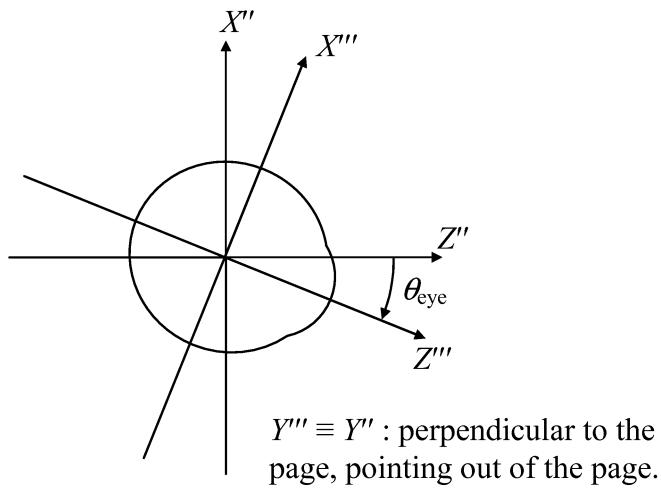


Fig. A.1: Rotation of the  $X''Y''Z''$  coordinate system about the  $Y''$ -axis by the pan angle  $\theta_{\text{eye}}$ .

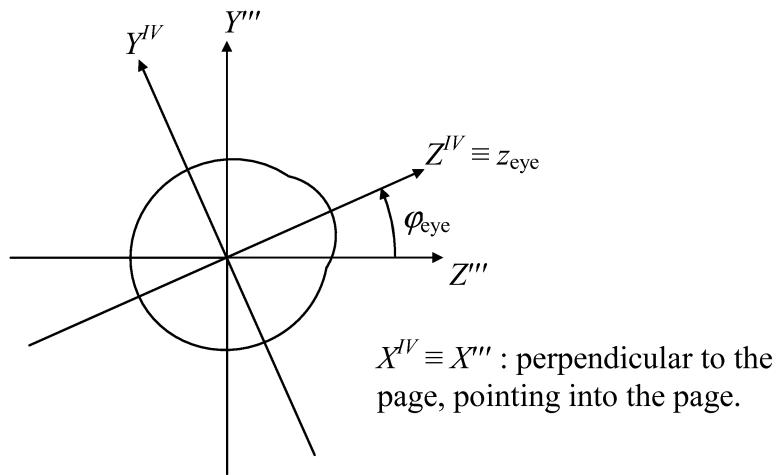


Fig. A.2: Rotation of the  $X''Y''Z''$  coordinate system about the  $X''$ -axis by the tilt angle  $\varphi_{\text{eye}}$ .

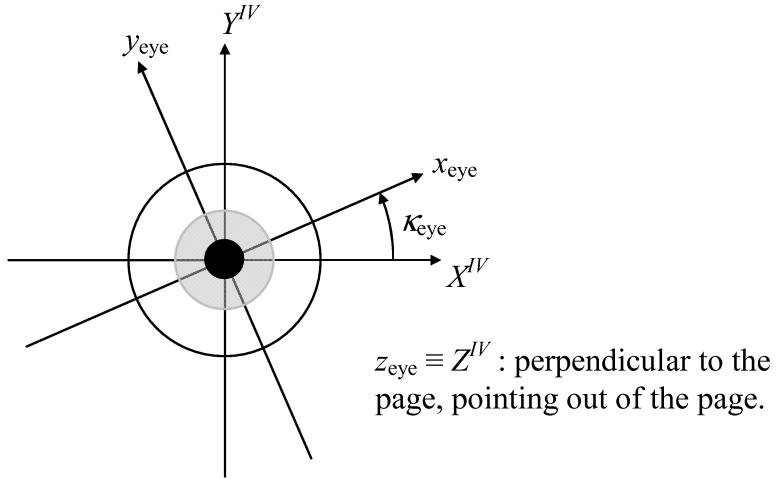


Fig. A.3: Rotation of the  $X^{IV}Y^{IV}Z^{IV}$  coordinate system about the  $Z^{IV}$ -axis by the torsion angle  $\kappa_{\text{eye}}$ .

## A.1 Notes on the torsion of the eye

The torsion angle  $\kappa_{\text{eye}}$  is the angle between the  $y_{\text{eye}}z_{\text{eye}}$ -plane of the ECS and the vertical plane that contains the optic axis of the eye. With the definition of the ECS given in Section 2.2.1, when the eye is in the primary position the  $y_{\text{eye}}z_{\text{eye}}$ -plane of the ECS is vertical and therefore  $\kappa_{\text{eye}} = 0$ .

The torsion of the eye depends on the position and orientation of the head (especially the orientation of the head with respect to gravity), the point-of-gaze and whether vision is monocular or binocular. The torsion of the eye is also affected by disorders of the visual and oculomotor systems.

Under a certain set of conditions (the head is erect and stationary, the point-of-gaze is stationary, monocular viewing conditions with the gaze at infinity), the torsion of the eye follows approximately Listing's and Donders' laws [96], [116], [117]. Before stating these laws, some definitions are necessary. The fixation line (a.k.a. line of fixation or fixation axis) is defined as the line connecting the center of rotation of the eye and the fixation point (point-of-gaze). For most practical situations, the fixation line is very close to the visual axis defined in Section 2.2 (they are strictly parallel and about 0.4-0.8 mm apart when the gaze is at infinity). The base line is the line connecting the centers of rotation of both eyes. The primary position can be defined as

the position of the eyes in the skull when the head is in a natural erect position and the fixation line is horizontal and perpendicular to the base line (another definition is provided later).

Listing's law states that each movement of the eyes from the primary position to any other position is described by a single rotation about an axis that is perpendicular to the plane that contains the initial and final positions of the fixation line. Since all possible axes of rotation are perpendicular to the fixation line in the primary position, they are all contained in a vertical plane called the equatorial plane or Listing's plane. In this context, any position of the eye can be fully described by specifying the orientation of the axis of rotation in Listing's plane and the magnitude of the rotation from the primary position. Therefore, two degrees of freedom suffice to describe any eye position completely.

A consequence of Listing's law is that, except for purely horizontal or purely vertical movements, the meridian of the eye that is vertical in the primary position is systematically tilted with respect to the objective vertical (gravity) in any other position (tertiary position). Based on this, the primary position can be defined as the position from which purely horizontal movements or purely vertical movements are not associated with any tilt of the vertical meridian of the eye with respect to the objective vertical [96], [97].

Donders' law states that the angle of the tilt of the vertical meridian of the eye with respect to the objective vertical at any given eye position is always the same regardless of the way the eyes reach that position. This implies that the torsion of the eye in any position is the same as would be observed if the eye moved to it from the primary position by rotation about a single axis in Listing's plane (analytical expressions of Listing's law are provided in Section A.1.1 and a quantification of the torsion predicted by Listing's law is provided in Section A.1.2).

It is important to note that in several recent publications [90]-[94], Listing's and Donders' laws are used as an integral part of gaze estimation methodologies as if they were always valid. However, this is not the case in general since, as mentioned previously, in the best case, Listing's and Donders' laws are only approximate. Listing's law does not apply when the head is not erect, when the head is non-stationary [96] or in binocular conditions, especially when viewing a relatively close object (convergence of the fixation lines, vergence eye movements) [96], [116]. In addition, Listing's and Donders' laws do not hold during smooth pursuit (eye movements that follow a moving target) [118].

Even under monocular viewing conditions when the head is erect and stationary and the point-of-gaze is also stationary, Listing's law is only approximate: physiological eye movements show considerable stochastic as well as systematical deviations from this law [117]. Some of the observations from [117] are: (a) even when the eye is in the primary position there can be considerable fluctuations of torsion; (b) there is asymmetry between the torsion in the nasal quadrants and the torsion in the temporal quadrants; (c) there are variations in torsion between the left and right eyes of any subject, and among subjects.

The preceding discussion suggests that the torsion of the eye should be measured independently, for example, by tracking the iris pattern and/or scleral blood vessels. The independent measurement of torsion, however, is beyond the scope of this thesis.

Despite the obvious limitations of Listing's law, analytical expressions are provided in the next section for the sake of completeness. A quantification of the torsion predicted by Listing's law is provided in Section A.1.2.

### A.1.1 Analytical expressions of Listing's law

In order to provide analytical expressions of Listing's law, let the orientation of the fixation line with respect to the WCS be described by the pan angle,  $\theta_{f.l.}$ , and the tilt angle,  $\varphi_{f.l.}$ , in the same way that the orientation of the optic axis of the eye with respect to the WCS is described by the pan angle,  $\theta_{eye}$ , and the tilt angle,  $\varphi_{eye}$ , defined above. Then, the unit vector in the direction of the fixation line with respect to the WCS is given by

$$\Phi \equiv \Phi(\theta_{f.l.}, \varphi_{f.l.}) = \begin{bmatrix} \sin \theta_{f.l.} \cos \varphi_{f.l.} \\ \sin \varphi_{f.l.} \\ -\cos \theta_{f.l.} \cos \varphi_{f.l.} \end{bmatrix}, \quad (\text{A.7})$$

which has the same form as the unit vector in the direction of the optic axis of the eye with respect to the WCS given by (2.26).

Next, assume that the line connecting the centers of rotation of the eyes (base line) is parallel to the  $X$ -axis of the WCS (e.g., the WCS is attached to a screen as described in Section 2.5.1 and the subject's head is in erect position with the plane of the face parallel to the screen) so that in the primary position the fixation line is parallel to the  $Z$ -axis of the WCS and therefore  $\theta_{f.l.} = \varphi_{f.l.}$

$= 0$  and  $\Phi = \Phi(0,0) = [0 \ 0 \ -1]^T$ . Then, the rotation from the primary position ( $\Phi(0,0)$ ) to any other position ( $\Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})$ ) as prescribed by Listing's law is described by the rotation vector  $\Gamma = \varepsilon \boldsymbol{\lambda}$ , where

$$\boldsymbol{\lambda} = \frac{\Phi(0,0) \times \Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})}{\|\Phi(0,0) \times \Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})\|} = \frac{[0 \ 0 \ -1]^T \times \Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})}{\|[0 \ 0 \ -1]^T \times \Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})\|} \quad (\text{A.8})$$

is the unit vector in the direction of the axis of rotation according to the right-hand rule and

$$\varepsilon = \arccos(\cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}) \quad (\text{A.9})$$

is the magnitude of the rotation relative to the primary position (eccentricity angle).

It can then be shown using Rodrigues' rotation formula [119] that the corresponding rotation matrix is given by

$$\mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) = \begin{bmatrix} 1 - \frac{\sin^2 \theta_{\text{f.l.}} \cos^2 \varphi_{\text{f.l.}}}{1 + \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}} & \frac{-\sin \varphi_{\text{f.l.}} \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}}{1 + \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}} & -\sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \\ \frac{-\sin \varphi_{\text{f.l.}} \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}}{1 + \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}} & \frac{\cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} + \cos^2 \varphi_{\text{f.l.}}}{1 + \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}} & -\sin \varphi_{\text{f.l.}} \\ \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} & \sin \varphi_{\text{f.l.}} & \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \end{bmatrix}, \quad (\text{A.10})$$

and the resulting torsion of the eye about the fixation line,  $\kappa_{\text{f.l.}}$ , is given by

$$\kappa_{\text{f.l.}} = \arcsin \frac{\sin \theta_{\text{f.l.}} \sin \varphi_{\text{f.l.}}}{1 + \cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}}}, \quad (\text{A.11})$$

where, as expected,  $\kappa_{\text{f.l.}} = 0$  in the primary position ( $\theta_{\text{f.l.}} = \varphi_{\text{f.l.}} = 0$ ).

The rotation matrix prescribed by Listing's law,  $\mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})$ , given by (A.10), and the rotation matrix of the ECS with respect to the WCS,  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ , given by (A.6), are related by

$$\mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \mathbf{R}_{\text{eye}}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}}, \kappa_{\text{eye, p.p.}} = 0) = \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}), \quad (\text{A.12})$$

where  $\theta_{\text{eye, p.p.}}$ ,  $\varphi_{\text{eye, p.p.}}$ , and  $\kappa_{\text{eye, p.p.}} = 0$  describe the rotation of the ECS with respect to the WCS when the eye is in the primary position.

In order to continue the derivation of the analytical expressions (including analytical expressions for  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$ ), assume that the fixation line is parallel to the visual axis (or, equivalently, redefine the fixation line as the line that goes through the center of rotation of the eye and is parallel to the visual axis). This is strictly true when the gaze is at infinity, in which case the fixation line and the visual axis are about 0.4-0.8 mm apart. When the point-of-gaze is at a finite but relatively large distance, the assumption that the fixation line and visual axis are parallel is reasonable. For example, for a point-of-gaze at 57 cm from the eye, the directions of the fixation line and the visual axis would differ by less than 5 minutes of arc.

Considering the fixation line and the visual axis to be parallel implies that

$$\begin{aligned} \Phi(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) &= \begin{bmatrix} \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \\ \sin \varphi_{\text{f.l.}} \\ -\cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \end{bmatrix} = \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \Phi(0,0) = \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \\ &= \mathbf{v}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}, \alpha_{\text{eye}}, \beta_{\text{eye}}) = \mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}}) \mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = \\ &= \frac{\mathbf{g} - \mathbf{c}}{\|\mathbf{g} - \mathbf{c}\|}, \end{aligned} \quad (\text{A.13})$$

where  $\mathbf{v}_{\text{ECS}}$  is the unit vector in the direction of the visual axis with respect to the ECS (Section 2.2.2, equation (2.1)),  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  describe the angular deviation of the visual axis from the optic axis of the eye with respect to the ECS (Section 2.2.2),  $\mathbf{v}$  is the unit vector in the direction of the visual axis with respect to the WCS (Section 2.5.3, equation (2.30)),  $\mathbf{c}$  is the nodal point of the eye / center of curvature of the spherical cornea, and  $\mathbf{g}$  is the point-of-gaze.

It then follows that in the primary position ( $\theta_{\text{f.l.}} = \varphi_{\text{f.l.}} = 0$ ),

$$\mathbf{R}_{\text{eye}}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}}, \kappa_{\text{eye, p.p.}} = 0) \mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = \Phi(0,0) = [0 \ 0 \ -1]^T. \quad (\text{A.14})$$

Solving this equation for  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$ , results in

$$\theta_{\text{eye, p.p.}} \equiv \theta_{\text{eye, p.p.}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = -\arctan \frac{\sin \alpha_{\text{eye}} \cos \beta_{\text{eye}}}{\sqrt{\cos^2 \alpha_{\text{eye}} \cos^2 \beta_{\text{eye}} + \sin^2 \beta_{\text{eye}}}} \approx -\alpha_{\text{eye}} \quad (\text{A.15})$$

and

$$\varphi_{\text{eye, p.p.}} \equiv \varphi_{\text{eye, p.p.}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = -\arctan \frac{\tan \beta_{\text{eye}}}{\cos \alpha_{\text{eye}}} \approx -\beta_{\text{eye}} . \quad (\text{A.16})$$

The application of the preceding analytical expressions of Listing's law depends on whether they are used for the generation of synthetic eye feature data, for the estimation of the point-of-gaze or for the single-point personal calibration procedure. These three cases are described in Sections A.1.1.1-A.1.1.3.

### A.1.1.1 Listing's law and the generation of synthetic eye feature data

The first step in the generation of synthetic eye feature data is to determine the orientation of the ECS with respect to the WCS ( $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$ ) given the 3-D positions of the center of the eyeball / center of rotation of the eye,  $\mathbf{d}$ , and the point-of-gaze,  $\mathbf{g}$ , and the deviation of the visual axis from the optic axis of the eye ( $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , or simply  $\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}})$  –see Section 2.2.2). The determination of  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  using Listing's law is carried out as follows. First note that from (2.26) and (A.12),

$$\boldsymbol{\omega}(\theta_{\text{eye}}, \varphi_{\text{eye}}) = \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \boldsymbol{\omega}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}}) , \quad (\text{A.17})$$

where  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$  are given by (A.15) and (A.16), respectively.

Then, combining (A.13), (A.17) and (2.55), it follows that

$$\frac{\mathbf{g} - \mathbf{d} - D \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \boldsymbol{\omega}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}})}{\|\mathbf{g} - \mathbf{d} - D \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \boldsymbol{\omega}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}})\|} = \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \\ \sin \varphi_{\text{f.l.}} \\ -\cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \end{bmatrix} . \quad (\text{A.18})$$

After solving (A.18) for  $\theta_{\text{f.l.}}$  and  $\varphi_{\text{f.l.}}$ ,  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  is determined with (A.12), where  $\mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}})$  is given by (A.10) and  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}}, \kappa_{\text{eye, p.p.}} = 0)$  is given by (A.6) with  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$  given by (A.15) and (A.16), respectively.

If the values of the angles  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$ , and  $\kappa_{\text{eye}}$  are desired, they are found as follows. First,  $\omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$  is found with (A.17) (if  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  has already been determined,  $\omega$  can be found simply as the last column of  $\mathbf{R}_{\text{eye}}$  as indicated in (2.26)). Having  $\omega$ ,  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$  are then calculated from (2.26) as

$$\theta_{\text{eye}} = -\arctan(\omega_x / \omega_z) \quad (\text{A.19})$$

and

$$\varphi_{\text{eye}} = \arcsin(\omega_y) . \quad (\text{A.20})$$

Finally, the angle  $\kappa_{\text{eye}}$  can be obtained from (A.12), or, as it can be shown, from

$$\kappa_{\text{eye}} = \arcsin \left\{ \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \begin{bmatrix} -\cos \theta_{\text{eye, p.p.}} \\ 0 \\ -\sin \theta_{\text{eye, p.p.}} \end{bmatrix} \cdot \begin{bmatrix} -\sin \theta_{\text{eye}} \sin \varphi_{\text{eye}} \\ \cos \varphi_{\text{eye}} \\ \cos \theta_{\text{eye}} \sin \varphi_{\text{eye}} \end{bmatrix} \right\}, \quad (\text{A.21})$$

$\mathbf{i}_{\text{eye}}$                                      $\mathbf{j}_{\text{eye-no torsion}}$

where  $\mathbf{i}_{\text{eye, p.p.}}$  is the unit vector in the direction of the  $x_{\text{eye}}$ -axis of the ECS with respect to the WCS when the eye is in the primary position,  $\mathbf{i}_{\text{eye}}$  is the unit vector in the direction of the  $x_{\text{eye}}$ -axis of the ECS with respect to the WCS when the eye is in any position, and  $\mathbf{j}_{\text{eye-no torsion}}$  is the unit vector in the direction of the  $Y^V$ -axis shown in Fig. A.3 with respect to the WCS, which is equivalent to  $\mathbf{j}_{\text{eye}}$  (the unit vector in the direction of the  $y_{\text{eye}}$ -axis of the ECS with respect to the WCS) when  $\kappa_{\text{eye}} = 0$ .

Note that an initial guess for  $\theta_{\text{f.l.}}$  and  $\varphi_{\text{f.l.}}$  for the solution of (A.18) can be found from  $\mathbf{d}$  and  $\mathbf{g}$  from

$$\begin{bmatrix} \sin \theta_{\text{f.l., initial guess}} & \cos \varphi_{\text{f.l., initial guess}} \\ \sin \varphi_{\text{f.l., initial guess}} & \\ -\cos \theta_{\text{f.l., initial guess}} & \cos \varphi_{\text{f.l., initial guess}} \end{bmatrix} = \mathbf{v}_{\text{initial guess}} = \frac{\mathbf{g} - \mathbf{d}}{\|\mathbf{g} - \mathbf{d}\|} \quad (\text{A.22})$$

as

$$\theta_{\text{f.l., initial guess}} = -\arctan(v_{\text{initial guess}, X} / v_{\text{initial guess}, Z}) \quad (\text{A.23})$$

and

$$\varphi_{\text{f.l., initial guess}} = \arcsin(v_{\text{initial guess}, Y}) . \quad (\text{A.24})$$

### A.1.1.2 Listing's law and the estimation of the point-of-gaze

In the case of the estimation of the point-of-gaze, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , and the direction of the optic axis of the eye with respect to the WCS (described either by the unit vector  $\omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$  or the pan and tilt angles  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ ) are calculated from the pupil and corneal reflections extracted from the eye images captured by the camera(s). Since the point-of-gaze is defined as the intersection of the visual axis with the scene, the problem here consists of finding the direction of the visual axis with respect to the WCS, which is described by the unit vector  $\mathbf{v}$ , given the direction of the optic axis of the eye and the deviation of the visual axis from the optic axis of the eye ( $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ ).

The calculation of the unit vector in the direction of the visual axis,  $\mathbf{v}$ , using Listing's law is carried out as follows. First,  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$  are calculated with (A.15) and (A.16). Next,  $\theta_{\text{f.l.}}$  and  $\varphi_{\text{f.l.}}$  are obtained from (A.17). Finally,  $\mathbf{v}$  is obtained from (A.13), that is,

$$\mathbf{v} = \mathbf{R}_{\text{Listing}}(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} . \quad (\text{A.25})$$

Having found  $\mathbf{v}$  and assuming that the scene is a plane with  $Z = 0$ , the point-of-gaze,  $\mathbf{g}$ , is estimated from (2.31) and (3.61) as

$$\mathbf{g} = \mathbf{c} - \frac{c_Z}{v_Z} \mathbf{v} . \quad (\text{A.26})$$

If the rotation matrix  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  is desired, it is determined with (A.12). If the values of

the angles  $\theta_{\text{eye}}$ ,  $\varphi_{\text{eye}}$  and  $\kappa_{\text{eye}}$  are desired, they can be found with (A.19)-(A.21).

Note that an initial guess for  $\theta_{\text{f.l.}}$  and  $\varphi_{\text{f.l.}}$  for the solution of (A.17) can be obtained as

$$\theta_{\text{f.l., initial guess}} = \theta_{\text{eye}} + \alpha_{\text{eye}} \quad (\text{A.27})$$

and

$$\varphi_{\text{f.l., initial guess}} = \varphi_{\text{eye}} + \beta_{\text{eye}} . \quad (\text{A.28})$$

### A.1.1.3 Listing's law and the single-point personal calibration procedure

When at least two cameras and at least two light sources are used, the optic axis of the eye can be reconstructed in 3-D space without knowing any subject-specific eye parameter. More specifically, the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , and the direction of the optic axis of the eye with respect to the WCS (described either by the unit vector  $\omega(\theta_{\text{eye}}, \varphi_{\text{eye}})$  or the pan and tilt angles  $\theta_{\text{eye}}$  and  $\varphi_{\text{eye}}$ ) are calculated from the pupil and corneal reflections extracted from the eye images captured by the cameras without knowing any subject-specific eye parameter. Since the point-of-gaze is defined as the intersection of the visual axis with the scene, a single-point personal calibration procedure is needed to determine the subject-specific deviation of the visual axis with respect to the optic axis of the eye ( $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$ , or simply  $\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}})$ —see Section 2.2.2).

The problem here, then, consists of finding  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  given  $\mathbf{c}$  and the direction of the optic axis of the eye with respect to the WCS when the subject fixates a known target calibration point  $\mathbf{g}$ . The calculation of  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  using Listing's law is carried out as follows. First,  $\theta_{\text{f.l.}}$  and  $\varphi_{\text{f.l.}}$  are obtained from  $\mathbf{c}$  and  $\mathbf{g}$  using (A.13), that is, from

$$\begin{bmatrix} \sin \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \\ \sin \varphi_{\text{f.l.}} \\ -\cos \theta_{\text{f.l.}} \cos \varphi_{\text{f.l.}} \end{bmatrix} = \mathbf{v} = \frac{\mathbf{g} - \mathbf{c}}{\|\mathbf{g} - \mathbf{c}\|} \quad (\text{A.29})$$

as

$$\theta_{\text{f.l.}} = -\arctan(v_x / v_z) \quad (\text{A.30})$$

and

$$\varphi_{\text{f.l.}} = \arcsin(v_y) . \quad (\text{A.31})$$

Next,  $\omega(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}})$  is obtained from (A.17) as

$$\underbrace{\omega(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}})}_{\boldsymbol{\omega}_{\text{p.p.}}} = \mathbf{R}_{\text{Listing}}^T(\theta_{\text{f.l.}}, \varphi_{\text{f.l.}}) \boldsymbol{\omega}(\theta_{\text{eye}}, \varphi_{\text{eye}}) . \quad (\text{A.32})$$

and then  $\theta_{\text{eye, p.p.}}$  and  $\varphi_{\text{eye, p.p.}}$  are calculated from (2.26) and (A.32) as

$$\theta_{\text{eye, p.p.}} = -\arctan(\omega_{\text{p.p., } X} / \omega_{\text{p.p., } Z}) \quad (\text{A.33})$$

and

$$\varphi_{\text{eye, p.p.}} = \arcsin(\omega_{\text{p.p., } Y}) . \quad (\text{A.34})$$

Finally, it follows from (A.14) that

$$\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}}) = \mathbf{R}_{\text{eye}}^T(\theta_{\text{eye, p.p.}}, \varphi_{\text{eye, p.p.}}, \kappa_{\text{eye, p.p.}} = 0) \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} , \quad (\text{A.35})$$

and, by solving (A.35) for  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  (with  $\mathbf{R}_{\text{eye}}(\theta_{\text{eye}}, \varphi_{\text{eye}}, \kappa_{\text{eye}})$  given by (A.6) and  $\mathbf{v}_{\text{ECS}}(\alpha_{\text{eye}}, \beta_{\text{eye}})$  given by (2.1)), it follows that

$$\alpha_{\text{eye}} = -\arctan \frac{\tan \theta_{\text{eye, p.p.}}}{\cos \varphi_{\text{eye, p.p.}}} \quad (\text{A.36})$$

and

$$\beta_{\text{eye}} = -\arcsin(\cos \theta_{\text{eye, p.p.}} \sin \varphi_{\text{eye, p.p.}}) . \quad (\text{A.37})$$

### A.1.2 Quantification of the torsion predicted by Listing's law

A quantification of the torsion predicted by Listing's law is provided in Table A.1 for a simulated right eye with  $\alpha_{\text{eye}} = -5^\circ$  and  $\beta_{\text{eye}} = 1.5^\circ$ . The last four rows of Table A.1 correspond approximately to the four corners of a 19" computer screen (5:4 aspect ratio) in typical viewing conditions with the eye at about 65 cm from the screen. Note that positive values of  $\theta_{\text{f.l.}}$  correspond to eye rotations to the right of the subject, positive values of  $\varphi_{\text{f.l.}}$  correspond to upward eye rotations, and positive values of  $\kappa_{\text{f.l.}}$  and  $\kappa_{\text{eye}}$  correspond, from the point of view of the subject, to clockwise torsions of the eye.

Table A.1: Quantification of the torsion predicted by Listing's law (all angles are in degrees)

$\theta_{\text{f.l.}}$	$\varphi_{\text{f.l.}}$	$\kappa_{\text{f.l.}}$	$\kappa_{\text{eye}}$
-30	30	-8.21	-5.23
30	30	8.21	10.86
-30	-30	8.21	5.40
30	-30	-8.21	-11.18
-20	10	-1.77	-0.88
20	10	1.77	2.63
-20	-20	3.56	1.75
20	-20	-3.56	-5.41

## Appendix B

### Outline of the Derivation of (2.4)-(2.6)

This appendix provides an outline of the derivation of (2.4)-(2.6), which express the parameters  $a$ ,  $b$ , and  $A$  of the ellipse described by (2.3) as functions of the radius of curvature at the apex of the cornea ( $\tau = 0$ ),  $R_0$ , the radius of curvature of the cornea at a distance  $\tau = \tau^*$  from the optic axis,  $R_{\tau^*}$ , and the distance between the apex of the cornea and the center of the eyeball,  $D_{\text{apex}}$ .

The radius of curvature of a curve described in explicit form as  $y = f(x)$  is given by [120]

$$\rho(x) = \frac{\left[ 1 + \left( \frac{dy}{dx} \right)^2 \right]^{3/2}}{\left| \frac{d^2y}{dx^2} \right|} . \quad (\text{B.1})$$

The ellipse from (2.3) can be represented in explicit form as

$$z_{\text{eye}} = z_{\text{eye}}(\tau) = A \pm \frac{a}{b} \sqrt{b^2 - \tau^2} , \quad |\tau| \leq b . \quad (\text{B.2})$$

From (B.1) and (B.2) it then follows that

$$\rho(\tau) = \frac{(b^4 - b^2\tau^2 + a^2\tau^2)^{3/2}}{ab^4} , \quad |\tau| \leq b . \quad (\text{B.3})$$

Setting  $\tau = 0$  and  $\rho(0) = R_0$  in (B.3) and solving for  $a$ , (2.5) follows. Next, substituting (2.5) into (B.3), setting  $\tau = \tau^*$  and  $\rho(\tau^*) = R_{\tau^*}$ , and solving for  $b$ , (2.4) follows. Finally, noting that  $(A + a)$  equals the distance between the apex of the cornea and the center of the eyeball,  $D_{\text{apex}}$ , (2.6) follows.

## Appendix C

### Relation Between the CCS and the WCS

The relation between the Camera Coordinate System (CCS) and the World Coordinate System (WCS) can be described by the following sequence of elementary transformations:

- 1) Translate the WCS  $XYZ$  to obtain a new coordinate system  $X'YZ'$  whose origin is located at the nodal point of the camera,  $\mathbf{o}$ :

$$\underbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}_{\text{WCS}} = \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \mathbf{o} . \quad (\text{C.1})$$

- 2) Rotate the  $X'YZ'$  coordinate system about the  $Y'$ -axis by the pan angle  $\theta_{\text{cam}}$  as shown in Fig. C.1 ( $\theta_{\text{cam}} > 0$  is shown) to obtain a new coordinate system  $X''Y''Z''$ :

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos \theta_{\text{cam}} & 0 & \sin \theta_{\text{cam}} \\ 0 & 1 & 0 \\ -\sin \theta_{\text{cam}} & 0 & \cos \theta_{\text{cam}} \end{bmatrix}}_{\mathbf{R}_{\theta_{\text{cam}}}} \begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} . \quad (\text{C.2})$$

Note that after this transformation, the  $X''Z''$ -plane remains horizontal / the  $Y''$ -axis remains vertical.

- 3) Rotate the  $X''Y''Z''$  coordinate system about the  $X''$ -axis by the tilt angle  $\varphi_{\text{cam}}$  as shown in Fig. C.2 ( $\varphi_{\text{cam}} > 0$  for upward rotations, as shown in the figure) to obtain a new coordinate system  $X'''Y'''Z'''$  whose  $Z'''$ -axis is coincident with the optic axis of the camera:

$$\begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_{\text{cam}} & \sin \varphi_{\text{cam}} \\ 0 & -\sin \varphi_{\text{cam}} & \cos \varphi_{\text{cam}} \end{bmatrix}}_{\mathbf{R}_{\varphi, \text{cam}}} \begin{bmatrix} X''' \\ Y''' \\ Z''' \end{bmatrix}. \quad (\text{C.3})$$

Note that after this transformation, the  $Y'''Z'''$ -plane remains vertical / the  $X'''$ -axis remains horizontal.

4) Rotate the  $X'''Y'''Z'''$  coordinate system about the  $Z'''$ -axis by the torsion angle  $\kappa_{\text{cam}}$  as shown in Fig. C.3 ( $\kappa_{\text{cam}} > 0$  for counterclockwise rotations, as shown in the figure) to finally obtain the CCS  $x_{\text{cam}}y_{\text{cam}}z_{\text{cam}}$ :

$$\begin{bmatrix} X''' \\ Y''' \\ Z''' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos \kappa_{\text{cam}} & -\sin \kappa_{\text{cam}} & 0 \\ \sin \kappa_{\text{cam}} & \cos \kappa_{\text{cam}} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_{\kappa, \text{cam}}} \underbrace{\begin{bmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ z_{\text{cam}} \end{bmatrix}}_{\text{CCS}}. \quad (\text{C.4})$$

Combining (C.1)-(C.4), (2.25) is obtained with

$$\mathbf{R}_{\text{cam}} \equiv \mathbf{R}_{\text{cam}}(\theta_{\text{cam}}, \varphi_{\text{cam}}, \kappa_{\text{cam}}) = \mathbf{R}_{\theta, \text{cam}} \mathbf{R}_{\varphi, \text{cam}} \mathbf{R}_{\kappa, \text{cam}}. \quad (\text{C.5})$$

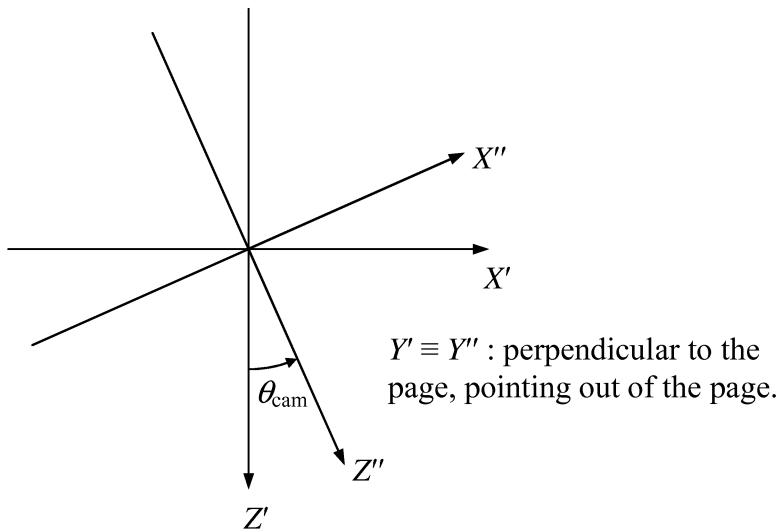


Fig. C.1: Rotation of the  $X'Y'Z'$  coordinate system about the  $Y'$ -axis by the pan angle  $\theta_{\text{cam}}$ .

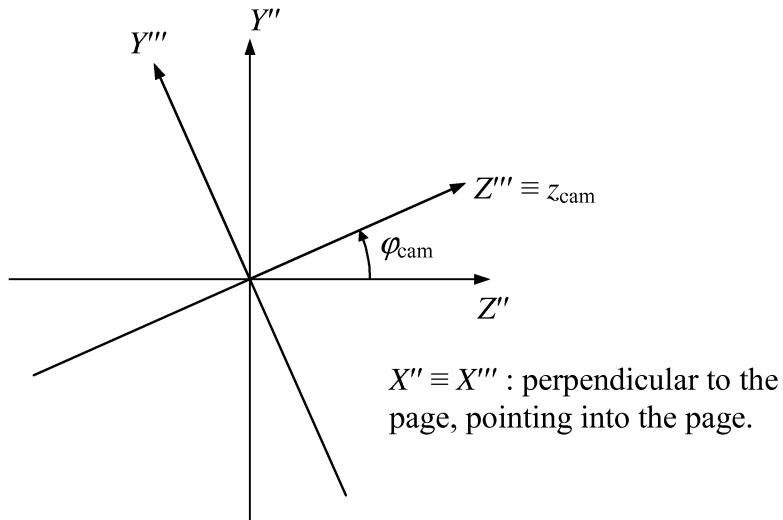


Fig. C.2: Rotation of the  $X''Y''Z''$  coordinate system about the  $X''$ -axis by the tilt angle  $\varphi_{\text{cam}}$ .

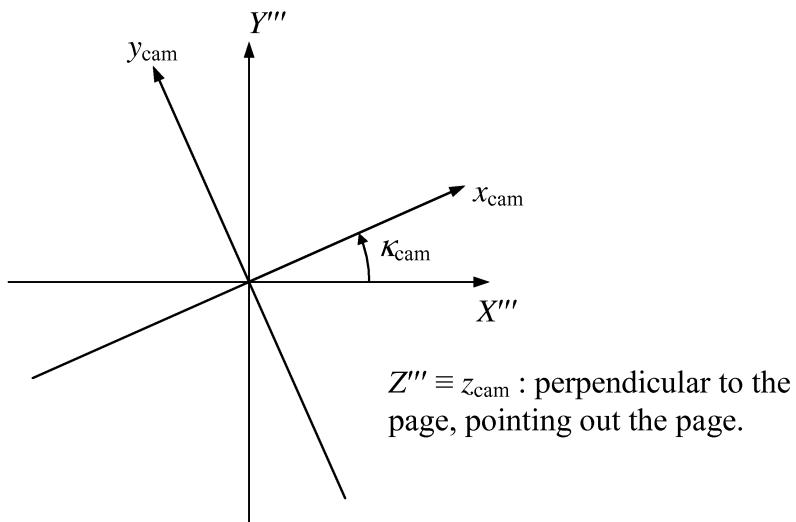


Fig. C.3: Rotation of the  $X'''Y'''Z'''$  coordinate system about the  $Z'''$ -axis by the torsion angle  $\kappa_{\text{cam}}$ .

## Appendix D

### Appendices to Chapter 3

#### D.1 Appendix to Section 3.2.2

Section 3.2.2 discusses the problem of determining the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , with one camera and two light sources. This problem is described by the system of equations constituted by (3.2)-(3.5) with constraint (3.6) (or, equivalently, by the system of equations constituted by (3.2) and (3.7)), with  $i = 1, 2$  and  $j = 1$  (since there is a single camera, the subscript  $j$  that identifies the camera is dropped hereafter for simplicity of notation).

When the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  are distinct and the radius of curvature of the cornea,  $R$ , is known (e.g., estimated through a personal calibration procedure), either system of equations is equivalent to 12 scalar equations with 11 scalar unknowns. Since in this case the number of equations exceeds the number of unknowns by one, it seems that it might be possible to determine  $\mathbf{c}$  without knowing  $R$  a priori. In such case,  $R$  would be treated as an additional scalar unknown, which can be eliminated by replacing (3.3) with

$$\|\mathbf{q}_1 - \mathbf{c}\| = \|\mathbf{q}_2 - \mathbf{c}\|. \quad (\text{D.1})$$

Without intending to consider all possible situations, this appendix presents a realistic example in which it is not possible to determine the coordinates of  $\mathbf{c}$  if  $R$  is not known a priori. In this example, the system setup is symmetrical and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_1$ , and  $\mathbf{u}_1$ , and the plane defined by  $\mathbf{o}$ ,  $\mathbf{l}_2$ , and  $\mathbf{u}_2$  are distinct. Without loss of generality, suppose that the plane of symmetry of the system is the  $YZ$ -plane of the WCS and, therefore,

$$\begin{aligned} \mathbf{o} &= [0 \quad o_Y \quad o_Z]^T, \\ \mathbf{l}_1 &= [-l_X \quad l_Y \quad l_Z]^T, \quad \mathbf{l}_2 = [l_X \quad l_Y \quad l_Z]^T. \end{aligned} \quad (\text{D.2})$$

It then follows that

$$\mathbf{o} = \mathbf{S}\mathbf{o},$$

$$\mathbf{l}_1 = \mathbf{S}\mathbf{l}_2, \quad \mathbf{l}_2 = \mathbf{S}\mathbf{l}_1, \quad (\text{D.3})$$

$$\mathbf{S} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

If the center of curvature of the cornea,  $\mathbf{c}$ , is somewhere in the plane of symmetry of the system (i.e.,  $\mathbf{c} = [0 \ c_Y \ c_Z]^T$ ,  $\mathbf{c} = \mathbf{Sc}$ ), then, due to the symmetry of the system, the points of reflection on the corneal surface are symmetrical (i.e.,  $\mathbf{q}_1 = \mathbf{Sq}_2$  and  $\mathbf{q}_2 = \mathbf{Sq}_1$ ) as are the lines connecting the points of reflection with their respective image points through the nodal point of the camera, i.e.,

$$\frac{\mathbf{o} - \mathbf{u}_1}{\|\mathbf{o} - \mathbf{u}_1\|} = \mathbf{S} \frac{\mathbf{o} - \mathbf{u}_2}{\|\mathbf{o} - \mathbf{u}_2\|}, \quad \frac{\mathbf{o} - \mathbf{u}_2}{\|\mathbf{o} - \mathbf{u}_2\|} = \mathbf{S} \frac{\mathbf{o} - \mathbf{u}_1}{\|\mathbf{o} - \mathbf{u}_1\|}. \quad (\text{D.4})$$

Now, consider the problem of finding the coordinates of  $\mathbf{c}$  from the coordinates of the images of a pair of corneal reflections that satisfy (D.4) using (3.2), (3.5) and (3.6) with  $i = 1, 2$  and  $j = 1$ , (3.8)-(3.9) (this pair of equations replaces (3.4) with  $i = 1, 2$  and  $j = 1$ ) and (D.1). From (D.2)-(D.4), it follows that vector  $\mathbf{b}$  given by (3.8) has the form  $[0 \ b_Y \ b_Z]^T$  and, therefore, from (3.9) and (D.2),

$$\mathbf{c} = \begin{bmatrix} 0 \\ o_Y + k_c b_Y \\ o_Z + k_c b_Z \end{bmatrix}, \quad (\text{D.5})$$

$$\mathbf{c} = \mathbf{Sc},$$

which indicates that  $\mathbf{c}$  is in plane of symmetry of the system, as expected.

If  $k_{q,1} = k_{q,2}$  in (3.2), it then follows from (D.4) that

$$\mathbf{q}_1 = \mathbf{Sq}_2, \quad \mathbf{q}_2 = \mathbf{Sq}_1. \quad (\text{D.6})$$

For each  $k_c > 0$  in **(D.5)** ( $\mathbf{b}$  is assumed to point towards the eye –see note after **(3.8)**) there is a unique value of  $k_{q,i}$  in **(3.2)** such that  $\mathbf{q}_i$  satisfies **(3.5)** and **(3.6)**. Suppose then that for a given  $\tilde{\mathbf{c}}$ ,  $\tilde{\mathbf{q}}_1$  satisfies **(3.5)** and **(3.6)** with  $i = 1$ . In such case, from **(3.5)**,

$$\left( \underbrace{\frac{\mathbf{l}_1 - \tilde{\mathbf{q}}_1}{\|\mathbf{l}_1 - \tilde{\mathbf{q}}_1\|} - \frac{\mathbf{o} - \tilde{\mathbf{q}}_1}{\|\mathbf{o} - \tilde{\mathbf{q}}_1\|}}_{[x_1 \quad x_2 \quad x_3]^T} \right) \bullet \underbrace{[\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3]^T}_{(\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}})} = x_1 y_1 + x_2 y_2 + x_3 y_3 = 0 . \quad (\text{D.7})$$

It then follows from **(D.3)**, **(D.5)** and **(D.6)** that

$$\begin{aligned} \frac{\mathbf{l}_1 - \tilde{\mathbf{q}}_1}{\|\mathbf{l}_1 - \tilde{\mathbf{q}}_1\|} - \frac{\mathbf{o} - \tilde{\mathbf{q}}_1}{\|\mathbf{o} - \tilde{\mathbf{q}}_1\|} &= \frac{\mathbf{l}_1 - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{l}_1 - \mathbf{S}\tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o} - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{o} - \mathbf{S}\tilde{\mathbf{q}}_2\|} = \frac{\mathbf{S}\mathbf{l}_2 - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{S}\mathbf{l}_2 - \mathbf{S}\tilde{\mathbf{q}}_2\|} - \frac{\mathbf{S}\mathbf{o} - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{S}\mathbf{o} - \mathbf{S}\tilde{\mathbf{q}}_2\|} = \\ \mathbf{S} \left( \frac{\mathbf{l}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{l}_2 - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o} - \tilde{\mathbf{q}}_2}{\|\mathbf{o} - \tilde{\mathbf{q}}_2\|} \right) &= [x_1 \quad x_2 \quad x_3]^T \Rightarrow \frac{\mathbf{l}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{l}_2 - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o} - \tilde{\mathbf{q}}_2}{\|\mathbf{o} - \tilde{\mathbf{q}}_2\|} = [-x_1 \quad x_2 \quad x_3]^T \end{aligned} \quad (\text{D.8})$$

and

$$\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}} = \mathbf{S}\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}} = \mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\tilde{\mathbf{c}} = \mathbf{S}(\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}) = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3]^T \Rightarrow \tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}} = [-\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3]^T , \quad (\text{D.9})$$

and, therefore,

$$\begin{aligned} \left( \frac{\mathbf{l}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{l}_2 - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o} - \tilde{\mathbf{q}}_2}{\|\mathbf{o} - \tilde{\mathbf{q}}_2\|} \right) \bullet (\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}) &= [-x_1 \quad x_2 \quad x_3]^T \bullet [-\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3]^T = \\ &= x_1 y_1 + x_2 y_2 + x_3 y_3 = 0 \end{aligned} \quad (\text{D.10})$$

This implies that  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$  satisfies **(3.5)** with  $i = 2$ .

Next, note that from **(3.6)** with  $i = 1$ , **(D.3)** and **(D.6)** it follows that

$$\left. \begin{aligned} \|\tilde{\mathbf{q}}_1 - \mathbf{o}\| &< \|\tilde{\mathbf{c}} - \mathbf{o}\| \\ \|\tilde{\mathbf{q}}_1 - \mathbf{o}\| &= \|\mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{o}\| = \|\mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\mathbf{o}\| = \|\mathbf{S}(\tilde{\mathbf{q}}_2 - \mathbf{o})\| = \|\tilde{\mathbf{q}}_2 - \mathbf{o}\| \end{aligned} \right\} \Rightarrow \|\tilde{\mathbf{q}}_2 - \mathbf{o}\| < \|\tilde{\mathbf{c}} - \mathbf{o}\| , \quad (\text{D.11})$$

and, therefore,  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$  satisfies **(3.6)** with  $i = 2$ .

In addition, whenever  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$ , it follows from **(D.5)** that

$$\|\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}}\| = \|\mathbf{S}\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}\| = \|\mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\tilde{\mathbf{c}}\| = \|\mathbf{S}(\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}})\| = \|\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}\| , \quad (\text{D.12})$$

and, therefore, **(D.1)** is also satisfied.

In summary, whenever  $\mathbf{q}_1$  satisfies **(3.5)** and **(3.6)** with  $i = 1$  and  $\mathbf{q}_2 = \mathbf{S}\mathbf{q}_1$ ,  $\mathbf{q}_2$  satisfies **(3.5)** and **(3.6)** with  $i = 2$ , and  $\mathbf{q}_1$  and  $\mathbf{q}_2$  satisfy **(D.1)**. This implies that **(3.2)**, **(3.5)** and **(3.6)** with  $i = 2$ , and **(D.1)** are replaced by **(D.6)**. Consequently, the system of equations constituted by **(3.2)** and **(3.5)** with  $i = 1, 2$ , **(3.8)-(3.9)** and **(D.1)**, subject to constraint **(3.6)** with  $i = 1, 2$ , is reduced to the system of equations constituted by **(3.2)** and **(3.5)** with  $i = 1$ , **(D.5)** and **(D.6)**, subject to constraint **(3.6)** with  $i = 1$ . This latter system of equations is equivalent to 10 scalar equations with 11 scalar unknowns and, therefore, it is not possible to find a unique solution for  $\mathbf{c}$ .

In other words, whenever the system setup is symmetrical (satisfies **(D.2)-(D.3)**) and the images of the corneal reflections satisfy **(D.4)**, for each  $k_c > 0$  in **(D.5)** there is a  $k_{q,1} = k_{q,2}$  in **(3.2)** such that  $\mathbf{q}_1$  satisfies **(3.5)** and **(3.6)** with  $i = 1$ ,  $\mathbf{q}_2$  satisfies **(3.5)** and **(3.6)** with  $i = 2$ , and  $\mathbf{q}_1$  and  $\mathbf{q}_2$  satisfy **(D.1)**. This means that there are infinite solutions for  $k_c$  and, therefore, infinite solutions for  $\mathbf{c}$ .

In conclusion, the preceding analysis shows that if the system configuration is symmetrical and the radius of curvature of the cornea,  $R$ , is not known, it is not possible to determine the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , when it is in the plane of symmetry of the system. This observation suggests that when  $\mathbf{c}$  is “close” to the plane of symmetry of the system, the calculation of the coordinates of  $\mathbf{c}$  can be very sensitive to noise when  $R$  is not known (ill-conditioned problem). Even when the system configuration is not symmetrical, not using a fixed valued of  $R$  that has been estimated through a personal calibration procedure can result in a very high sensitivity to noise. Furthermore, as discussed in Section 3.5, when a single camera is used, there is no real practical benefit in avoiding the determination of  $R$  through a personal calibration procedure. So, in general, calibrating for  $R$  is recommended.

## D.2 Appendix to Section 3.2.4

Section 3.2.4 discusses the problem of determining the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , with two cameras and one light source. This problem is described by the system of equations constituted by (3.2)-(3.5) with constraint (3.6) (or, equivalently, by the system of equations constituted by (3.2) and (3.7)), with  $i = 1$  and  $j = 1, 2$  (since there is a single light source, the subscript  $i$  that identifies the light source is dropped hereafter for simplicity of notation).

When the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1$  and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2$  are distinct and the radius of curvature of the cornea,  $R$ , is known (e.g., estimated through a personal calibration procedure), either system of equations is equivalent to 12 scalar equations with 11 scalar unknowns. Since in this case the number of equations exceeds the number of unknowns by one, it seems that it might be possible to determine  $\mathbf{c}$  without knowing  $R$  a priori. In such case,  $R$  would be treated as an additional scalar unknown, which can be eliminated by replacing (3.3),  $j = 1, 2$ , with (D.1).

Without intending to consider all possible situations, this appendix presents a realistic example in which it is not possible to determine the coordinates of  $\mathbf{c}$  if  $R$  is not known a priori. In this example, the system setup is symmetrical and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_1$ , and  $\mathbf{u}_1$ , and the plane defined by  $\mathbf{l}$ ,  $\mathbf{o}_2$ , and  $\mathbf{u}_2$  are distinct. Without loss of generality, suppose that the plane of symmetry of the system is the  $YZ$ -plane of the WCS and, therefore,

$$\begin{aligned} \mathbf{l} &= [0 \quad l_Y \quad l_Z]^T, \\ \mathbf{o}_1 &= [-o_X \quad o_Y \quad o_Z]^T, \quad \mathbf{o}_2 = [o_X \quad o_Y \quad o_Z]^T. \end{aligned} \tag{D.13}$$

It then follows that

$$\begin{aligned} \mathbf{l} &= \mathbf{S}\mathbf{l}, \\ \mathbf{o}_1 &= \mathbf{S}\mathbf{o}_2, \quad \mathbf{o}_2 = \mathbf{S}\mathbf{o}_1, \end{aligned} \tag{D.14}$$

$$\mathbf{S} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

If the center of curvature of the cornea,  $\mathbf{c}$ , is somewhere in the plane of symmetry of the system (i.e.,  $\mathbf{c} = [0 \ c_Y \ c_Z]^T$ ,  $\mathbf{c} = \mathbf{Sc}$ ), then, due to the symmetry of the system, the points of reflection on the corneal surface are symmetrical (i.e.,  $\mathbf{q}_1 = \mathbf{Sq}_2$  and  $\mathbf{q}_2 = \mathbf{Sq}_1$ ) as are the lines connecting the points of reflection with their respective image points through the nodal points of the respective cameras, i.e.,

$$\frac{\mathbf{o}_1 - \mathbf{u}_1}{\|\mathbf{o}_1 - \mathbf{u}_1\|} = \mathbf{S} \frac{\mathbf{o}_2 - \mathbf{u}_2}{\|\mathbf{o}_2 - \mathbf{u}_2\|}, \quad \frac{\mathbf{o}_2 - \mathbf{u}_2}{\|\mathbf{o}_2 - \mathbf{u}_2\|} = \mathbf{S} \frac{\mathbf{o}_1 - \mathbf{u}_1}{\|\mathbf{o}_1 - \mathbf{u}_1\|}. \quad (\text{D.15})$$

Now, consider the problem of finding the coordinates of  $\mathbf{c}$  from the coordinates of the images of a corneal reflection that satisfy (D.15) using (3.2), (3.5) and (3.6) with  $i = 1$  and  $j = 1, 2$ , (3.14)-(3.15) (this pair of equations replaces (3.4) with  $i = 1$  and  $j = 1, 2$ ) and (D.1). From (D.13)-(D.15), it follows that vector  $\mathbf{b}$  given by (3.14) has the form  $[0 \ b_Y \ b_Z]^T$  and, therefore, from (3.15) and (D.13),

$$\mathbf{c} = \begin{bmatrix} 0 \\ l_Y + k_c b_Y \\ l_Z + k_c b_Z \end{bmatrix}, \quad (\text{D.16})$$

$$\mathbf{c} = \mathbf{Sc},$$

which indicates that  $\mathbf{c}$  is in plane of symmetry of the system, as expected.

If  $k_{q,1} = k_{q,2}$  in (3.2), it then follows from (D.14) and (D.15) that

$$\mathbf{q}_1 = \mathbf{Sq}_2, \quad \mathbf{q}_2 = \mathbf{Sq}_1. \quad (\text{D.17})$$

Suppose now that for a given  $\tilde{\mathbf{c}}$ ,  $\tilde{\mathbf{q}}_1$  satisfies (3.5) and (3.6) with  $j = 1$ . In such case, from (3.5),

$$\underbrace{\left( \frac{\mathbf{l} - \tilde{\mathbf{q}}_1}{\|\mathbf{l} - \tilde{\mathbf{q}}_1\|} - \frac{\mathbf{o}_1 - \tilde{\mathbf{q}}_1}{\|\mathbf{o}_1 - \tilde{\mathbf{q}}_1\|} \right)}_{[x_1 \quad x_2 \quad x_3]^T} \bullet \underbrace{[\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}}]}_{[y_1 \quad y_2 \quad y_3]^T} = x_1 y_1 + x_2 y_2 + x_3 y_3 = 0 . \quad (\text{D.18})$$

It then follows from **(D.14)**, **(D.16)** and **(D.17)** that

$$\begin{aligned} \frac{\mathbf{l} - \tilde{\mathbf{q}}_1}{\|\mathbf{l} - \tilde{\mathbf{q}}_1\|} - \frac{\mathbf{o}_1 - \tilde{\mathbf{q}}_1}{\|\mathbf{o}_1 - \tilde{\mathbf{q}}_1\|} &= \frac{\mathbf{l} - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{l} - \mathbf{S}\tilde{\mathbf{q}}_2\|} - \frac{\mathbf{S}\mathbf{o}_2 - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{S}\mathbf{o}_2 - \mathbf{S}\tilde{\mathbf{q}}_2\|} = \frac{\mathbf{S}\mathbf{l} - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{S}\mathbf{l} - \mathbf{S}\tilde{\mathbf{q}}_2\|} - \frac{\mathbf{S}\mathbf{o}_2 - \mathbf{S}\tilde{\mathbf{q}}_2}{\|\mathbf{S}\mathbf{o}_2 - \mathbf{S}\tilde{\mathbf{q}}_2\|} = \\ \mathbf{S} \left( \frac{\mathbf{l} - \tilde{\mathbf{q}}_2}{\|\mathbf{l} - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{o}_2 - \tilde{\mathbf{q}}_2\|} \right) &= [x_1 \quad x_2 \quad x_3]^T \Rightarrow \frac{\mathbf{l} - \tilde{\mathbf{q}}_2}{\|\mathbf{l} - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{o}_2 - \tilde{\mathbf{q}}_2\|} = [-x_1 \quad x_2 \quad x_3]^T \end{aligned} \quad (\text{D.19})$$

and

$$\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}} = \mathbf{S}\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}} = \mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\tilde{\mathbf{c}} = \mathbf{S}(\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}) = [y_1 \quad y_2 \quad y_3]^T \Rightarrow \tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}} = [-y_1 \quad y_2 \quad y_3]^T , \quad (\text{D.20})$$

and, therefore,

$$\begin{aligned} \left( \frac{\mathbf{l} - \tilde{\mathbf{q}}_2}{\|\mathbf{l} - \tilde{\mathbf{q}}_2\|} - \frac{\mathbf{o}_2 - \tilde{\mathbf{q}}_2}{\|\mathbf{o}_2 - \tilde{\mathbf{q}}_2\|} \right) \bullet (\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}) &= [-x_1 \quad x_2 \quad x_3]^T \bullet [-y_1 \quad y_2 \quad y_3]^T = \\ &= x_1 y_1 + x_2 y_2 + x_3 y_3 = 0 \end{aligned} \quad (\text{D.21})$$

This implies that  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$  satisfies **(3.5)** with  $j = 2$ .

Next, note that from **(3.6)** with  $j = 1$ , **(D.14)**, **(D.16)** and **(D.17)** it follows that

$$\left. \begin{aligned} \|\tilde{\mathbf{q}}_1 - \mathbf{o}_1\| &< \|\tilde{\mathbf{c}} - \mathbf{o}_1\| \\ \|\tilde{\mathbf{q}}_1 - \mathbf{o}_1\| &= \|\mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\mathbf{o}_2\| = \|\mathbf{S}(\tilde{\mathbf{q}}_2 - \mathbf{o}_2)\| = \|\tilde{\mathbf{q}}_2 - \mathbf{o}_2\| \\ \|\tilde{\mathbf{c}} - \mathbf{o}_1\| &= \|\mathbf{S}\tilde{\mathbf{c}} - \mathbf{S}\mathbf{o}_2\| = \|\mathbf{S}(\tilde{\mathbf{c}} - \mathbf{o}_2)\| = \|\tilde{\mathbf{c}} - \mathbf{o}_2\| \end{aligned} \right\} \Rightarrow \|\tilde{\mathbf{q}}_2 - \mathbf{o}_2\| < \|\tilde{\mathbf{c}} - \mathbf{o}_2\| , \quad (\text{D.22})$$

and, therefore,  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$  satisfies **(3.6)** with  $j = 2$ .

In addition, whenever  $\tilde{\mathbf{q}}_2 = \mathbf{S}\tilde{\mathbf{q}}_1$ , it follows from **(D.16)** that

$$\|\tilde{\mathbf{q}}_1 - \tilde{\mathbf{c}}\| = \|\mathbf{S}\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}\| = \|\mathbf{S}\tilde{\mathbf{q}}_2 - \mathbf{S}\tilde{\mathbf{c}}\| = \|\mathbf{S}(\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}})\| = \|\tilde{\mathbf{q}}_2 - \tilde{\mathbf{c}}\| , \quad (\text{D.23})$$

and, therefore, **(D.1)** is also satisfied.

In summary, whenever  $\mathbf{q}_1$  satisfies **(3.5)** and **(3.6)** with  $j = 1$  and  $\mathbf{q}_2 = \mathbf{S}\mathbf{q}_1$ ,  $\mathbf{q}_2$  satisfies **(3.5)** and **(3.6)** with  $j = 2$ , and  $\mathbf{q}_1$  and  $\mathbf{q}_2$  satisfy **(D.1)**. This implies that **(3.2)**, **(3.5)** and **(3.6)** with  $j = 2$ , and **(D.1)** are replaced by **(D.17)**. Consequently, the system of equations constituted by **(3.2)** and **(3.5)** with  $j = 1, 2$ , **(3.14)-(3.15)** and **(D.1)**, subject to constraint **(3.6)** with  $j = 1, 2$ , is reduced to the system of equations constituted by **(3.2)** and **(3.5)** with  $j = 1$ , **(D.16)** and **(D.17)**, subject to constraint **(3.6)** with  $j = 1$ . This latter system of equations is equivalent to 10 scalar equations with 11 scalar unknowns and, therefore, it is not possible to find a unique solution for  $\mathbf{c}$ .

In other words, whenever the system setup is symmetrical (satisfies **(D.13)-(D.14)**) and the images of the corneal reflection satisfy **(D.15)**, for each  $k_c > 0$  in **(D.16)** ( $\mathbf{b}$  is assumed to point from the light source to the eye –see note after **(3.14)**) there is a  $k_{q,1} = k_{q,2}$  in **(3.2)** such that  $\mathbf{q}_1$  satisfies **(3.5)** and **(3.6)** with  $j = 1$ ,  $\mathbf{q}_2$  satisfies **(3.5)** and **(3.6)** with  $j = 2$ , and  $\mathbf{q}_1$  and  $\mathbf{q}_2$  satisfy **(D.1)**. This means that there are infinite solutions for  $k_c$  and, therefore, infinite solutions for  $\mathbf{c}$ .

In conclusion, the preceding analysis shows that if the system configuration is symmetrical and the radius of curvature of the cornea,  $R$ , is not known, it is not possible to determine the coordinates of the center of curvature of the cornea,  $\mathbf{c}$ , when it is in the plane of symmetry of the system. This observation suggests that when  $\mathbf{c}$  is “close” to the plane of symmetry of the system, the calculation of the coordinates of  $\mathbf{c}$  can be very sensitive to noise when  $R$  is not known (ill-conditioned problem). Even when the system configuration is not symmetrical, not using a fixed valued of  $R$  that has been estimated through a personal calibration procedure can result in a very high sensitivity to noise.

As discussed in Section 3.2.5, by adding a second light source, it is possible to avoid situations like the one just described while being able to calculate the coordinates of  $\mathbf{c}$  without the need to determine  $R$  through a personal calibration procedure and regardless of the relative positions of the eye, the cameras and the light sources. For this reason, a system configuration with two cameras and a single light source is of limited practical interest.

## Appendix E

### Sensitivity to Errors in System Parameters

This appendix presents results of a simulation study of the effect of errors in the values of several system parameters for the system setup studied in Section 4.7.2. In this simulation study, the synthetic eye features are generated with the spherical corneal model and no noise is added to the eye features. The center of the eyeball (**d**) of the simulated right eye adopts 27 different positions (Table 4.34) arranged in a 3-by-3 grid spanning, approximately, the entire volume of operation of the system. For each position of **d**, the simulated eye gazes at 25 points arranged in a 5-by-5 rectangular grid spanning the entire screen (same 25 points-of-gaze as in most of the simulations in Chapter 4). The point-of-gaze is estimated using the actual values of the system parameters, except for the parameter(s) being studied, after calibrating  $\alpha_{\text{eye}}$  and  $\beta_{\text{eye}}$  for the screen center. In most cases, a single parameter is studied at a time. The parameters studied can be divided into (1) coordinates of the light sources, and (2) camera parameters. Before showing the effect of errors in the system parameters, the inherent bias (bias in the absence of errors in the system parameters) of the different point-of-gaze estimation methods tested is provided in Table E.1 (MAX Bias) and Table E.2 (RMS Bias) for reference.

#### E.1 Sensitivity to errors in the coordinates of the light sources

Each coordinate of each light source is tested individually. The simulated error (10 mm) is relatively large since in practice it should be possible to easily measure the coordinates of the light sources within 1-2 mm. The results of these simulations are summarized in Table E.3 through Table E.26. In most cases, the point-of-gaze estimation bias resulting from an error in a coordinate of a light source is smaller than the error in the coordinate of the light source itself. The sensitivity to error in the *Z* coordinate of the light sources tends to be smaller than the sensitivity to error in the *X* and *Y* coordinates. The ... + *PCTriang* methods tend to exhibit the largest bias (primarily due to the inherent bias –compare with Table E.1 and Table E.2) and the *CNoPlInt-3CR/4CR + PB/PB-EF* methods tend to exhibit the smallest bias.

## E.2 Sensitivity to errors in the camera parameters

Each coordinate of the nodal point, the pan, tilt and torsion angles and the focusing distance of camera 1 (left camera from the point of view of the subject) are tested individually. The pan angle,  $\theta_{\text{cam}}$ , the tilt angle,  $\varphi_{\text{cam}}$ , and the torsion angle,  $\kappa_{\text{cam}}$ , of a camera are defined in Appendix C. The general procedure to introduce errors in these angles is as follows. It can be shown from Appendix C that the unit vector in the direction of the optic axis of camera  $j$  is related to the pan and tilt angles of the camera by

$$\mathbf{k}_{\text{cam},j} = \mathbf{R}_{\text{cam},j}(\theta_{\text{cam},j}, \varphi_{\text{cam},j}, \kappa_{\text{cam},j}) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin \theta_{\text{cam},j} \cos \varphi_{\text{cam},j} \\ \sin \varphi_{\text{cam},j} \\ \cos \theta_{\text{cam},j} \cos \varphi_{\text{cam},j} \end{bmatrix}. \quad (\text{E.1})$$

Note that this expression can be also derived directly from Fig. E.1.

Using (E.1), the actual pan and tilt angles of a camera can be readily determined from the actual unit vector in the direction of the optic axis of the camera given by (4.3) as

$$\theta_{\text{cam},j} = \arctan(k_{\text{cam},j,X} / k_{\text{cam},j,Z}) \quad (\text{E.2})$$

and

$$\varphi_{\text{cam},j} = \arcsin(k_{\text{cam},j,Y}), \quad (\text{E.3})$$

respectively, while the actual torsion angle is

$$\kappa_{\text{cam},j} = 0 \quad (\text{E.4})$$

by definition.

Then, given the errors  $\Delta\theta_{\text{cam},j}$ ,  $\Delta\varphi_{\text{cam},j}$  and  $\Delta\kappa_{\text{cam},j}$ , the new camera rotation matrix,  $\tilde{\mathbf{R}}_{\text{cam},j}$ , is obtained by substituting  $\theta_{\text{cam},j}$  by  $(\theta_{\text{cam},j} + \Delta\theta_{\text{cam},j})$ ,  $\varphi_{\text{cam},j}$  by  $(\varphi_{\text{cam},j} + \Delta\varphi_{\text{cam},j})$  and  $\kappa_{\text{cam},j}$  by  $\Delta\kappa_{\text{cam},j}$  in (C.2)-(C.5):

$$\tilde{\mathbf{R}}_{\text{cam},j} = \mathbf{R}_{\text{cam},j}(\theta_{\text{cam},j} + \Delta\theta_{\text{cam},j}, \varphi_{\text{cam},j} + \Delta\varphi_{\text{cam},j}, \underbrace{\kappa_{\text{cam},j} + \Delta\kappa_{\text{cam},j}}_{=0}) . \quad (\text{E.5})$$

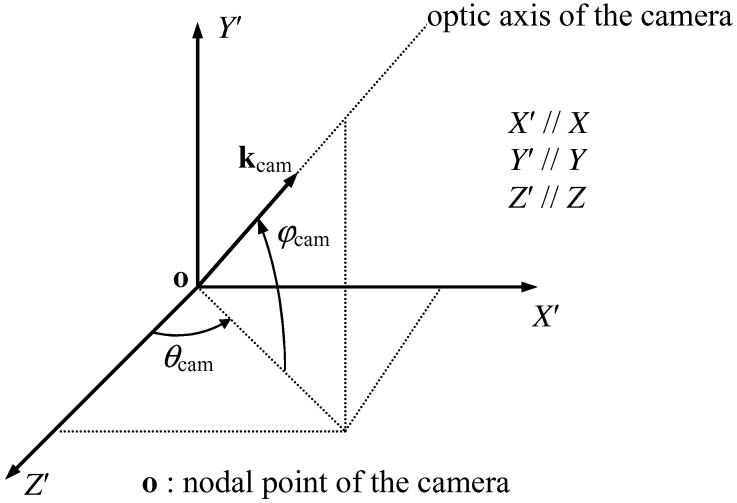


Fig. E.1: Orientation of the optic axis of a camera with respect to the WCS.

The point-of-gaze estimation bias resulting from errors in each coordinate of the nodal point, the pan, tilt and torsion angles and the focusing distance of camera 1 are summarized in Table E.27 through Table E.46. The key observation from these tables is that the only method that consistently exhibits a reasonable performance for all the errors tested is the *CPlIntTriang-4CR + PCTriang* method. Beyond this, a few other observations can be made in general terms. The sensitivity to error in the  $Y$  coordinate of the nodal point of camera 1 ( $o_{1,Y}$ ) is larger than the sensitivity to error in its  $Z$  coordinate ( $o_{1,Z}$ ), which in turn is larger than the sensitivity to error in its  $X$  coordinate ( $o_{1,X}$ ) (Table E.27 through Table E.32). The sensitivity to error in the tilt angle of camera 1 ( $\varphi_{\text{cam},1}$ ) is larger than the sensitivity to error in its pan angle ( $\theta_{\text{cam},1}$ ), which in turn is larger than the sensitivity to error in its torsion angle ( $\kappa_{\text{cam},1}$ ) (Table E.33 through Table E.44). The sensitivity of the ...-3CR + ... methods is larger than the sensitivity of the ...-4CR + ... methods. For the more severe errors in the extrinsic camera parameters tested, the ... + *PB/PB-EF* methods are the most sensitive, followed by the ... + *PCTriangRefr* methods. For five of the errors tested ( $\Delta o_{1,Y} = 1$  mm –see Table E.29 and Table E.30,  $\Delta\theta_{\text{cam},1} = 1^\circ$  –see Table E.35 and

Table E.36,  $\Delta\varphi_{cam,1} = 0.1^\circ$  –see Table E.37 and Table E.38,  $\Delta\varphi_{cam,1} = 1^\circ$  –see Table E.39 and Table E.40, and  $\Delta\kappa_{cam,1} = 1^\circ$  –see Table E.43 and Table E.44) the *CNoPlInt-3CR + PB/PB-EF* methods fail to reconstruct the optic axis of the eye in 3-D space for some, several or most combinations of 3-D positions of the center of the eyeball and the point-of-gaze. The reason of the failure in those cases is that at least one back-projected ray from a pupil boundary image point does not intersect the corneal surface<sup>13</sup>. For two of the errors tested ( $\Delta\theta_{cam,1} = 1^\circ$  –see Table E.35 and Table E.36, and  $\Delta\varphi_{cam,1} = 1^\circ$  –see Table E.39 and Table E.40) the *CNoPlInt-4CR + PB/PB-EF* methods also fail to reconstruct the optic axis of the eye in 3-D space in many cases for the same reason. For one of the errors tested ( $\Delta\varphi_{cam,1} = 1^\circ$  –see Table E.39 and Table E.40) the ... + *PCTriangRefr* methods fail to reconstruct the optic axis of the eye in 3-D space in many cases as, at least for one of the cameras, the back-projected ray from the center of the pupil image does not intersect the corneal surface<sup>14</sup>. From the preceding observations it follows that  $\Delta\varphi_{cam,1} = 1^\circ$  is the most severe of all the errors in system parameters tested. For this particular case, the *CPlIntTriang-4CR + PCTriang* method is the only method that exhibits an acceptable performance.

In summary, except for the *CPlIntTriang-4CR + PCTriang* method, all the other methods are highly sensitive to relatively small errors in one, several or most of the parameters that describe the position and orientation of the cameras. It must be noted, however, that the key problem is in errors in the relative position and orientation of the cameras with respect to each other. If the camera pair is treated as a unit in which the relative position and orientation of the cameras with respect to each other remain constant, relatively large errors in the position and orientation of the camera pair result in relatively small point-of-gaze estimation bias. For example, consider the case of estimating the point-of-gaze using the camera parameters that correspond to translating the camera pair up by 10 mm, i.e., having an error of 10 mm in the  $Y$  coordinate of the nodal points of both cameras. The point-of-gaze estimation bias corresponding to this case is summarized in Table E.47 (MAX Bias) and Table E.48 (RMS Bias). As it can be observed from these tables, the point-of-gaze estimation bias is relatively small. In contrast, a 1 mm error in the  $Y$  coordinate of the nodal point of only one of the cameras (Table E.29 and Table E.30) has a

<sup>13</sup>  $k_{r_b,jk}$  in (3.48) is not a real number.

<sup>14</sup>  $k_{r_j}$  in (3.29) is not a real number.

significant effect and only the *CPIIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods exhibit an acceptable performance (the performance of the *CNoPInt-4CR + PCTriangRefr* method could also be acceptable in some cases). As another example, consider the case of estimating the point-of-gaze using the camera parameters that correspond to tilting the camera pair by  $1^\circ$  about the line defined by the nodal points of the cameras, i.e., using the camera rotation matrix given by

$$\tilde{\mathbf{R}}_{\text{cam},j} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \Delta\varphi_{\text{cam.pair}} & \sin \Delta\varphi_{\text{cam.pair}} \\ 0 & -\sin \Delta\varphi_{\text{cam.pair}} & \cos \Delta\varphi_{\text{cam.pair}} \end{bmatrix} \mathbf{R}_{\text{cam},j}, \quad j = 1, 2, \quad (\text{E.6})$$

with  $\Delta\varphi_{\text{cam.pair}} = 1^\circ$ , where  $\mathbf{R}_{\text{cam},j}$  is the actual rotation matrix of camera  $j$ . The point-of-gaze estimation bias corresponding to this case is summarized in Table E.49 (MAX Bias) and Table E.50 (RMS Bias). As it can be observed from these tables, the point-of-gaze estimation bias is also relatively small. In contrast, a  $0.1^\circ$  error in the tilt angle of only one of the cameras (Table E.37 and Table E.38) has a very significant effect, except on the *CPIIntTriang-4CR + PCTriang* and the *CNoPInt-4CR + PCTriang* methods. When there is an error of  $1^\circ$  in the tilt angle of only one of the cameras (Table E.39 and Table E.40), the situation is dramatically worse and only the *CPIIntTriang-4CR + PCTriang* method exhibits a reasonable performance.

In summary, the simulation results from this appendix indicate that the structure that holds the cameras must be robust enough to ensure that the relative position and orientation of the cameras with respect to each other remain constant over time after the system has been set up and calibrated. In the presence of errors in the relative position and orientation of the cameras with respect to each other, only the *CPIIntTriang-4CR + PCTriang* method exhibits a reasonable performance for the magnitude of the errors tested.

Table E.1: Inherent bias (no error in system parameters) – *MAX Bias* [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	9.65	N/A	N/A	N/A
<i>CNoPPIInt-3CR</i>	9.65	3.66	$9.36 \cdot 10^{-5}$	1.45
<i>CNoPPIInt-4CR</i>	9.65	3.66	$8.70 \cdot 10^{-5}$	1.45

Table E.2: Inherent bias (no error in system parameters) – *RMS Bias* [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	2.57	N/A	N/A	N/A
<i>CNoPPIInt-3CR</i>	2.57	$9.53 \cdot 10^{-1}$	$1.29 \cdot 10^{-5}$	$2.37 \cdot 10^{-1}$
<i>CNoPPIInt-4CR</i>	2.57	$9.53 \cdot 10^{-1}$	$1.29 \cdot 10^{-5}$	$2.37 \cdot 10^{-1}$

Table E.3: Effect of error in the coordinates of the light sources – Error in  $l_{1,X} = 10$  mm – *MAX Bias* [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	9.65	N/A	N/A	N/A
<i>CNoPPIInt-3CR</i>	9.65	5.56	5.49	5.52
<i>CNoPPIInt-4CR</i>	9.65	5.39	5.40	5.44

Table E.4: Effect of error in the coordinates of the light sources – Error in  $l_{1,X} = 10$  mm – *RMS Bias* [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	4.45	N/A	N/A	N/A
<i>CNoPPIInt-3CR</i>	3.19	2.35	2.18	2.22
<i>CNoPPIInt-4CR</i>	3.23	2.40	2.24	2.28

Table E.5: Effect of error in the coordinates of the light sources – Error in  $l_{1,y} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.65	N/A	N/A	N/A
CNoPPIInt-3CR	9.65	7.40	7.10	7.10
CNoPPIInt-4CR	9.65	6.16	5.90	5.90

Table E.6: Effect of error in the coordinates of the light sources – Error in  $l_{1,y} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	3.70	N/A	N/A	N/A
CNoPPIInt-3CR	3.31	2.45	2.25	2.27
CNoPPIInt-4CR	3.53	2.79	2.54	2.57

Table E.7: Effect of error in the coordinates of the light sources – Error in  $l_{1,z} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.65	N/A	N/A	N/A
CNoPPIInt-3CR	9.65	3.66	3.16	3.17
CNoPPIInt-4CR	9.65	3.66	3.14	3.15

Table E.8: Effect of error in the coordinates of the light sources – Error in  $l_{1,z} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	3.39	N/A	N/A	N/A
CNoPPIInt-3CR	2.91	1.51	1.02	1.03
CNoPPIInt-4CR	2.93	1.53	1.05	1.06

Table E.9: Effect of error in the coordinates of the light sources – Error in  $l_{2,X} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.76	N/A	N/A	N/A
CNoPPIInt-3CR	6.58	5.61	5.05	5.04
CNoPPIInt-4CR	5.76	5.37	5.10	4.96

Table E.10: Effect of error in the coordinates of the light sources – Error in  $l_{2,X} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	5.11	N/A	N/A	N/A
CNoPPIInt-3CR	3.45	3.25	3.28	3.30
CNoPPIInt-4CR	3.35	3.50	3.62	3.63

Table E.11: Effect of error in the coordinates of the light sources – Error in  $l_{2,Y} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	13.21	N/A	N/A	N/A
CNoPPIInt-3CR	15.05	9.38	8.07	8.10
CNoPPIInt-4CR	12.75	8.05	7.90	7.93

Table E.12: Effect of error in the coordinates of the light sources – Error in  $l_{2,Y} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	4.96	N/A	N/A	N/A
CNoPPIInt-3CR	6.33	5.40	4.87	4.80
CNoPPIInt-4CR	5.89	4.97	4.50	4.47

Table E.13: Effect of error in the coordinates of the light sources – Error in  $l_{2,Z} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	7.76	N/A	N/A	N/A
CNoPPIInt-3CR	7.66	3.08	2.64	2.89
CNoPPIInt-4CR	7.31	3.05	2.62	3.23

Table E.14: Effect of error in the coordinates of the light sources – Error in  $l_{2,Z} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.75	N/A	N/A	N/A
CNoPPIInt-3CR	2.46	1.64	1.70	1.75
CNoPPIInt-4CR	2.33	1.69	1.84	1.89

Table E.15: Effect of error in the coordinates of the light sources – Error in  $l_{3,X} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.65	N/A	N/A	N/A
CNoPPIInt-3CR	9.65	6.96	6.67	6.69
CNoPPIInt-4CR	9.65	5.90	5.82	5.84

Table E.16: Effect of error in the coordinates of the light sources – Error in  $l_{3,X} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.50	N/A	N/A	N/A
CNoPPIInt-3CR	3.88	3.50	3.52	3.53
CNoPPIInt-4CR	3.51	2.88	2.83	2.85

Table E.17: Effect of error in the coordinates of the light sources – Error in  $l_{3,y} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.79	N/A	N/A	N/A
CNoPPIInt-3CR	9.65	6.40	6.18	6.15
CNoPPIInt-4CR	9.65	6.33	5.64	5.63

Table E.18: Effect of error in the coordinates of the light sources – Error in  $l_{3,y} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	5.05	N/A	N/A	N/A
CNoPPIInt-3CR	4.09	3.21	2.94	2.93
CNoPPIInt-4CR	3.98	3.20	3.00	2.98

Table E.19: Effect of error in the coordinates of the light sources – Error in  $l_{3,z} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.65	N/A	N/A	N/A
CNoPPIInt-3CR	9.65	3.66	3.64	3.65
CNoPPIInt-4CR	9.65	3.66	3.15	3.15

Table E.20: Effect of error in the coordinates of the light sources – Error in  $l_{3,z} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	3.58	N/A	N/A	N/A
CNoPPIInt-3CR	2.90	1.78	1.58	1.62
CNoPPIInt-4CR	3.03	1.87	1.59	1.64

Table E.21: Effect of error in the coordinates of the light sources – Error in  $l_{4,X} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	14.98	N/A	N/A	N/A
CNoPPIInt-3CR	13.75	9.31	9.00	8.99
CNoPPIInt-4CR	15.35	9.55	9.22	9.21

Table E.22: Effect of error in the coordinates of the light sources – Error in  $l_{4,X} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.62	N/A	N/A	N/A
CNoPPIInt-3CR	7.62	6.57	5.86	5.82
CNoPPIInt-4CR	7.89	6.76	6.02	5.97

Table E.23: Effect of error in the coordinates of the light sources – Error in  $l_{4,Y} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.50	N/A	N/A	N/A
CNoPPIInt-3CR	6.70	6.51	5.14	5.20
CNoPPIInt-4CR	6.84	3.60	3.00	3.42

Table E.24: Effect of error in the coordinates of the light sources – Error in  $l_{4,Y} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	4.34	N/A	N/A	N/A
CNoPPIInt-3CR	3.42	3.39	3.15	3.24
CNoPPIInt-4CR	2.68	2.17	2.06	2.15

Table E.25: Effect of error in the coordinates of the light sources – Error in  $l_{4,Z} = 10$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	8.21	N/A	N/A	N/A
CNoPPIInt-3CR	8.15	2.45	1.49	2.42
CNoPPIInt-4CR	8.64	2.66	0.92	1.99

Table E.26: Effect of error in the coordinates of the light sources – Error in  $l_{4,Z} = 10$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.55	N/A	N/A	N/A
CNoPPIInt-3CR	2.37	0.86	0.55	0.67
CNoPPIInt-4CR	2.44	0.85	0.37	0.51

Table E.27: Effect of error in the camera parameters – Error in  $o_{1,X} = 1$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.76	N/A	N/A	N/A
CNoPPIInt-3CR	9.85	3.86	0.55	1.00
CNoPPIInt-4CR	9.78	3.79	0.49	1.05

Table E.28: Effect of error in the camera parameters – Error in  $o_{1,X} = 1$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.61	N/A	N/A	N/A
CNoPPIInt-3CR	2.63	1.01	0.21	0.21
CNoPPIInt-4CR	2.62	1.00	0.20	0.23

Table E.29: Effect of error in the camera parameters – Error in  $o_{1,y} = 1$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.53	N/A	N/A	N/A
CNoPPIInt-3CR	67.02	91.41	failed	failed
CNoPPIInt-4CR	10.92	12.57	33.23	33.82

Table E.30: Effect of error in the camera parameters – Error in  $o_{1,y} = 1$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.53	N/A	N/A	N/A
CNoPPIInt-3CR	33.28	39.08	failed	failed
CNoPPIInt-4CR	3.29	6.27	14.00	15.27

Table E.31: Effect of error in the camera parameters – Error in  $o_{1,z} = 1$  mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.16	N/A	N/A	N/A
CNoPPIInt-3CR	35.53	42.41	46.72	44.41
CNoPPIInt-4CR	8.41	4.18	13.58	19.60

Table E.32: Effect of error in the camera parameters – Error in  $o_{1,z} = 1$  mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.43	N/A	N/A	N/A
CNoPPIInt-3CR	16.60	17.06	18.28	17.90
CNoPPIInt-4CR	2.10	1.83	5.73	7.48

Table E.33: Effect of error in the camera parameters – Error in  $\theta_{cam,1} = 0.1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.78	N/A	N/A	N/A
CNoPPIInt-3CR	11.72	10.39	9.03	8.68
CNoPPIInt-4CR	9.99	3.88	2.08	3.49

Table E.34: Effect of error in the camera parameters – Error in  $\theta_{cam,1} = 0.1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.62	N/A	N/A	N/A
CNoPPIInt-3CR	4.63	4.35	3.90	3.73
CNoPPIInt-4CR	2.72	1.05	0.70	1.17

Table E.35: Effect of error in the camera parameters – Error in  $\theta_{cam,1} = 1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	11.05	N/A	N/A	N/A
CNoPPIInt-3CR	154.20	732.38	failed	failed
CNoPPIInt-4CR	13.03	71.58	failed	failed

Table E.36: Effect of error in the camera parameters – Error in  $\theta_{cam,1} = 1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	3.29	N/A	N/A	N/A
CNoPPIInt-3CR	55.57	105.13	failed	failed
CNoPPIInt-4CR	4.18	19.79	failed	failed

Table E.37: Effect of error in the camera parameters – Error in  $\varphi_{cam,1} = 0.1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.71	N/A	N/A	N/A
CNoPPIInt-3CR	116.81	254.61	failed	failed
CNoPPIInt-4CR	11.47	32.42	89.67	90.57

Table E.38: Effect of error in the camera parameters – Error in  $\varphi_{cam,1} = 0.1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.60	N/A	N/A	N/A
CNoPPIInt-3CR	52.62	76.88	failed	failed
CNoPPIInt-4CR	3.63	14.32	30.97	31.93

Table E.39: Effect of error in the camera parameters – Error in  $\varphi_{cam,1} = 1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	10.18	N/A	N/A	N/A
CNoPPIInt-3CR	1412.09	failed	failed	failed
CNoPPIInt-4CR	81.10	failed	failed	failed

Table E.40: Effect of error in the camera parameters – Error in  $\varphi_{cam,1} = 1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.95	N/A	N/A	N/A
CNoPPIInt-3CR	589.08	failed	failed	failed
CNoPPIInt-4CR	25.76	failed	failed	failed

Table E.41: Effect of error in the camera parameters – Error in  $\kappa_{\text{cam},1} = 0.1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.45	N/A	N/A	N/A
CNoPPIInt-3CR	11.62	7.00	6.44	6.22
CNoPPIInt-4CR	9.41	3.45	1.53	3.26

Table E.42: Effect of error in the camera parameters – Error in  $\kappa_{\text{cam},1} = 0.1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.57	N/A	N/A	N/A
CNoPPIInt-3CR	3.18	2.19	1.83	1.77
CNoPPIInt-4CR	2.51	0.89	0.39	0.68

Table E.43: Effect of error in the camera parameters – Error in  $\kappa_{\text{cam},1} = 1^\circ$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	8.24	N/A	N/A	N/A
CNoPPIInt-3CR	80.97	102.79	failed	failed
CNoPPIInt-4CR	7.37	24.99	54.52	57.95

Table E.44: Effect of error in the camera parameters – Error in  $\kappa_{\text{cam},1} = 1^\circ$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	3.36	N/A	N/A	N/A
CNoPPIInt-3CR	21.29	25.69	failed	failed
CNoPPIInt-4CR	2.62	7.43	14.94	16.17

Table E.45: Effect of error in the camera parameters – Error in  $focusing\ distance_1 = 50\ mm$  – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	10.61	N/A	N/A	N/A
CNoPPIInt-3CR	12.73	7.56	6.82	6.53
CNoPPIInt-4CR	10.34	4.41	1.24	2.30

Table E.46: Effect of error in the camera parameters – Error in  $focusing\ distance_1 = 50\ mm$  – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.90	N/A	N/A	N/A
CNoPPIInt-3CR	3.43	2.45	1.99	1.89
CNoPPIInt-4CR	2.86	1.22	0.40	0.68

Table E.47: Effect of error in the camera parameters – Error in  $o_Y$  for both cameras = 10 mm – MAX Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	9.73	N/A	N/A	N/A
CNoPPIInt-3CR	8.91	3.07	2.88	4.10
CNoPPIInt-4CR	9.51	3.68	2.35	3.57

Table E.48: Effect of error in the camera parameters – Error in  $o_Y$  for both cameras = 10 mm – RMS Bias [mm]

Method	PCTriang	PCTriangRefr	PB	PB-EF
CPIIntTriang-4CR	2.64	N/A	N/A	N/A
CNoPPIInt-3CR	2.42	0.92	0.95	1.11
CNoPPIInt-4CR	2.50	0.96	0.85	0.99

Table E.49: Effect of error in the camera parameters – Error in  $\varphi_{\text{cam. pair}} = 1^\circ$  – MAX Bias [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	9.59	N/A	N/A	N/A
<i>CNoPInt-3CR</i>	9.62	3.64	1.29	1.31
<i>CNoPInt-4CR</i>	9.80	3.82	1.32	1.30

Table E.50: Effect of error in the camera parameters – Error in  $\varphi_{\text{cam. pair}} = 1^\circ$  – RMS Bias [mm]

<i>Method</i>	<i>PCTriang</i>	<i>PCTriangRefr</i>	<i>PB</i>	<i>PB-EF</i>
<i>CPIIntTriang-4CR</i>	2.68	N/A	N/A	N/A
<i>CNoPInt-3CR</i>	2.55	0.99	0.48	0.51
<i>CNoPInt-4CR</i>	2.54	1.03	0.60	0.60