

# **Modeling and Calibration of Automated Zoom Lenses**

Reg G. Willson

January 1994

CMU-RI-TR-94-03

The Robotics Institute  
Carnegie Mellon University  
Pittsburgh, Pennsylvania 15213

This document was originally published as a Ph.D. dissertation in the  
Department of Electrical and Computer Engineering, Carnegie Mellon University.

Thesis Committee:

Prof. Steve Shafer (co-advisor)  
Prof. Takeo Kanade (co-advisor)  
Prof. B. V. K. Vijaya Kumar  
Dr. Eric Krotkov

Copyright © 1994 Reg G. Willson

This research is sponsored by the Department of the Army, Army Research Office under grant number DAAH04-94-G-0006, and by Bell-Northern Research Ltd. Views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing official policies or endorsements, either expressed or implied, of the Department of the Army, the United States Government, or Bell-Northern Research.



# Abstract

Camera systems with automated zoom lenses are inherently more useful than those with fixed-parameter lenses. Variable-parameter lenses enable us to produce better images by matching the camera's sensing characteristics to the conditions in a scene. They also allow us to make measurements by noting how the scene's image changes as the parameters are varied. The reason variable-parameter lenses are not more commonly used in machine vision is that they are difficult to model for continuous ranges of lens settings.

We show in this thesis that traditional modeling approaches cannot capture the complex relationships between control parameters and imaging processes. Furthermore, we demonstrate that the assumption of idealized behavior in traditional models can lead to significant performance problems in color imaging and focus ranging. By using more complex models and control strategies we were able to reduce or eliminate these performance problems.

The principal contribution of our research is a methodology for empirically producing accurate camera models for systems with variable-parameter lenses. We also developed a comprehensive taxonomy for the property of "image center." To demonstrate the effectiveness of our methodology we applied it to produce an "adjustable," perspective-projection camera model based on Tsai's fixed camera model. We calibrated and tested our model on two different automated camera systems. In both cases the calibrated model operated across continuous ranges of focus and zoom with an average error of less than 0.14 pixels between the predicted and the measured positions of features in the image plane. We also calibrated and tested our model on one automated camera system across a continuous range of aperture and achieved similar results.



## ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors, Steve Shafer and Takeo Kanade, for all their help and support during my years here. I thank Steve for his instruction and direction, for his great insights and ideas, and mostly for the freedom he gave me to explore. I also owe Takeo Kanade many thanks for his early guidance, and especially for providing such an amazingly fertile environment to work in – CMU’s Vision and Autonomous Systems Center (VASC).

I would like to thank the other members of my thesis committee, Eric Krotkov and B.V.K. Vijaya Kumar, for their time and for their thoughtful recommendations.

Many others have contributed to this work and deserve my thanks. Mark DeLouis for building the automated lenses for our lab – his ingenuity and attention to detail were a critical factor in the success of this work. Jim Moody, Bill Ross, and Ralph Hyre for our group’s systems support. Jutta Denton for reviewing this manuscript – her many comments and suggestions have greatly improved the clarity of this thesis. John Krumm and Carol Novak for the many great discussions, from the technical to the non-technical. Greg Martin at Eastman Kodak for the Kodak Megaplus XRC data used in Appendix B. Karl Kluge for Figure 1.1. And all the members of VASC, both past and present – they’ve been a tremendous source of support and a lot of fun to work with.

Thanks also to Bell-Northern Research and ARPA for their financial support.

In the years I’ve spent at CMU I’ve easily learned as much outside of school as I’ve learned inside. Many of the best friends I have have come from the years I’ve spent here. I’d like to thank Matt Vea, gReg Silvus, Dagmara Berztiss, Roly Schaefer, Linda Zirnitis, Mark Stahl, Kevin Lorei, Chris Paredis, Debbie Scapatura, Nick Papanikopoulos, Sandy Ramos, Jim Rehg, Laura Kelmar, R.S. Sreenivas, and many others for their friendship and the time spent together.

Finally, I would like to thank my parents for their unwavering encouragement and support. The farther on I go the more I realize how much I’ve really learned from them.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Why use adjustable lenses? . . . . .	1
1.2	Physical versus abstract operating spaces . . . . .	1
1.3	The modeling and calibration problem . . . . .	4
1.4	Review of related work . . . . .	5
1.5	Hardware overview . . . . .	9
1.6	Thesis organization . . . . .	11
<b>2</b>	<b>Active Lens Control for Measurement</b>	<b>13</b>
2.1	Active Color Imaging . . . . .	13
2.1.1	Chromatic aberration . . . . .	14
2.1.2	Compensation approach . . . . .	18
2.1.3	Experimental results . . . . .	21
2.1.4	Conclusion . . . . .	22
2.2	Constant Magnification Focusing . . . . .	25
2.2.1	Focus magnification . . . . .	25
2.2.2	Compensation approach . . . . .	28
2.2.3	Experimental results . . . . .	29
2.2.4	Conclusion . . . . .	30
2.3	Summary . . . . .	32
<b>3</b>	<b>Image Center</b>	<b>33</b>
3.1	Camera calibration and image center . . . . .	33
3.2	Real lenses . . . . .	34
3.3	A taxonomy of image centers . . . . .	39
3.3.1	Non-imaging definitions . . . . .	40
3.3.2	Single image definitions . . . . .	41
3.3.3	Multi-image definitions . . . . .	48
3.3.4	Experimental results . . . . .	50
3.4	Image center in variable-parameter lenses . . . . .	52
3.5	Summary . . . . .	55
<b>4</b>	<b>Modeling and Calibrating Automated Zoom Lenses</b>	<b>57</b>
4.1	Fixed- versus variable-parameter lenses . . . . .	57
4.1.1	$R_1$ — Image-formation process and the hardware configuration . . . . .	58

4.1.2	$R_2$ — Hardware configuration and the control parameters . . . . .	58
4.1.3	Formulation issues . . . . .	59
4.1.4	Calibration issues . . . . .	59
4.2	Adjustable camera models . . . . .	59
4.2.1	Notation . . . . .	61
4.2.2	Collecting calibration data . . . . .	62
4.2.3	Characterizing variations in fixed model parameters . . . . .	63
4.3	Summary . . . . .	66
<b>5</b>	<b>A Fixed Perspective-Projection Camera Model</b>	<b>69</b>
5.1	Perspective-projection camera models . . . . .	69
5.2	Formulation of Tsai's model . . . . .	70
5.3	Performance metrics . . . . .	72
5.4	Calibration data . . . . .	73
5.5	Model calibration . . . . .	74
5.5.1	Initial parameter estimation . . . . .	74
5.5.2	Full non-linear optimization . . . . .	75
5.5.3	Public archive for code . . . . .	75
5.6	Recalibrating exterior orientation . . . . .	77
5.7	Calibration example . . . . .	78
5.8	Exterior orientation recalibration example . . . . .	81
5.9	Summary . . . . .	81
<b>6</b>	<b>Adjustable Perspective-Projection Camera Models</b>	<b>85</b>
6.1	Definitions . . . . .	85
6.2	Performance metrics . . . . .	86
6.3	Fujinon/Photometrics model for focus and zoom . . . . .	87
6.3.1	Operating range . . . . .	87
6.3.2	Sampling strategy . . . . .	88
6.3.3	Calibration data . . . . .	88
6.3.4	Initial fixed camera model calibration . . . . .	88
6.3.5	Formulating the parameter models . . . . .	93
6.3.6	Fitting the parameter models . . . . .	93
6.3.7	Validating the adjustable model . . . . .	95
6.4	Recalibrating exterior orientation . . . . .	100
6.5	Cosmicar/Panasonic model for focus and zoom . . . . .	101
6.5.1	Operating range . . . . .	101
6.5.2	Sampling strategy . . . . .	101
6.5.3	Calibration data . . . . .	101
6.5.4	Initial fixed camera model calibration . . . . .	103
6.5.5	Formulating the parameter models . . . . .	103
6.5.6	Fitting the parameter models . . . . .	103
6.5.7	Validating the adjustable model . . . . .	107
6.5.8	Exterior orientation recalibration example . . . . .	110
6.6	Fujinon/Photometrics model for aperture . . . . .	115

6.7 Discussion . . . . .	118
6.8 Summary . . . . .	120
<b>7 Conclusions</b>	<b>123</b>
7.1 Contributions . . . . .	123
7.2 Directions for future research . . . . .	125
<b>A Calibrated Imaging Lab</b>	<b>127</b>
A.1 General facilities . . . . .	127
A.2 Cosmicar/Panasonic camera system . . . . .	128
A.3 Fujinon/Photometrics camera system . . . . .	129
A.4 Lens automation . . . . .	130
A.5 Summary of computer-controlled parameters . . . . .	130
<b>B Chromatic Aberration Measurement</b>	<b>133</b>
B.1 Measurement approach . . . . .	133
B.2 Examples . . . . .	140
<b>C Sub-pixel Centroid Measurement</b>	<b>153</b>
<b>D Aperture Calibration</b>	<b>157</b>
D.1 Aperture stops . . . . .	157
D.2 Relative and photometric aperture . . . . .	157
D.3 Depth of field . . . . .	161
D.4 Lens aberrations and aperture . . . . .	164



# List of Figures

1.1	Example of a complex natural scene . . . . .	2
1.2	Basic pinhole image-formation model . . . . .	3
1.3	Basic thin-lens image-formation model . . . . .	3
1.4	Cosmicar lens and Panasonic camera . . . . .	10
1.5	Fujinon lens and Photometrics camera . . . . .	10
2.1	Chromatic aberration in a thin lens . . . . .	14
2.2	Checkerboard test target used to measure lateral chromatic aberration . . . . .	15
2.3	Image of grid of black lines on a white background . . . . .	16
2.4	Pixel intensity profiles near center of image (cross section 1) . . . . .	17
2.5	Pixel intensity profiles near edge of image (cross section 2) . . . . .	17
2.6	Active Color Imaging calibration target . . . . .	19
2.7	Active Color Imaging demonstration target . . . . .	20
2.8	Active Color Imaging results . . . . .	20
2.9	Full field blue/red misregistration — uncompensated image . . . . .	23
2.10	Full field blue/red misregistration — compensated image . . . . .	23
2.11	Color histogram — uncompensated image . . . . .	24
2.12	Color histogram — compensated image . . . . .	24
2.13	Focus magnification in a thin lens . . . . .	25
2.14	Simple image magnification versus focus and zoom motors . . . . .	26
2.15	Focus magnification causing feature clipping . . . . .	27
2.16	Focus magnification causing perimeter scaling . . . . .	27
2.17	Focus magnification causing gradient scaling . . . . .	27
2.18	Box and bar targets for perimeter scaling experiment . . . . .	29
2.19	Focus versus zoom settings for both compensated and uncompensated focusing . . . . .	30
2.20	Criterion function peak position versus target type and focusing approach . . . . .	31
2.21	Graded bar and solid bar targets for gradient scaling experiment . . . . .	31
3.1	Different image centers for the Fujinon/Photometrics camera system . . . . .	35
3.2	Shift in image center as a function of the focus and zoom motors . . . . .	35
3.3	Decentration for a simple lens . . . . .	37
3.4	Fixed focal length lens (From [25]) . . . . .	37
3.5	Variable focal length (zoom) lens (From [25]) . . . . .	37
3.6	Non-linear motion of lens groups during mechanical compensation (From [43]) . . . . .	38
3.7	Mechanical compensation cam (From [43]) . . . . .	38
3.8	Vanishing points of a right-angled cube . . . . .	42

3.9	Center of lines of interpretation . . . . .	43
3.10	Center of field of view . . . . .	44
3.11	Center of an autocollimated laser . . . . .	45
3.12	Pixel intensity profile for row 200 from Figure 3.14 . . . . .	46
3.13	Vignetting in a lens (From [25]) . . . . .	47
3.14	Image of a uniform white field showing sharply defined vignetting . . . . .	47
3.15	Field curvature in the image produced by a thin lens . . . . .	48
3.16	Mean image plane error as a function of image center used during calibration . . . . .	51
3.17	Mechanical repeatability of shift in laser image due to focus motor . . . . .	53
3.18	Mechanical repeatability of shift in laser image due to zoom motor . . . . .	53
3.19	Mechanical hysteresis in shift in laser image due to focus motor . . . . .	54
3.20	Mechanical hysteresis in shift in laser image due to zoom motor . . . . .	54
4.1	Lens control parameters and the image-formation process . . . . .	58
4.2	Checkerboard target at 130 mm focal length and 2 m range . . . . .	60
4.3	Checkerboard target at 10 mm focal length and 2 m range . . . . .	60
5.1	Fixed perspective-projection camera model geometry . . . . .	70
5.2	Transformation from undistorted sensor to distorted frame coordinates . . . . .	71
5.3	Calibration target and translation stage . . . . .	76
5.4	Calibration target at 1.5 m . . . . .	79
5.5	Calibration target at 2.5 m . . . . .	79
5.6	Residual error for calibration target at 1.5 m (magnified 100 $\times$ ) . . . . .	80
5.7	Residual error for calibration target at 2.5 m (magnified 100 $\times$ ) . . . . .	80
5.8	Calibration target at 1.5 m with camera shifted . . . . .	82
5.9	Calibration target at 2.5 m with camera shifted . . . . .	82
5.10	Experiment in carrying intrinsic parameters to new pose . . . . .	83
6.1	Variation in X coordinate of autocollimated laser's image with focus and zoom . . . . .	89
6.2	Variation in Y coordinate of autocollimated laser's image with focus and zoom . . . . .	89
6.3	Fujinon fixed camera model $R_x$ versus focus and zoom motors . . . . .	90
6.4	Fujinon fixed camera model $R_y$ versus focus and zoom motors . . . . .	90
6.5	Fujinon fixed camera model $R_z$ versus focus and zoom motors . . . . .	90
6.6	Fujinon fixed camera model $T_x$ versus focus and zoom motors . . . . .	90
6.7	Fujinon fixed camera model $T_y$ versus focus and zoom motors . . . . .	91
6.8	Fujinon fixed camera model $T_z$ versus focus and zoom motors . . . . .	91
6.9	Fujinon fixed camera model $C_x$ versus focus and zoom motors . . . . .	91
6.10	Fujinon fixed camera model $C_y$ versus focus and zoom motors . . . . .	91
6.11	Fujinon fixed camera model $f$ versus focus and zoom motors . . . . .	92
6.12	Fujinon fixed camera model $\kappa_1$ versus focus and zoom motors . . . . .	92
6.13	Fujinon fixed camera model $s_x$ versus focus and zoom motors . . . . .	92
6.14	Fujinon fixed camera model M_UIPE versus focus and zoom motors . . . . .	92
6.15	Fujinon adjustable camera model $f$ versus focus and zoom motors . . . . .	96
6.16	Fujinon adjustable camera model $T_z$ versus focus and zoom motors . . . . .	96
6.17	Fujinon adjustable camera model $C_x$ versus focus and zoom motors . . . . .	96
6.18	Fujinon adjustable camera model $C_y$ versus focus and zoom motors . . . . .	96

6.19 Fujinon adjustable camera model $\kappa_1$ versus focus and zoom motors . . . . .	97
6.20 Fujinon adjustable camera model M_UIPE versus focus and zoom . . . . .	97
6.21 Difference between the final adjustable model M_UIPE and the initial fixed model M_UIPE . . . . .	97
6.22 M_UIPE for model calibrated with dataset 1 and tested on dataset 1 . . . . .	99
6.23 M_UIPE for model calibrated with dataset 2 and tested on dataset 2 . . . . .	99
6.24 M_UIPE for model calibrated with dataset 1 and tested on dataset 2 . . . . .	99
6.25 M_UIPE for model calibrated with dataset 2 and tested on dataset 1 . . . . .	99
6.26 Extrinsic (and intrinsic) parameter changes with lens settings . . . . .	100
6.27 Variation in X coordinate of autocollimated laser's image with focus and zoom	102
6.28 Variation in Y coordinate of autocollimated laser's image with focus and zoom	102
6.29 Cosmicar fixed camera model $R_x$ versus focus and zoom motors . . . . .	104
6.30 Cosmicar fixed camera model $R_y$ versus focus and zoom motors . . . . .	104
6.31 Cosmicar fixed camera model $R_z$ versus focus and zoom motors . . . . .	104
6.32 Cosmicar fixed camera model $T_x$ versus focus and zoom motors . . . . .	104
6.33 Cosmicar fixed camera model $T_y$ versus focus and zoom motors . . . . .	105
6.34 Cosmicar fixed camera model $T_z$ versus focus and zoom motors . . . . .	105
6.35 Cosmicar fixed camera model $C_x$ versus focus and zoom motors . . . . .	105
6.36 Cosmicar fixed camera model $C_y$ versus focus and zoom motors . . . . .	105
6.37 Cosmicar fixed camera model $f$ versus focus and zoom motors . . . . .	106
6.38 Cosmicar fixed camera model $\kappa_1$ versus focus and zoom motors . . . . .	106
6.39 Cosmicar fixed camera model $s_x$ versus focus and zoom motors . . . . .	106
6.40 Cosmicar fixed camera model M_UIPE versus focus and zoom motors . . . . .	106
6.41 Cosmicar adjustable camera model $f$ versus focus and zoom motors . . . . .	108
6.42 Cosmicar adjustable camera model $T_z$ versus focus and zoom motors . . . . .	108
6.43 Cosmicar adjustable camera model $C_x$ versus focus and zoom motors . . . . .	108
6.44 Cosmicar adjustable camera model $C_y$ versus focus and zoom motors . . . . .	108
6.45 Cosmicar adjustable camera model $\kappa_1$ versus focus and zoom motors . . . . .	109
6.46 Cosmicar adjustable camera model M_UIPE versus focus and zoom . . . . .	109
6.47 Difference between the final adjustable model M_UIPE and the initial fixed model M_UIPE . . . . .	109
6.48 M_UIPE for model calibrated with dataset 1 and tested on dataset 1 . . . . .	111
6.49 M_UIPE for model calibrated with dataset 2 and tested on dataset 2 . . . . .	111
6.50 M_UIPE for model calibrated with dataset 1 and tested on dataset 2 . . . . .	111
6.51 M_UIPE for model calibrated with dataset 2 and tested on dataset 1 . . . . .	111
6.52 Experiment in carrying intrinsic parameter models to new pose . . . . .	112
6.53 M_UIPE for model calibrated at pose 1 . . . . .	113
6.54 M_UIPE for model calibrated at pose 2 . . . . .	113
6.55 M_UIPE for extrinsic recalibration using 1 base setting . . . . .	113
6.56 M_UIPE for extrinsic recalibration using 4 base settings . . . . .	113
6.57 Fujinon camera model $R_x$ versus aperture motor . . . . .	116
6.58 Fujinon camera model $R_y$ versus aperture motor . . . . .	116
6.59 Fujinon camera model $R_z$ versus aperture motor . . . . .	116
6.60 Fujinon camera model $T_x$ versus aperture motor . . . . .	116
6.61 Fujinon camera model $T_y$ versus aperture motor . . . . .	116

6.62 Fujinon camera model $T_z$ versus aperture motor . . . . .	116
6.63 Fujinon camera model $C_x$ versus aperture motor . . . . .	117
6.64 Fujinon camera model $C_y$ versus aperture motor . . . . .	117
6.65 Fujinon camera model $f$ versus aperture motor . . . . .	117
6.66 Fujinon camera model $\kappa_1$ versus aperture motor . . . . .	117
6.67 Fujinon camera model $s_x$ versus aperture motor . . . . .	117
6.68 Fujinon camera model M_UIPE versus aperture motor . . . . .	117
A.1 Motor/lens pushrod assembly . . . . .	131
A.2 Motor/lens double pushrod assembly . . . . .	131
A.3 Motor/lens pulley assembly . . . . .	131
B.1 Checkerboard target used for measuring lateral chromatic aberration . . . . .	135
B.2 Thresholded edge image of checkerboard . . . . .	135
B.3 Test mask for finding horizontal reference edges on checkerboard . . . . .	136
B.4 Edge image with highlighted reference edges . . . . .	136
B.5 Blue/Red vertical misregistration . . . . .	137
B.6 Blue/Green vertical misregistration . . . . .	137
B.7 Blue/Red horizontal misregistration . . . . .	138
B.8 Blue/Green horizontal misregistration . . . . .	138
B.9 Blue/Red total misregistration . . . . .	139
B.10 Blue/Green total misregistration . . . . .	139
B.11 Kowa 12.5-75 mm at 55 mm - blue/red misregistration . . . . .	141
B.12 Kowa 12.5-75 mm at 55 mm - blue/green misregistration . . . . .	141
B.13 Kowa 12.5-75 mm at 55 mm - red/green misregistration . . . . .	141
B.14 Nikon 35-105 mm at 85 mm - blue/red misregistration . . . . .	142
B.15 Nikon 35-105 mm at 85 mm - blue/green misregistration . . . . .	142
B.16 Nikon 35-105 mm at 85 mm - red/green misregistration . . . . .	142
B.17 Tokina 35-200 mm at 150 mm - blue/red misregistration . . . . .	143
B.18 Tokina 35-200 mm at 150 mm - blue/green misregistration . . . . .	143
B.19 Tokina 35-200 mm at 150 mm - red/green misregistration . . . . .	143
B.20 Nikon 200 mm - blue/red misregistration . . . . .	144
B.21 Nikon 200 mm - blue/green misregistration . . . . .	144
B.22 Nikon 200 mm - red/green misregistration . . . . .	144
B.23 Nikon 50 mm AF Nikkor - blue/red misregistration . . . . .	145
B.24 Nikon 50 mm AF Nikkor - blue/green misregistration . . . . .	145
B.25 Nikon 50 mm AF Nikkor - red/green misregistration . . . . .	145
B.26 Nikon 50 mm Nikkor - blue/red misregistration . . . . .	146
B.27 Nikon 50 mm Nikkor - blue/green misregistration . . . . .	146
B.28 Nikon 50 mm Nikkor - red/green misregistration . . . . .	146
B.29 Nikon 50 mm Nikkor - blue/red misregistration . . . . .	147
B.30 Nikon 50 mm Nikkor - blue/green misregistration . . . . .	147
B.31 Nikon 50 mm Nikkor - red/green misregistration . . . . .	147
B.32 Nikon 24 mm Nikkor - blue/red misregistration . . . . .	148
B.33 Nikon 24 mm Nikkor - blue/green misregistration . . . . .	148

B.34 Nikon 24 mm Nikkor - red/green misregistration . . . . .	148
B.35 Nikon 20 mm AF Nikkor - blue/red misregistration . . . . .	149
B.36 Nikon 20 mm AF Nikkor - blue/green misregistration . . . . .	149
B.37 Nikon 20 mm AF Nikkor - red/green misregistration . . . . .	149
B.38 Nikon 28 mm Nikkor - blue/red misregistration . . . . .	150
B.39 Nikon 28 mm Nikkor - blue/green misregistration . . . . .	150
B.40 Nikon 28 mm Nikkor - red/green misregistration . . . . .	150
B.41 Nikon 55 mm micro-Nikkor - blue/red misregistration . . . . .	151
B.42 Nikon 55 mm micro-Nikkor - blue/green misregistration . . . . .	151
B.43 Nikon 55 mm micro-Nikkor - red/green misregistration . . . . .	151
B.44 Elicar 90 mm V-HQ - blue/red misregistration . . . . .	152
B.45 Elicar 90 mm V-HQ - blue/green misregistration . . . . .	152
B.46 Elicar 90 mm V-HQ - red/green misregistration . . . . .	152
C.1 Typical image of a black reference point on a white background . . . . .	153
C.2 Intensity profiles when blur width is greater than width of dot's cross section	155
C.3 Intensity profiles when blur width is less than width of dot's cross section .	155
D.1 Basic pinhole model . . . . .	158
D.2 Basic thin-lens model with wide and narrow apertures . . . . .	158
D.3 Light flux versus aperture motor . . . . .	160
D.4 Light flux (log scale) versus aperture motor . . . . .	162
D.5 Error between flux data and flux model versus aperture motor . . . . .	162
D.6 <i>f</i> -number versus aperture motor . . . . .	163
D.7 T-number versus aperture motor . . . . .	163
D.8 Depth of field for a thin lens . . . . .	164
D.9 Image magnification versus aperture motor . . . . .	165



# List of Tables

2.1	Typical camera and lens settings for corrected color images . . . . .	21
3.1	Different image centers for the same camera system. . . . .	51
5.1	Example of a calibrated fixed camera model . . . . .	78
5.2	Calibration results from carrying intrinsic parameters to new pose . . . . .	83
6.1	Choice of polynomial orders for Fujinon/Photometrics parameter models . .	94
6.2	Fitting sequence for Fujinon/Photometrics parameter models for focus and zoom . . . . .	95
6.3	Zero-order parameter models for two sets of calibration data from the same pose . . . . .	98
6.4	Performance of two adjustable models on two sets of calibration data from the same pose . . . . .	98
6.5	Fitting sequence for Cosmicar/Panasonic parameter models for focus and zoom	107
6.6	Zero-order parameter models for two sets of calibration data from the same pose . . . . .	110
6.7	Performance of two adjustable models on two sets of calibration data from the same pose . . . . .	110
6.8	Calibration results when carrying intrinsic parameter models to new pose . .	114
6.9	Fitting sequence for Fujinon/Photometrics parameter models for aperture . .	118
A.1	Controllable parameters for the Cosmicar/Panasonic camera system and jig .	132
A.2	Controllable parameters for the Fujinon/Photometrics camera system and jig	132
B.1	Description of lenses measured for lateral chromatic aberration . . . . .	140
D.1	Effect of aperture radius on primary lens aberrations (From [43]) . . . . .	165



# Chapter 1

## Introduction

The goal of machine vision is to understand the visible world by inferring 3D properties from 2D images. Making an inference requires models describing how a camera and its lens form images. Adjustable lenses offer far greater capability and flexibility than fixed-parameter lenses, however, most machine-vision systems have been limited to cameras with fixed lenses because of the difficulty of modeling cameras with adjustable lenses.

This thesis analyzes the issues involved and develops a methodology for empirically constructing camera models for adjustable lenses.

### 1.1 Why use adjustable lenses?

Consider the complex natural scene shown in Figure 1.1. This scene contains considerable variations in light intensity, spatial detail, and physical distance. In the real world the range of conditions that a camera system may need to image under, be it radiometric sensitivity, spatial resolution, or focused distance, can often exceed the capabilities of a camera with a fixed-parameter lens. To adapt to the conditions at hand camera systems require lenses whose imaging parameters can be changed under computer control. Variable-parameter lenses can also be used to determine properties of a scene (e.g. range) by noting how the scene's image changes as the camera's image-formation process is changed.

Whether for adaptation or measurement, to effectively use adjustable lenses we need to have models of the camera's image-formation process that work across ranges of lens settings.

### 1.2 Physical versus abstract operating spaces

In machine vision we need to know aspects of a camera's image-formation process that range from simple properties, such as magnification and focused distance, to more complex image properties, such as perspective projection and image defocus. In order to have computa-



Figure 1.1: Example of a complex natural scene

tionally efficient, closed-form equations for the more complex properties we use models that are based on simplifications or abstractions of the lens's true image-formation process. The two most common abstract models are the pinhole camera model and the thin-lens camera model, used respectively to explain perspective projection and image defocus.

In the basic pinhole camera model, illustrated in Fig. 1.2, the 3D coordinates of a point  $P(x_c, y_c, z_c)$  are related to the 2D coordinates of its image  $P'(X_u, Y_u)$  by the equations

$$\begin{aligned} X_u &= f \frac{x_c}{z_c}, \\ Y_u &= f \frac{y_c}{z_c} \end{aligned}$$

where  $f$  is the effective focal length of the pinhole camera model.

In the thin-lens camera model, illustrated in Fig. 1.3, the position of a point  $P$  in front of the lens is related to the position of the point's focused image  $P'$  behind the lens by the equation

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$

where  $f$  is the focal length of the thin-lens model,  $s$  is the object to lens separation and  $s'$  in the lens to focused image plane separation.

For fixed lenses the set of abstract model parameters (e.g.  $f$ ,  $s$ , and  $s'$ ) consists of constants. For adjustable lenses the parameters vary with different lens settings. The range of the

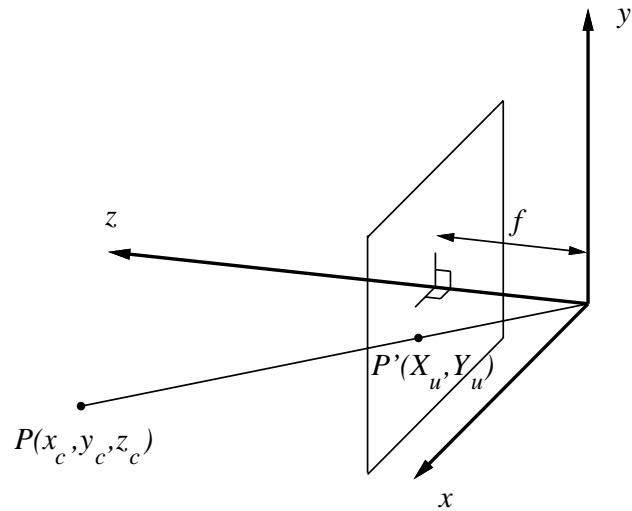


Figure 1.2: Basic pinhole image-formation model

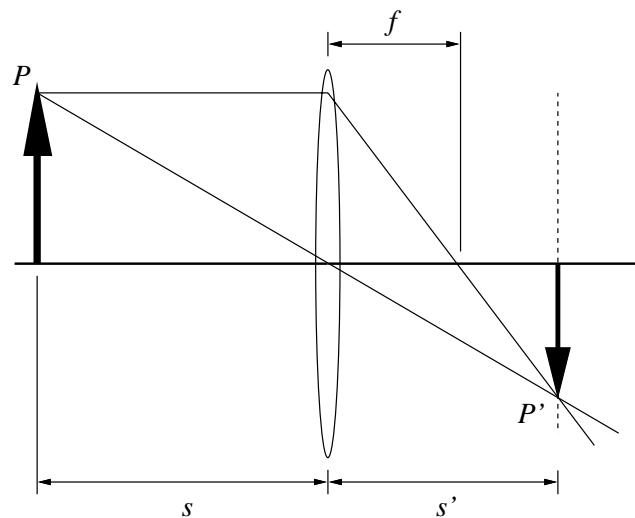


Figure 1.3: Basic thin-lens image-formation model

abstract model parameters defines an *abstract operating space* for the camera system. Unfortunately we cannot directly control the values of the abstract camera model parameters. Instead we control the camera's physical configuration using actuators, such as focus, zoom, and aperture motors. The range of settings for these actuators defines a *physical operating space* for the camera system. In machine-vision tasks the interpretation of image data or the planning of a sensing strategy (i.e. the reasoning) is typically carried out in the abstract operating space of the camera system using the simplified model of the image-formation process. However, the actual implementation of the task must be carried out in the camera's physical operating space. Calibrated camera models provide the bridge between the physical and abstract operating spaces of the camera system.

## 1.3 The modeling and calibration problem

In camera systems with fixed-parameter lenses the camera's image-formation process is fixed. All we need to do to calibrate a camera model is take measurements of the camera system and estimate the values of the constant terms in the model's formulation.

For cameras with automated lenses the image-formation process varies with the lens's actuators, and thus many model terms are functions of the actuator settings. So, instead of calibrating constants we have to formulate and calibrate functions that describe the relationships between model terms and actuator settings. This is difficult for several reasons:

1. Abstract models (e.g. the pinhole projection model or the thin-lens defocus model) are too far removed from the real image-formation process to provide very useful insights into how the actuator settings are related to the parameters of the final image. They cannot, for example, explain how geometric lens distortion varies with lens settings.
2. The relationship between the lens's optical configuration and the actuator settings is a design choice made by the manufacturer and typically unknown to the user. Moreover, the relationship is mechanical and exhibits complex phenomena, such as hysteresis and discontinuity, that are difficult to model and calibrate.
3. The dimensionality of the required calibration data and the wide range of imaging conditions over which it must be taken makes data acquisition for variable-parameter lenses much more complex than for fixed-parameter lenses.
4. To develop complex adjustable camera models (e.g. for perspective projection) functional relationships must be jointly calibrated across its full range of lens settings. The number of parameters and the volume of data required can make this problem computationally prohibitive.

## 1.4 Review of related work

While camera and lens technology has improved significantly in the 25 years since variable-parameter lenses were first developed for machine-vision tasks, the approaches used to model them have changed little. Much of the previous work on building predictive models for variable-parameter camera systems centered on finding ways to use the parameters in the lens's abstract operating space for accommodation or for measurement. In this section we concentrate on previous efforts to solve the "calibration problem" by relating abstract model parameters to the physical control settings of the lens.

Horn's focusing work [23] in 1968 on MIT's MAC system was one of the first uses of computer-controlled cameras in machine vision. The MAC system featured computer-controlled focus and aperture. His objective was to automatically focus a lens by servoing the focus actuator until the camera's image sharpness peaked. A predictive model relating the focused distance of the lens to the actuator settings was proposed but not actually developed.

In the early 70s Sobel and Tenenbaum used computer-controlled cameras in their work at Stanford's AI laboratory. The Stanford system consisted of a camera with four fixed-focal-length lenses and a color-filter wheel mounted on a pan-tilt mechanism. The camera's focus, aperture, the choice of lens and color filter, the pan and tilt, and the camera's gain and digitization parameters were all under computer control.

Sobel's objective was to design a 3D to 2D perspective-projection camera model for acquiring and centering objects in a scene (viewpoint planning) [44][45]. His model of perspective projection was based on a simple pinhole camera with no distortion correction. The relatively low precision of the hardware enabled Sobel to use simple zero and first-order polynomials to capture the basic relationships between the actuator settings and the camera model parameters. A total of 16 coefficients had to be estimated for each of the four lenses.

The objective of Tenenbaum's work was auto-focusing and auto-irising accommodation for edge tracking[51]. While Tenenbaum used a thin lens to explain defocus in his auto-focus work, he did not develop a model to relate the focused distance of the lens to the actual actuator settings of the camera system.

In the early 80s Bracho, Schlag, et al. [10][41] developed the POPEYE system at Carnegie Mellon as a test-bed for machine-vision research in visual inspection, object classification, and interactive control tasks. The system featured an automated lens with continuously variable focus, zoom, iris, as well as computer-controlled camera pan and tilt. Two empirical models relating imaging parameters to actuator settings were developed for the system. The first model related the camera's focused distance to the focus-motor setting, the second related the camera's angular field-of-view to the zoom-motor setting. In both cases the models were implemented with sparse lookup tables and interpolation. Second-order lens behaviors, such as changes in focused distance with zooming and changes in the field-of-view with focusing, were not modeled.

One of the first serious attempts to model and calibrate the relationships between imaging parameters and control motors was made by Krotkov at the University of Pennsylvania's

GRASP laboratory in the mid 80s [27]. Krotkov's objective was active exploratory sensing of spatial layout using focus ranging and stereopsis. The agile stereo system he developed featured two cameras with continuously variable focus, zoom, and aperture. Both cameras were mounted on a platform that allowed them to pan, tilt, and verge.

To model focused distance as a function of the focus motor setting, Krotkov used a thick-lens model and assumed that the focus-motor position was linearly related to the distance between the lens's rear nodal point and the image plane (parameter  $s'$  for the thin lens in Figure 1.3). Coefficients for the model were estimated from 30 measurements of focused distance and focus-motor position. The model was calibrated at a single focal length. While Krotkov noted that the image centers actually move as the focus and zoom of the lens are varied, he used a fixed value for image center in his stereo system. A value for image center was calculated for each lens by varying the zoom motor and calculating the focus of expansion. A simple model relating aperture diameter to the aperture motor was presented, but the model was not calibrated. While he measured relative magnification as a function of zoom, Krotkov did not build a model relating the magnification to the zoom-motor setting.

A system similar to Krotkov's was developed at the University of Illinois at Urbana-Champaign by Abbott [1] in the late 80s. The objective of Abbott's system was shape recovery for textured surfaces by dynamic integration of focus, camera vergence, and stereo cues. Abbott's system also featured two cameras with continuously variable focus, zoom, and aperture. Both cameras were mounted on a platform that allowed them to pan, tilt, and verge.

Abbott built three models relating image parameters to the lens-control parameters. The first modeled focused distance as a function of focus-motor settings. The formulation of this model was based on the assumption that the focus motor was linearly related to the lens-to-image-plane distance in a thin lens (parameter  $s'$  in Figure 1.3). The model was calibrated at one fixed zoom position. Abbott's second model used relative image magnification as a function of zoom motor. A general, second-order polynomial was used for the model formulation. The model was calibrated at one fixed focus position. The third model built by Abbott modeled relative magnification as a function of focus-motor position. Again a general, second-order polynomial was used for the model formulation. The model was calibrated at one fixed zoom position. Like Krotkov, Abbott assumed fixed image centers for his camera lenses. Calibration for image center was accomplished by measuring the focus of expansion with zoom.

One of the first researchers to make extensive use of variable-parameter camera models for viewpoint planning was Tarabanis [48][49]. Developed at IBM's T.J. Watson Research Center, Tarabanis' system featured a single camera with variable focus, zoom, and aperture. The position and orientation of the camera were also under computer control. The objective of Tarabanis' work was to determine the position, orientation, and lens settings for a camera such that objects in a scene were resolvable to a given specification, in focus, within the field-of-view of the camera, and not occluded.

Tarabanis calibrated the relationships between the focus and zoom motors and the following five image parameters: effective focal length; distance from the back principal point of the lens to the sensor plane; diameter of the lens entrance pupil; location of the entrance pupil

center; and location of the lens's front principal point. The pupil diameter model was also a function of the aperture motor. The relationships were stored in sparse look-up tables and interpolation used to calculate values for intermediate lens settings. To calibrate the effective focal length and the location of the front and back principal points, the lens was removed from the camera and placed in a special, optical-bench setup.

One of the more recent systems built for variable-parameter-camera research was the SPARCS system at SUNY, Stony Brook. This system featured computer-controlled focus and manually adjusted zoom and aperture. The primary objective of the SPARCS system was auto-focus and shape-from-focus research. Surya and Subbarao [46] noted that the relationship between the inverse of the focused distance and the focus-motor setting is “almost linear.” This relationship was subsequently modeled using a first-order polynomial. The model was calibrated for one zoom position.

Wiley[55] at the University of Illinois, Urbana-Champaign, was among the first to present results for calibrating a complex model of perspective projection for a variable-parameter lens. He calibrated a 17-parameter photogrammetric camera model for 14 separate focal-length settings on a manually adjusted zoom lens. Plots of the terms of the camera model versus the focal length (i.e. zoom) marked on the lens body showed a high degree of regularity, and Wiley concluded that they might potentially be modeled as first- and second-degree polynomial functions of the focal-length setting. No models were formulated or calibrated though.

Lavest et al. [29] calibrated a simple perspective-projection camera model for 11 zoom positions. The objective for the work was 3D reconstruction from zooming. Experimental results were presented for the reconstruction of three objects. No parametric models relating the lens's zoom to the perspective-projection camera model's terms were presented, and we assume that the reconstruction of the objects was performed using the tabulated model parameters from the 11 calibrated zoom positions. To determine image center Lavest used the focus-of-expansion technique presented in [30]. Measurements of their system showed little variation in the position of image center, therefore it was modeled as a constant in their perspective-projection camera model. This model did not include terms for geometric distortion.

In work very similar to Lavest's, Rodin[40] used another simple perspective-projection camera model to perform 3D reconstruction from zooming. Rodin calibrated his camera model for four (apparently manually set) zoom positions. Experimental results were presented for the reconstruction of several objects. We assume Rodin, like Lavest, reconstructed objects by using the tabulated model parameters from the four calibrated zoom positions. In the perspective projection camera model used by Rodin, the image center was calibrated at each zoom position, but there was no explicit modeling of lens distortion. Both Lavest and Rodin comment that the low accuracy of their results may have been due in part to unmodeled lens distortion effects.

Pahlavan's work [37] with the KTH-head at Stockholm University was one of the more recent efforts to build predictive models for more than one lens parameter. The KTH-head was developed to study active vision from a biological perspective. It featured a pair of

cameras with automated zoom lenses mounted in a highly mobile camera jig. Among other motions the jig allowed the cameras to be independently panned and tilted. Normally as the parameters of a zoom lens are varied the optical center (the center of perspective projection) for the lens moves back and forth along the optical axis. In the KTH-head a translational stage under each camera allowed the lens's optical centers to be re-positioned directly over the rotational centers of the panning motors, more closely emulating the geometry of biological vision systems and "simplifying the task of seeing." To model the necessary scalar camera displacement for any given lens setting, Pahlavan measured the position for the optical center of the lens for a  $32 \times 32$  array of focus and zoom positions. The data was then fit with a six-coefficient polynomial. The model was third-order in the focus motor and first-order in the zoom motor. The method by which the formulation was determined was not discussed.

## The need for this research

To date few predictive models of the relationships between image parameters and the control parameters of adjustable lenses have been implemented. With the exception of Sobel's and Pahlavan's work the models we reviewed have typically been for scalar image properties in only one control variable. Furthermore, these predictive models have been either implicitly or explicitly based on the premise that the imaging behavior of the lens can be completely described by the pinhole, thin-lens, or thick-lens abstractions. Lens behaviors that cannot be explained by these abstractions, such as changes in geometric distortion or the shifting of the camera's field-of-view with changing lens parameters, have not been modeled.

The control of optical parameters is becoming more important in active vision systems. While present day automated camera systems typically have very good control capability, the lack of adjustable camera models prevents this capability from being fully utilized. To be able to effectively use these systems we need to be able to build accurate camera models that hold calibration across continuous ranges of one, two, or even three lens parameters. To do this we must understand the problems and issues involved in the modeling and calibrating real automated lenses. Moreover, we must have a comprehensive and systematic approach to building adjustable camera models for simple to very complex imaging properties.

Previous modeling and calibration for automated lenses has been based on the premise that lens behavior can be inferred from pinhole, thin-lens, or thick-lens camera models. This thesis shows that these approaches do not sufficiently capture the complex relationships between the control parameters and the imaging process. Moreover, we demonstrate how the idealized behavior assumed with these models can lead to significant performance problems in two machine-vision tasks: color imaging and focus ranging. We subsequently discuss how these performance problems can be reduced or eliminated with more complex models and lens control strategies. An outgrowth of our examination of the behavior of real lenses in both fixed- and variable-parameter systems is the comprehensive taxonomy for the property of "image center" we developed.

Our principal contribution is a methodology for producing accurate adjustable camera models for automated camera systems. This approach involves first calibrating a conventional

fixed camera model at a number of lens settings spanning the lens's control space. We then model how the terms of the fixed camera model vary with different lens settings by alternately fitting polynomials to individual model terms and reestimating the unfitted terms using the calibration data. This process is repeated until all of the fixed camera model's terms have been replaced with polynomial functions of the lens control parameters. The result is a predictive camera model that can interpolate between the original sampled lens settings to produce — for any lens setting — a set of values for the terms in the fixed camera model. We have used this approach to produce an adjustable, perspective-projection camera model based on Tsai's fixed camera model and calibrated and tested our model on two different, automated camera systems. In both cases the calibrated models operated across continuous ranges of focus and zoom with an average error of less than 0.14 pixels between the predicted and the measured positions of features in the image plane. We also calibrated and tested our model on one system across a continuous range of aperture and achieved similar results.

In summary then, this thesis provides a methodology for producing adjustable-lens camera models from fixed-lens camera models.

## 1.5 Hardware overview

The results presented in this thesis are largely based on experimental work conducted with two automated camera systems designed and built at Carnegie Mellon's Calibrated Imaging Laboratory (CIL). The first camera system, shown in Fig. 1.4, consists of a Cosmicar 12.5-75 mm ( $6\times$ ) zoom lens mounted on a Panasonic video camera connected to a Matrox frame grabber. The automation for this lens is provided by digital microstepping motors, which are connected to the lens body by backlash-free pushrod assemblies. The Cosmicar lens has 3900 steps of resolution for focus, 4000 steps for zoom, and 2700 steps for aperture.

The second camera system, shown in Fig. 1.5, consists of Fujinon 10-130 mm ( $13\times$ ) zoom lens mounted on a Photometrics scientific camera. The automation for this lens is also provided by digital microstepping motors. The Fujinon lens has 5200 steps of resolution for focus, 11100 steps for zoom, and 2700 steps for aperture. The Fujinon/Photometrics camera system also has a built-in seven slot computer-controlled filter wheel.

A more detailed description of both camera systems and the CIL laboratory is contained in Appendix A.



Figure 1.4: Cosmicar lens and Panasonic camera

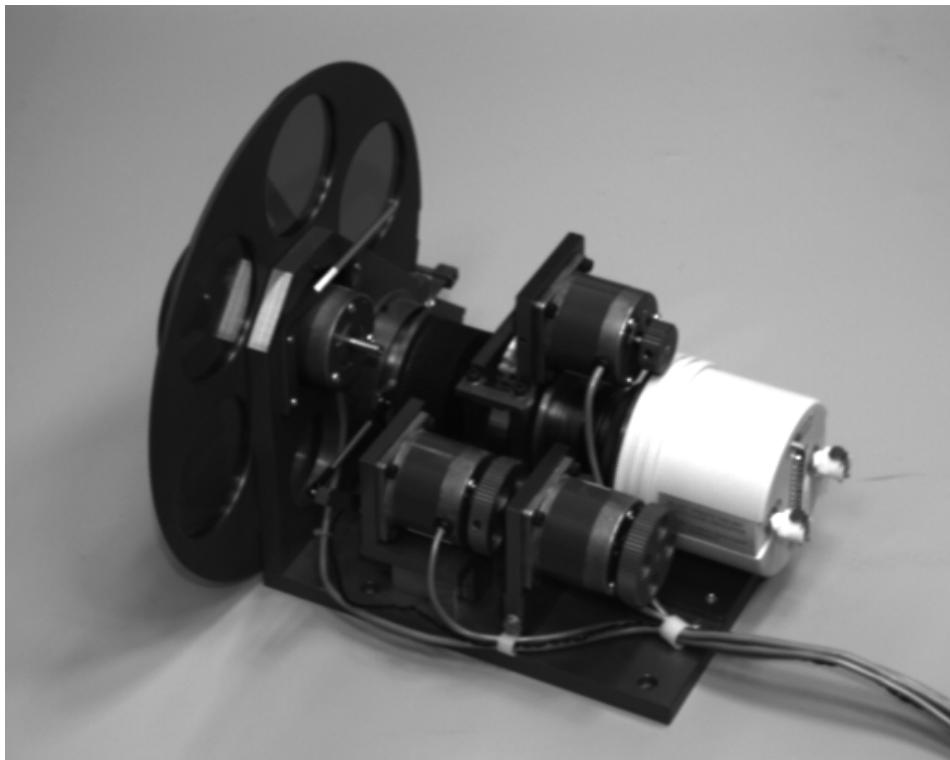


Figure 1.5: Fujinon lens and Photometrics camera

## 1.6 Thesis organization

The remainder of the thesis is structured as follows:

Chapter 2 illustrates how first-order (i.e. idealized) models of variable-parameter lens behavior often fail to capture significant second-order effects (e.g. chromatic aberration and focus magnification). We then show how these effects can cause significant problems in vision tasks and how more realistic modeling of the automated lens’s behavior and more sophisticated control strategies can significantly improve performance.

In Chapter 3 we use the property of “image center” as an example to discuss the differences between the ideal image-formation process and the behavior of real lenses and the serious consequences they can have both for fixed- and variable-parameter lenses. We also present a comprehensive taxonomy for image center that we developed as an outgrowth of our examination of the behavior of real lenses.

Chapter 4 explains why modeling and calibrating variable-parameter camera systems is considerably more difficult than fixed-parameter systems. We then present our new, empirical methodology for developing calibrated camera models for automated zoom lenses.

Chapter 5 presents a detailed discussion of the formulation, calibration, and testing of the fixed camera model that we used as the basis for developing the adjustable camera models described in the following chapter.

Chapter 6 discusses how we applied our methodology (described in Chapter 4) to the fixed camera model (presented in Chapter 5) to produce an adjustable, perspective-projection camera model. We calibrated and tested this model on two distinctly different systems for focus and zoom and on one system for aperture.

Chapter 7 summarizes the specific contributions of the research described in the thesis and discusses remaining open questions and areas for future research.

Appendix A contains a detailed description of the facilities and camera systems of Carnegie Mellon’s Calibrated Imaging Lab. Appendix B describes the details of the approach we used to measure chromatic aberration. Appendix C describes the approach we used to measure the center of our autocollimated laser and the centers of the reference points on our calibration targets. Appendix D describes the calibration of relative and photometric aperture models for one of our camera systems.



# Chapter 2

## Active Lens Control for Measurement

Adjustable lenses allow us to measure properties of a scene by noting how the scene's image changes as the lens parameters are varied. Naturally the accuracy of our measurements depends on how well we model the relationships between the camera's image-formation process and the lens parameters.

In this chapter we present two vision tasks, color imaging and focus ranging, which are prone to significant measurement error when first-order approximations of the lens's imaging behavior are used. We then show that more complete models of the imaging behavior allow more sophisticated lens control strategies and as a result improve task performance.

### 2.1 Active Color Imaging

Color-image analysis uses the information contained in three spectral bands to determine properties of the scene being imaged. Implicit in any color-image analysis is the assumption that the per-pixel information in each band corresponds to the same point, region, or volume in object space. As we will demonstrate in the following section, this is not always so.

To simplify color band alignment, virtually all color imaging processes use a single lens. Three basic approaches are used to digitize the bands. The simplest approach, found in most commercially sold color cameras, uses a single sensor that has its pixel array covered with a mosaic of color filters. Spatial interpolation is then used to fill the gaps in each of the image bands.

The second approach, used in color cameras for commercial television, is to split the image into three paths and project each path through bandpass filters onto separate sensors.

The third approach, used in still color imaging, is to place three bandpass filters in the optical path of the camera system, one after another, and take the three images with the

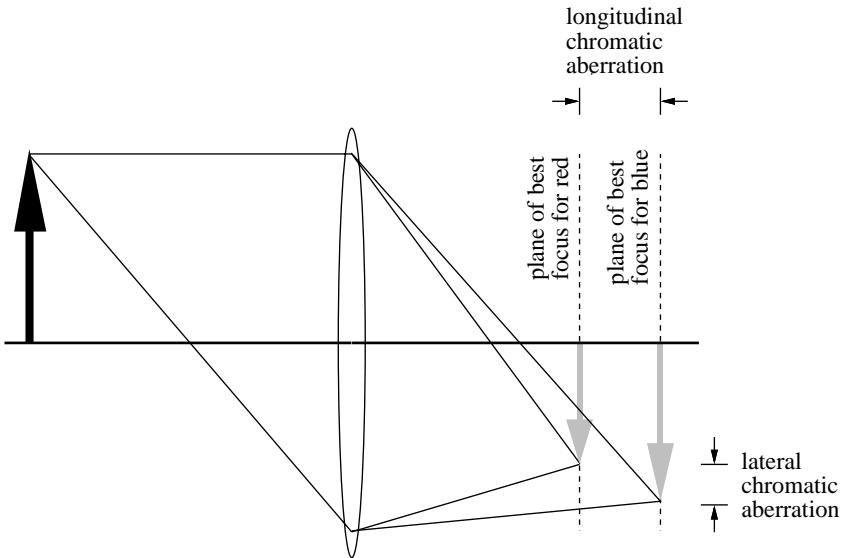


Figure 2.1: Chromatic aberration in a thin lens

same sensor. This is the approach we use in the Calibrated Imaging Lab. It is commonly used for machine-vision research in color physics because it allows superior registration between images and an improved SNR for each band.

For the experiments described in this chapter we used the Cosmicar/Panasonic camera system, a computer-controlled filter wheel, and a computer-controlled six-degree-of-freedom camera platform. The filter wheel houses six filters including, Wratten #25 (red), 58 (green), and 47B (blue). For a more complete description of the equipment see Appendix A.

### 2.1.1 Chromatic aberration

Chromatic aberration exists in camera lenses because the index of refraction of optical components varies as a function of wavelength. This difference in refractive index causes the different wavelengths of light to be refracted or bent to different degrees by the elements of the lens. For example, given a simple uncorrected thin lens with incident white light rays, the blue components of these rays will be brought to a focus closer to the lens than the red components (see Fig. 2.1). Funt and Ho [19] make use of this property for extracting spectral information to address the problem of color constancy.

Chromatic aberration is an intrinsic property of a camera lens. It can be partially compensated for in the lens optics by using pairs of lens elements with offsetting dispersion factors. Unfortunately such compensation is usually done for only two wavelengths (red and blue) and then only at two points in the image field: at the optical axis and at some specified radial distance from the optical axis [43]. For a more detailed discussion of chromatic aberration in lenses the reader is referred to [42] or [43].

To measure chromatic aberration we used black-on-white checkerboard targets like the one

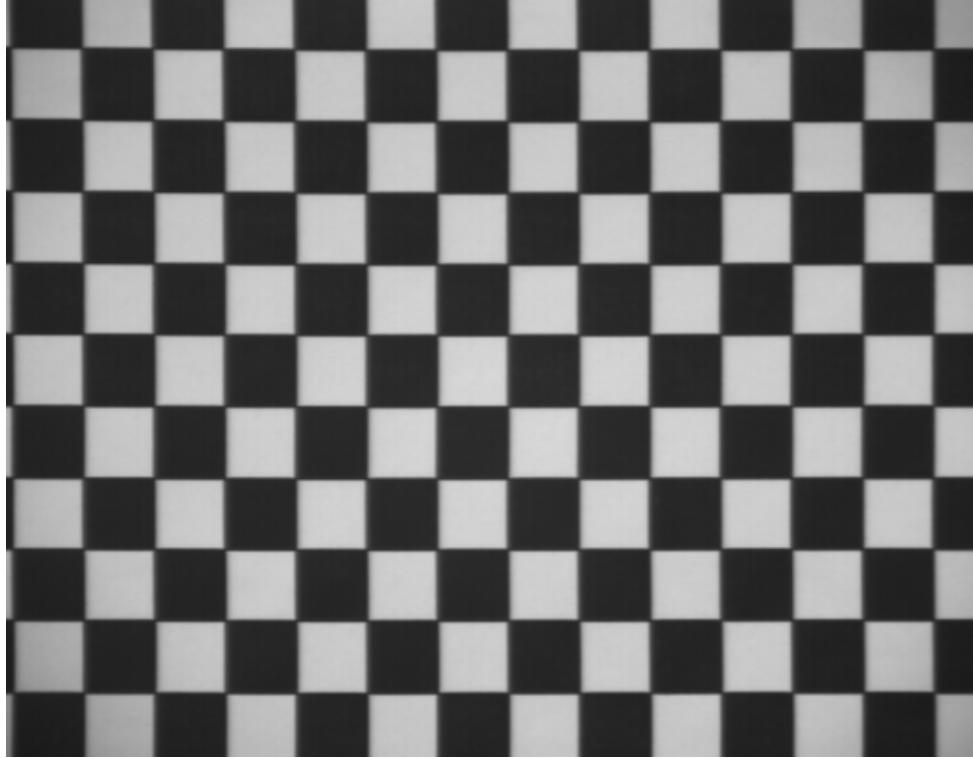


Figure 2.2: Checkerboard test target used to measure lateral chromatic aberration

shown in Fig. 2.2. The relative positions of the vertical and horizontal black to white step edges, measured in each of the bands, provide a measure of the amount of lateral chromatic aberration in the camera system. Appendix B describes the measurement procedure in detail and presents the measurement results for 12 different lenses.

## Effects of chromatic aberration on images

Chromatic aberration has three effects on color images. It causes differences in magnification, focus, and centering between the bands. The first two of these effects are illustrated in Fig. 2.1: the scale difference (i.e. magnification) between images, also called lateral chromatic aberration, and the difference in focus plane position, also called longitudinal chromatic aberration.

The effects of chromatic aberration can most readily be observed in color images of scenes containing sharp black to white transitions, such as the grid of black lines on a white background shown in Fig. 2.3. Figure 2.4 is a graph of the red, green, and blue pixel values for the scanline in section 1 of the grid image. The vertical black grid line that cuts across section 1 is near the optical axis of the lens and shows no noticeable misregistration between the red, green, and blue color bands. The graph in Figure 2.5 shows the pixels in section 2 of the same scanline. The vertical black line being examined in this figure is relatively far from the optical axis of the lens and clearly shows the misregistration of the three color bands due

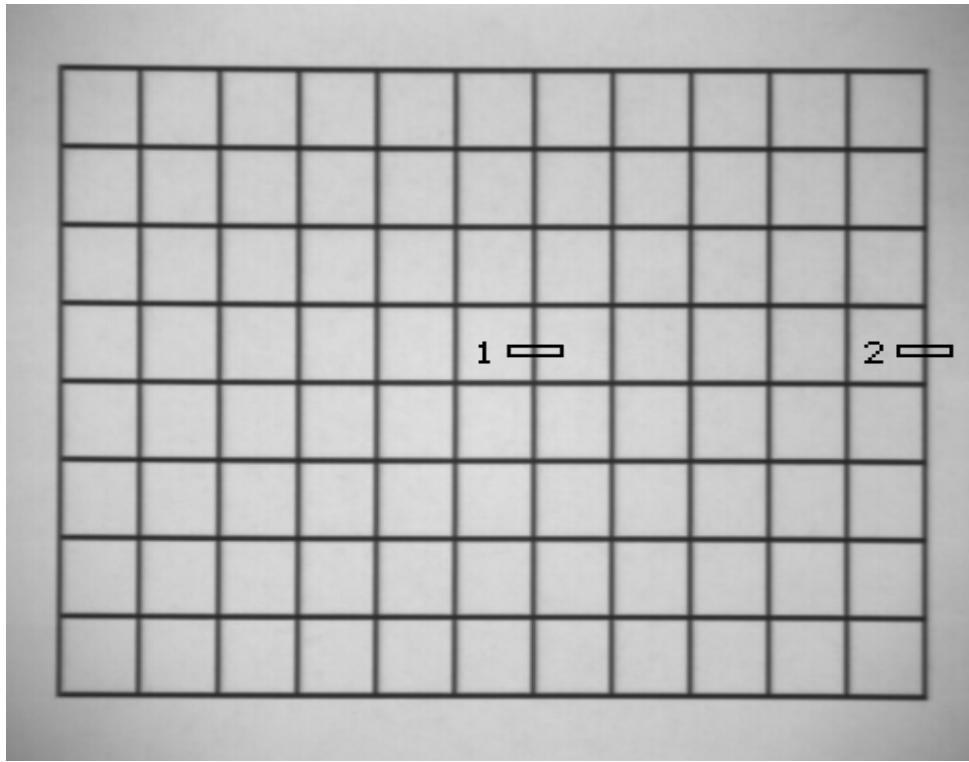


Figure 2.3: Image of grid of black lines on a white background

to lateral chromatic aberration. The slightly broader width of the red valley in both figures shows the defocus in the red image (when lens is initially focused with blue light). With this camera system and at this lens setting the difference in the magnification between the uncorrected red and blue images is on the order of 0.5%, or 1.2 pixels near the outer edges of images that are 512 pixels wide (i.e. a 1.2 pixel displacement in one direction at the left edge and a 1.2 pixel displacement in the other direction at the right edge).

While magnification and defocus result from intrinsic properties of the optical elements, the third chromatic effect, decentering between the image bands, is the result of misalignment in the optical elements of the lens. Misalignment causes light rays of different wavelengths to take slightly different optical paths through the lens. Furthermore, the decentering effect is compounded by optical “rollout.” Changing the lens’s focus and zoom settings changes the degree of misalignment and causes a drift in the image center. Whereas the magnification differences are radially symmetric with respect to the center of the optical system and have a smaller magnitude nearer the center than at the edge, decentering differences affect the entire image uniformly. In our system the measured translation between the uncorrected red, green, and blue bands ranges from 0.1 to 0.4 pixel widths, depending upon the image band and the lens settings.

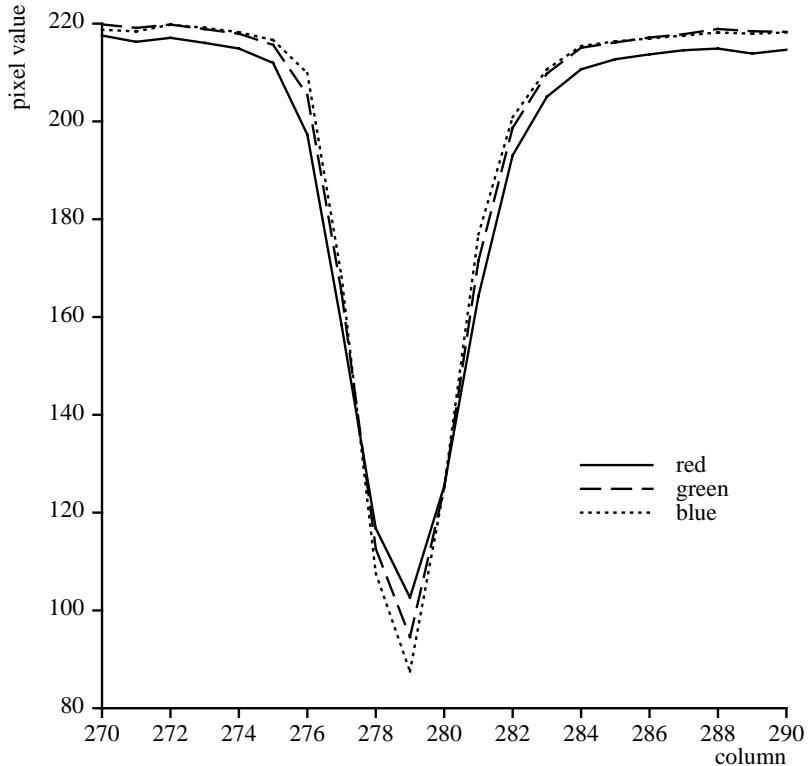


Figure 2.4: Pixel intensity profiles near center of image (cross section 1)

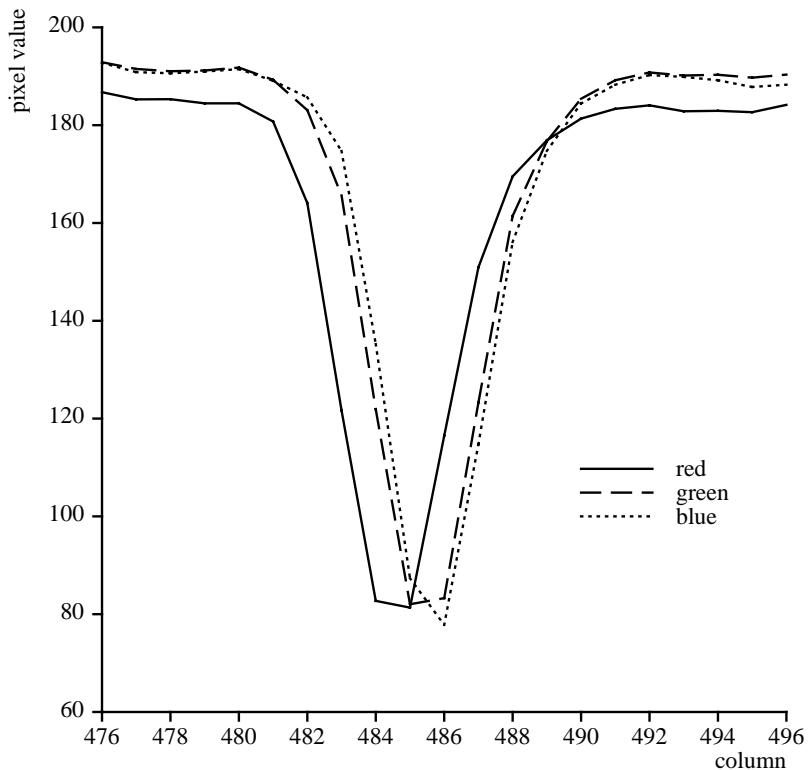


Figure 2.5: Pixel intensity profiles near edge of image (cross section 2)

## Effects on color analysis

The result of the magnification, focus, and centering differences between the color bands is that the three values for each pixel in the color image do not correspond to the same point in object space. This has serious implications for color image analysis. For example, regions close to rapid color changes (e.g. near highlights or the edges of features) will show hue shifts. Looking closely at the pixels in section 2 of Fig. 2.3 (shown in profile in Fig. 2.5), going left to right, we see the pixel colors change from white, to bluish, to black, to reddish, and back to white. As we will show later, these hue shifts show up as erroneous points in color histograms of the image, which in turn can lead to problems in the segmentation and analysis of color images [26][36]. Edge detection and edge localization in color images are also confounded in obvious ways.

### 2.1.2 Compensation approach

The most direct approach to compensating for chromatic aberration is by modifying the focus and zoom for each of the color bands to null out the defocus and magnification differences. Correction for image decentering can be accomplished by shifting the camera in the  $XY$  (i.e. sensor) plane. This is the compensation approach we use and call Active Color Imaging (ACI).

We determine the lens settings for the compensated images by first placing a black-on-white calibration target (shown in Fig. 2.6) directly in front of the scene being imaged. The black-to-white edges on the target are used to focus on and to determine the relative image magnification and shifting between the color bands. Once the correct focus, zoom, and shift settings for the lens have been determined, the calibration target is removed and a compensated color image of the scene taken.

In our Active Color Imaging approach the blue image is used as “ground truth.” The lens settings for the red and the green bands are determined by minimizing the errors between the blue and red images and the blue and green images. To determine the compensated lens settings for the red and green images we first find the best focus position for each of the three bands by using the automated focusing algorithm developed by Krotkov [27]. Next we take three focus-corrected images and determine the magnification differences between the blue and red and the blue and green bands. Using a calibrated lens model that relates the image magnification to the positions of the focus and zoom motors we then change the zoom for the red and the green bands to correct the magnification differences that result from chromatic aberration and from refocusing. Finally, to correct the color-dependent image translation, we take three focus- and magnification-corrected images and determine the amount of camera shifting that is needed in the red and green bands to compensate for translation introduced by the optics.

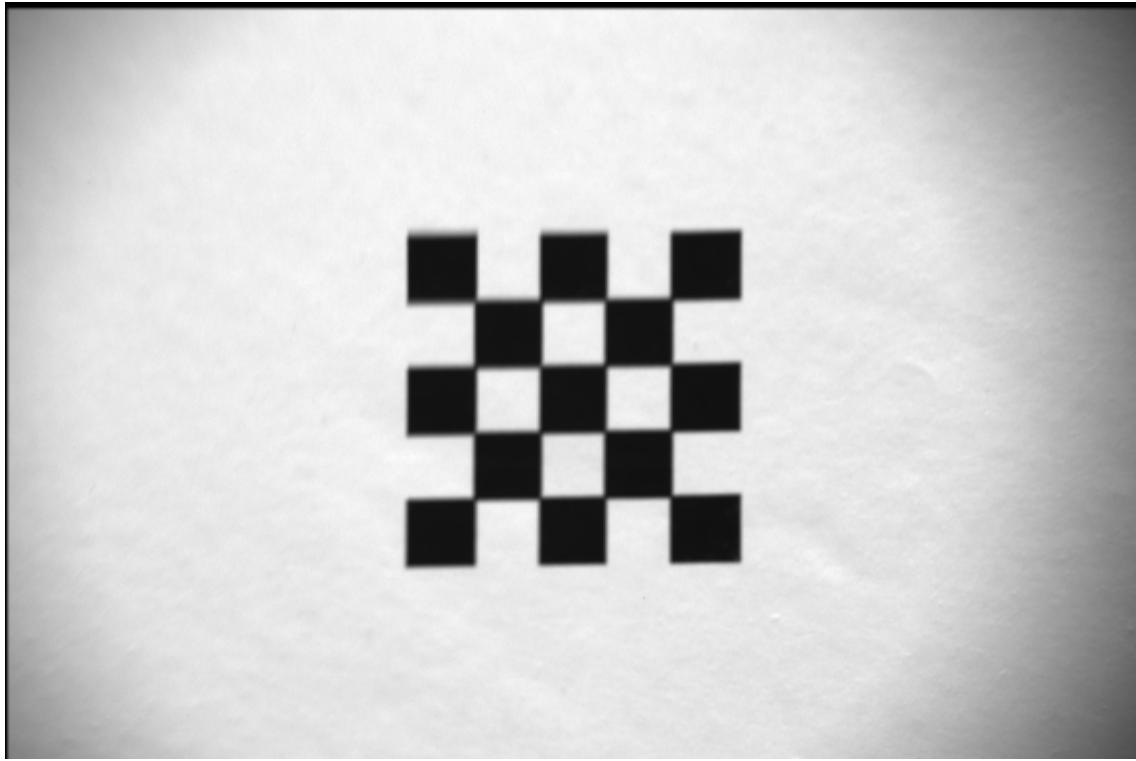


Figure 2.6: Active Color Imaging calibration target

### Calibration example

To demonstrate the steps in the Active Color Imaging approach we use the test target in Fig. 2.7. By examining the relative positions of the 12 edges in a scanline crossing the center of the target we can graphically show the effects of each of the steps of the compensation procedure. Plotting the difference in the edge positions (in pixel widths) between the uncorrected blue and red images versus the edge number we obtain the solid line in Fig. 2.8. (For a lens without chromatic aberration the line would be zero everywhere, indicating the same edge positions in the blue and red images.) The plot in Fig. 2.8 shows two effects: the relative magnification of the blue and red images indicated by the slope of the blue-red line, and the  $X$  component of the image translation indicated by the displacement of the line up or down. In this example the magnification difference between the uncompensated blue and red images is -0.45% and between the blue and green images is -0.14%<sup>1</sup>. The  $X$  shift between the blue and red images is 0.12 pixels and between the blue and green images is 0.08 pixels.

The dotted line shows the blue-red differences after refocusing the red image. The magnification difference has now shifted in the other direction as a result of the magnification change introduced by refocusing. After focus compensation the magnification difference is 1.01% between the blue and red images and 0.35% between the blue and green images. The  $X$  shift is 0.35 pixels between the blue and red images and 0.14 pixels between the blue and

---

<sup>1</sup>For clarity the blue-green plot is not shown on this graph.

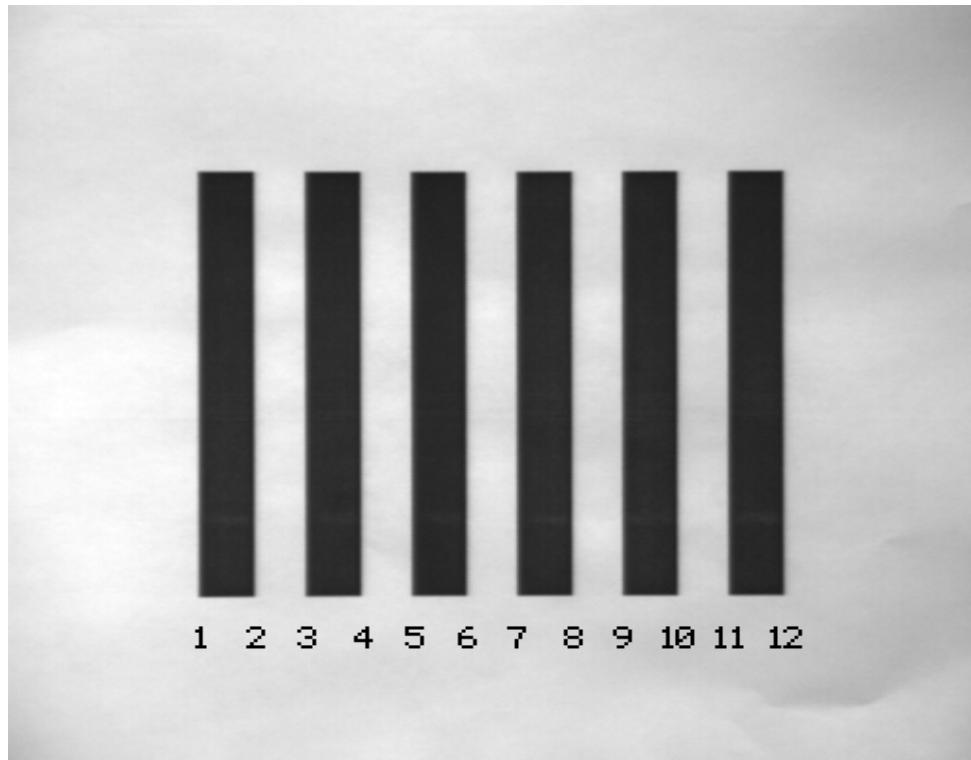


Figure 2.7: Active Color Imaging demonstration target

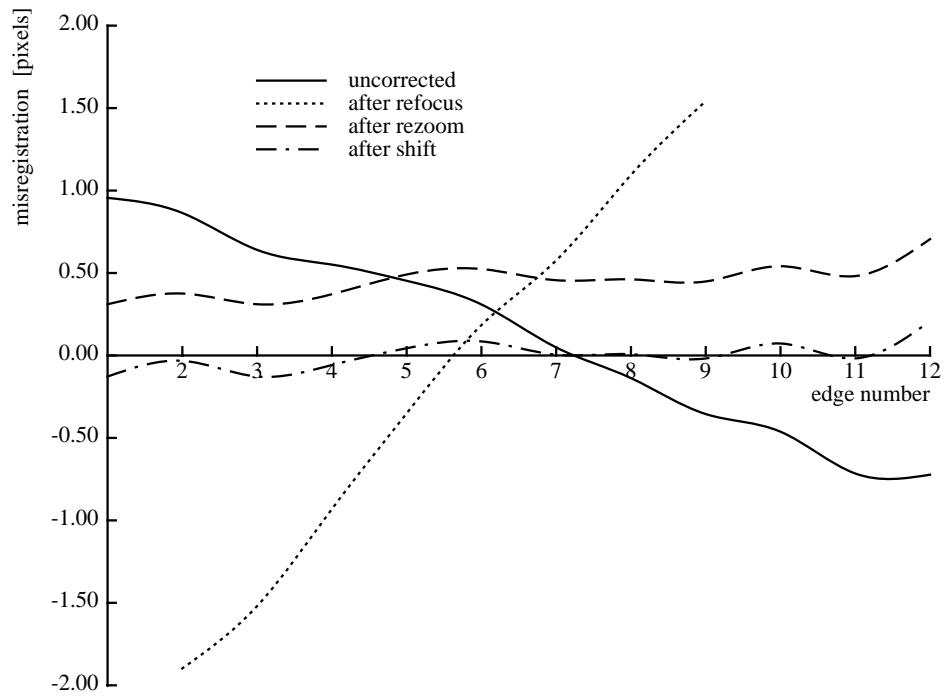


Figure 2.8: Active Color Imaging results

Table 2.1: Typical camera and lens settings for corrected color images

color filter	focus motor (out of 3900)	difference from blue setting	zoom motor (out of 4000)	difference from blue setting	camera <i>X</i> shift
red	1407	-6.36%	533	+0.83%	0.15 mm
green	1592	-1.62%	508	+0.20%	0.05 mm
blue	1655		500		

green images.

The dashed line shows the blue-red differences after re-zooming the red image. While the slope of the line is approximately zero, indicating that the magnification differences between the bands have been eliminated, the image translation is now obvious. After focus and zoom compensation the magnification difference is 0.07% between the blue and red images and 0.01% between the blue and green images. The *X* shift is 0.46 pixels between the blue and red images and 0.12 pixels between the green and blue images.

The dash-dot line shows the final result after the camera has been shifted. The plot is now close to zero everywhere, emulating ideal lens behavior. At the edge of the image the misregistration has dropped from approximately one pixel to approximately one-tenth of a pixel. The final magnification difference is 0.05% between the blue and red images and -0.01% between the blue and green images is -0.01%. The final *X* shift is 0.00 pixels between the blue and red images and 0.01 pixels between the blue and green images.

Active Color Imaging requires very fine control of the lens parameters. For this example the lens and camera parameters were changed by the amounts shown in Table 2.1. The image was initially composed with the zoom motor set to 500 units (focal length  $\approx$  67 mm). At the start of the ACI procedure the lens was focused with the blue, green, and red filters. The final focus motor positions are listed in column 2, with the percentage deviation from the blue settings listed in column 3. Using the refocused images, the compensating zoom motor settings were calculated for the green and red images. The final zoom settings are listed in column 4 and the percentage deviation from the blue settings in column 5. Finally, using the focus- and zoom-compensated images, the required *X* camera shift was calculated for the green and red images. These values are shown in column 6. Shifting in the *Y* direction was not performed for this example.

### 2.1.3 Experimental results

Figures 2.9 and 2.10 show the magnitude of the image misregistration (in pixel widths) between the blue and red images coded in a gray scale ranging from 0 pixels (black) to 1.4 pixels (white). Figure 2.9 shows the blue-red misregistration for the target imaged without lens compensation. The misregistration across the image ranges from 0 to 1.2 pixel widths. For the blue-red case the zero error region is slightly to the right of the image center. In general the location of the zero error region for the blue-green case will not be the same.

Figure 2.10 shows the magnitude of the blue-red misregistration for the target imaged with lens compensation. The remaining misregistration is now less than 0.1 pixel widths over most of the image. This result indicates that with active lens compensation it is possible to reduce the image band misregistration introduced by chromatic aberration by over an order of magnitude.

Another way of showing the effects of active lens compensation is to plot image pixels in RGB space. Figures 2.11 and 2.12 contain color histograms of the region containing the rightmost vertical line of color image of a black grid on a white background shown in Fig. 2.3. Ideally all pixels in the grid image should be black, white, or shades of gray. The plotted pixels from the region containing the grid line should form a straight line of grays between the lower left corner of the color cube (corresponding to black) and the upper right corner of the cube (corresponding to white). This is not the case in the uncorrected image where pixels show two distinct paths from black to white (see Fig. 2.11). The first path bends towards the red corner of the color cube while the second path bends to the blue corner. These deviations would pose a serious problem for any algorithm trying to segment the image by using regions in the color space. In Fig. 2.12 the histogram of the pixels from the image taken with active lens compensation shows a much tighter grouping of pixel values along the gray line.

### 2.1.4 Conclusion

The lenses typically used in computer vision and the processes employed for taking color images will generally produce images that have magnification, focus- and image-center differences between the color bands. We have measured comparable levels of chromatic aberration in a wide variety of lenses, including CCTV lenses, 35mm SLR lenses and ENG/EFP color TV lenses, including lenses that are advertised as being corrected to eliminate chromatic aberration (See Appendix B for results). These effects are due to both chromatic aberration and optical misalignment in the camera lens and result in significant levels of misregistration and defocus in the final composite image. With active compensation of the lens settings and the camera position, misregistration between the color bands of an image can be reduced by an order of magnitude.

Chromatic aberration effects can also be removed by post processing the image bands. Boult [7] compares our ACI approach to their image warping method. Image warping has the advantages that it can be applied to cameras that take RGB images in parallel (e.g. color video cameras), does not require special lens-control hardware, and can handle chromatically-varying geometric distortions that cannot be corrected by focusing or zooming. On the other hand, the image warping approach cannot compensate for defocus between image bands and usually requires a relatively large amount of calibration data. Boult found that while image warping compared reasonably well, both quantitatively and qualitatively with our approach, our approach produces better overall results for uncorrected camera lenses because of its ability to correct for focus differences between bands.

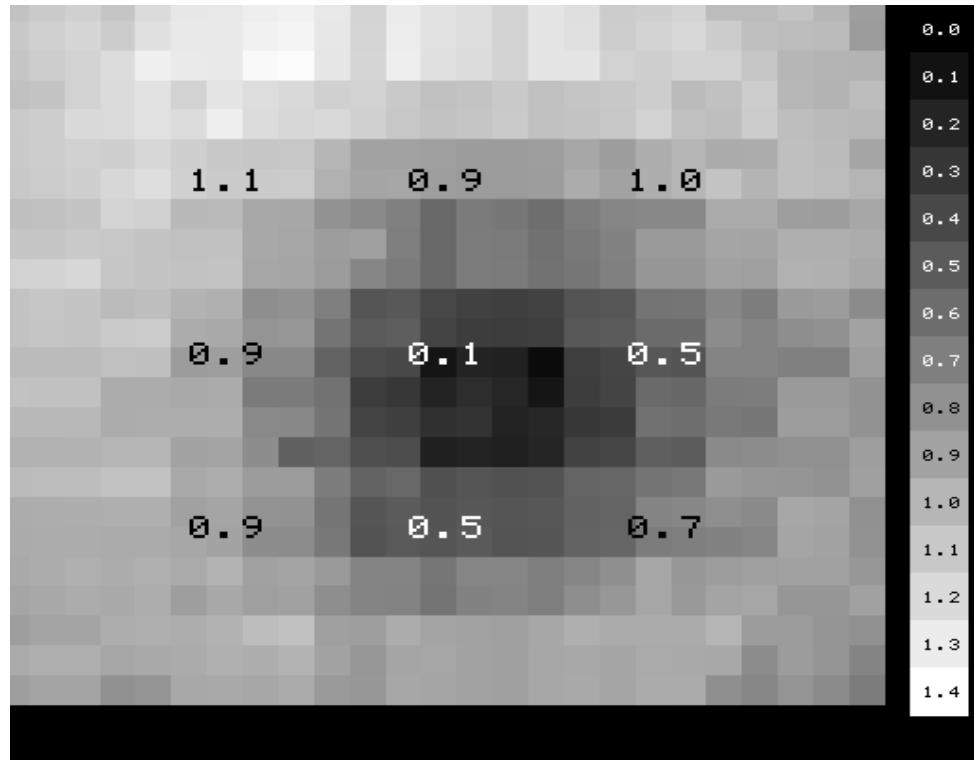


Figure 2.9: Full field blue/red misregistration — uncompensated image

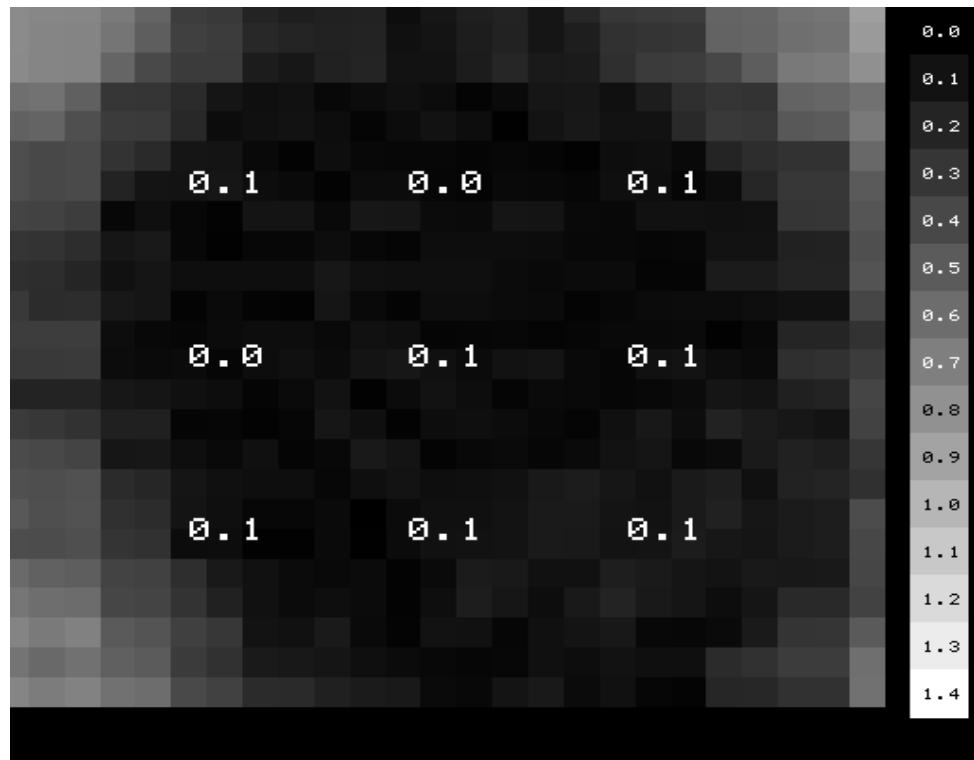


Figure 2.10: Full field blue/red misregistration — compensated image

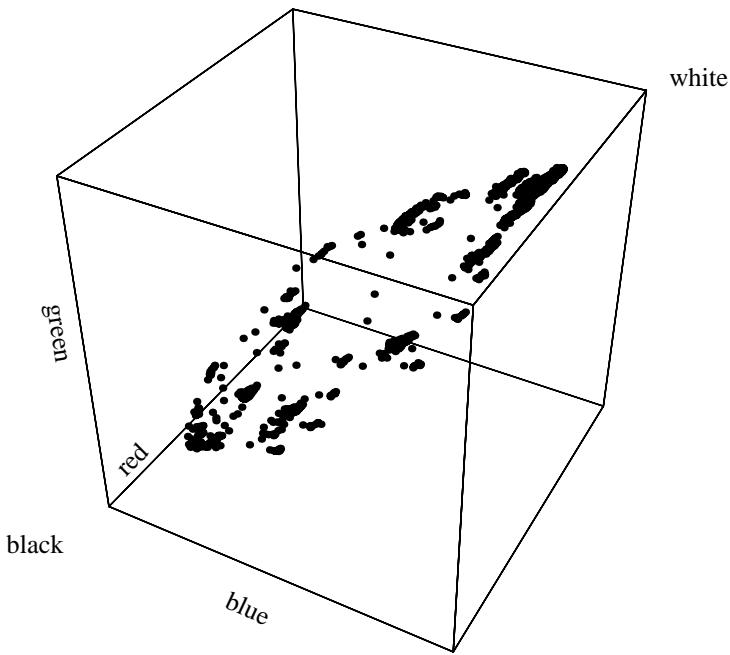


Figure 2.11: Color histogram — uncompensated image

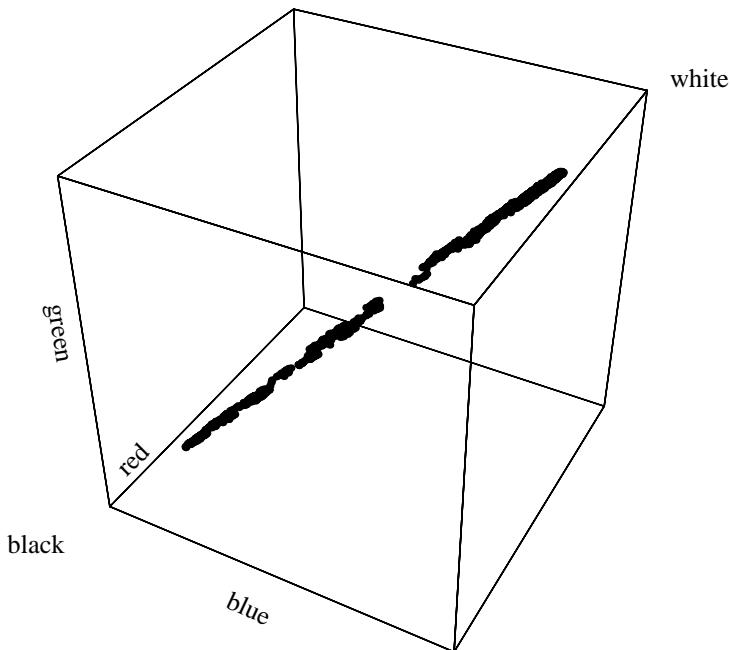


Figure 2.12: Color histogram — compensated image

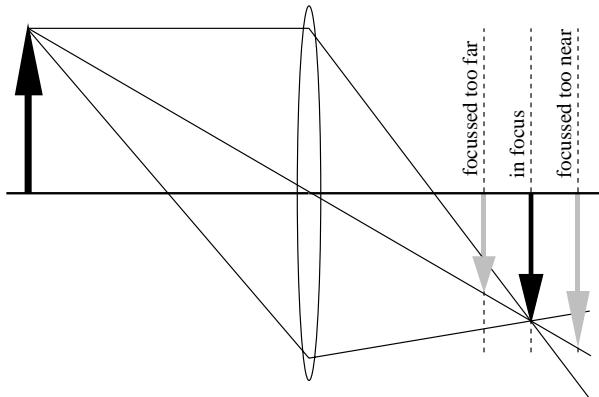


Figure 2.13: Focus magnification in a thin lens

## 2.2 Constant Magnification Focusing

In range-from-focus the distance to a target is estimated by moving the camera lens through a series of focus positions, while the sharpness of the target’s image at each position is determined by evaluating a sharpness criterion function. Given the focus position of the lens where the criterion function peaks, the distance between the camera and the target can be determined using a calibrated lens model. Implicit in the range-from-focus procedure is the assumption that the position of the criterion function’s peak is dependent only on the distance between the target and the camera. In the following section we demonstrate how, for the most widely used class of criterion functions, the position of the peak is also dependent on the lens’s aperture and on the content of the image.

Many sharpness criterion functions have been suggested for focus ranging [27]. In our experiments we make use of a function based on the sum of squared image gradients, also called the Tenengrad function. Most of our discussion transcends a specific choice of criterion function.

### 2.2.1 Focus magnification

For a fixed focal length lens, focus magnification is the change in image magnification that results as the camera’s sensing plane is moved along the optical axis to vary the lens’s focused distance. This process is illustrated for a simple thin lens in Fig. 2.13. Focus magnification causes the image to scale up as the lens is focused from far to near. For zoom lenses focus magnification is conceptually the same.

If one considers focus magnification the net magnification of a zoom lens is actually a function of both the focus motor and the zoom motor. By measuring the dimensions of a target over a range of focus and zoom settings ( $m_f$  and  $m_z$ ) we can produce a calibrated lens model for the relative image magnification  $M = g(m_f, m_z)$ . Figure 2.14 shows a plot of this model.

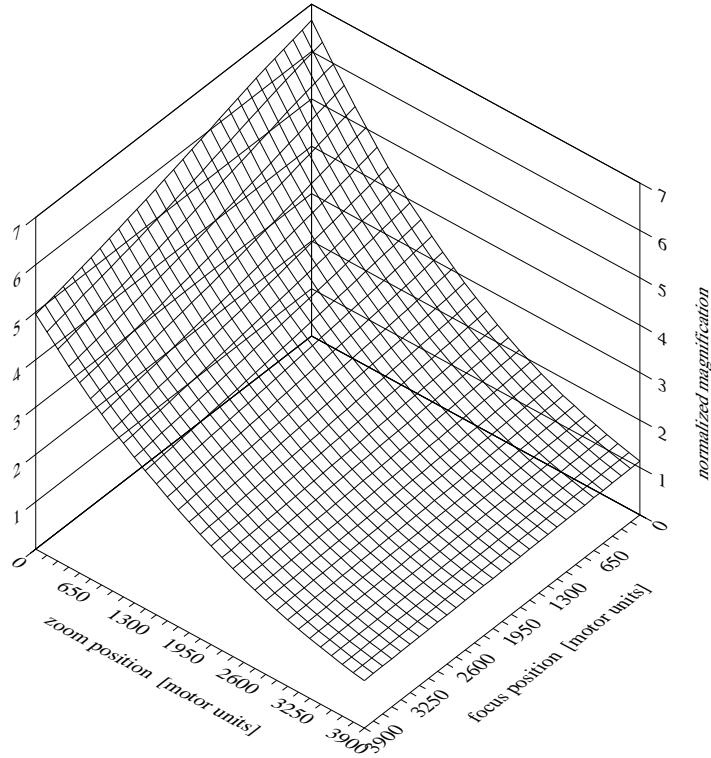


Figure 2.14: Simple image magnification versus focus and zoom motors

### Effects on the image

Focus magnification causes three effects in range-from-focus: feature clipping, perimeter scaling, and gradient scaling. Feature clipping, illustrated in Fig. 2.15, involves the movement of features into or out of the window of interest as the image magnification changes. This effect is pretty straightforward and has been noted by several researchers, including Krotkov [27], Darrell [17], Nair [33], and Nayar [35]. Perimeter scaling, illustrated in Fig. 2.16, involves the proportional scaling of the length of feature perimeters as image magnification is changed. Essentially, perimeter scaling encompasses image scaling effects in directions *perpendicular* to intensity gradients in the image. Gradient scaling, illustrated in Fig. 2.17, involves the inverse scaling of the width and slope of intensity gradients and encompasses image-scaling effects in directions *parallel* to the direction of the image's intensity gradients.

### Effects on focus ranging

For accurate focus ranging it would be convenient if the position of the criterion function peak would depend solely on the focused distance of the lens. As we will demonstrate, focus magnification can make the position of the peak sensitive to additional factors.

Feature clipping causes abrupt changes in the criterion function value as features enter or leave the evaluation window. These changes, in turn, may result in multiple criterion

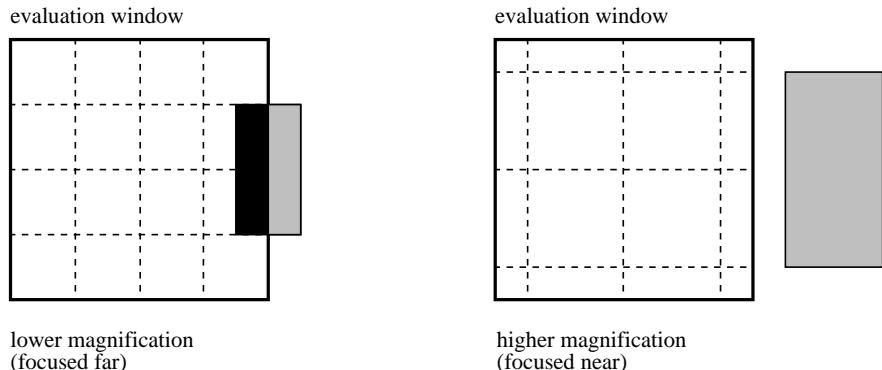


Figure 2.15: Focus magnification causing feature clipping

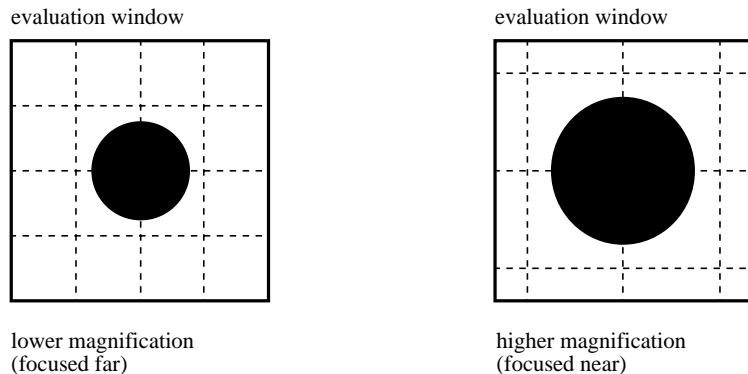


Figure 2.16: Focus magnification causing perimeter scaling

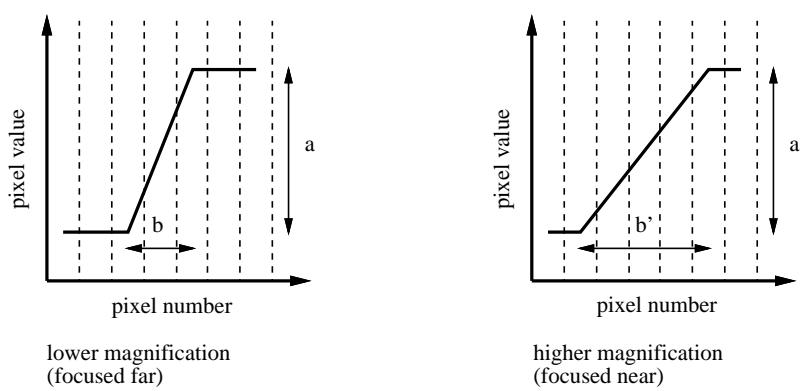


Figure 2.17: Focus magnification causing gradient scaling

function peaks, which in turn can lead to erroneous range estimates. The other two focus magnification effects, perimeter scaling and gradient scaling, cause more subtle biases in the position of the criterion function peak.

With focus magnification, the size of an image scales up as the lens is focused from far to near. As the image scales up feature perimeters will also be proportionately scaled up, causing the value of the criterion function to increase. This increase, in turn, contributes a bias in the position of the criterion function peak, making the point of best focus appear too near. However, as the image scales up with focus magnification, the intensity gradients in the image spread out laterally. In the one-dimensional case, shown in Fig. 2.17, the value of a sum-of-squared gradients criterion function would be

$$\text{value} = \frac{a^2}{b}.$$

Thus, as the image — and  $b$  — are scaled up, the value of the criterion function is proportionately scaled down. This decrease contributes a bias in the position of the criterion function peak, making the point of best focus appear too far away.

Both the perimeter and gradient scaling effects are functions of image magnification and the rate of blurring as the lens is moved away from the point of best focus. The blur rate is effectively a function of the lens's aperture. At higher blur rates (wider apertures, narrower depths of field) the value of the criterion function falls off too quickly for focus magnification effects to significantly bias the position of the function's peak. At lower blur rates (narrower apertures, wider depths of field) the value of the criterion function falls off more slowly and focus magnification has a more pronounced effect.

The relative contributions of the effects of perimeter scaling and gradient scaling depend on the content of the scene being imaged. Perimeter-scaling effects dominate if the intensity gradients in the image are sharp (e.g. black to white step edges). Gradient-scaling effects dominate if the intensity gradients are gradual (e.g. shadows on curved surfaces).

The net result of the two competing focus magnification effects is an additive bias to the position of the criterion function peak. This bias is dependent on both the lens's aperture and the image content.

### 2.2.2 Compensation approach

To avoid focus magnification effects in focus ranging, researchers have suggested either keeping the image magnification constant [17] or keeping the evaluation window large enough to completely encompass the features being ranged [27]. However, as we have suggested above, simply keeping the features being ranged entirely within the evaluation window will not overcome all of the focus magnification effects. Only a constant image magnification approach can avoid the problems that result from focus magnification.

Darrel [17] suggests two approaches for dealing with focus magnification. The first approach, and the one he implemented, is to cancel out the magnification change by scaling the images

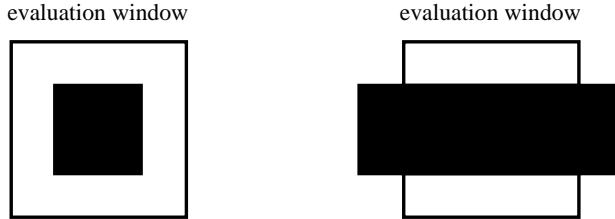


Figure 2.18: Box and bar targets for perimeter scaling experiment

in software after they are taken. This approach is relatively simple and works even for fixed-focal-length lenses. One major drawback of this approach is that, as the image is scaled, the spatial interpolation and resampling process will smooth the image and influence the value of the criterion function. An additional drawback is the computational effort required for scaling the images.

The second approach suggested by Darrell, is to null out the focus magnification by changing the lens's zoom before the images are taken. While it is more efficient and effective than scaling images, this approach requires precise, active lens control as the focus of the lens is varied. In our system we use the current focus and zoom settings, the new focus setting, and our relative magnification model to find a new zoom setting that maintains the same relative magnification. We call this compensation approach "constant magnification focusing."

### 2.2.3 Experimental results

To illustrate the effects of focus magnification on the position of the Tenengrad criterion function peak we use the two targets illustrated in Fig. 2.18. Target 1, completely enclosed in the evaluation window, is a black square on white background. Target 2 is a black bar on white background, and only the center region of the bar is contained in the evaluation window. For both targets the majority of the value of the criterion function results from the black-to-white edge, while the totally white and totally black regions contribute insignificant amounts. For target 1 the length of the perimeter of the black region changes with the focus magnification, while for target 2 the perimeter length remains constant. For both targets the width of the intensity gradient at the black to white edge is not large enough to produce significant gradient scaling effects. The box and bar targets are both located on the same plane, 1.5 m away from the camera's sensor plane.

For the sharpness criterion function in these experiments we use a sum of squared image gradients. The image gradients are calculated using a  $3 \times 3$  Sobel operator. The position of the criterion function peak is determined by fitting the criterion function values to a quadratic curve and solving for the curve's peak. To highlight any bias in the position of the peak we use a narrow aperture.

For the conventional focus-ranging approach the focus motor is varied while the zoom motor is held constant at 500 motor units. In this example the image magnification over the range of focus changes by a factor of 0.751.

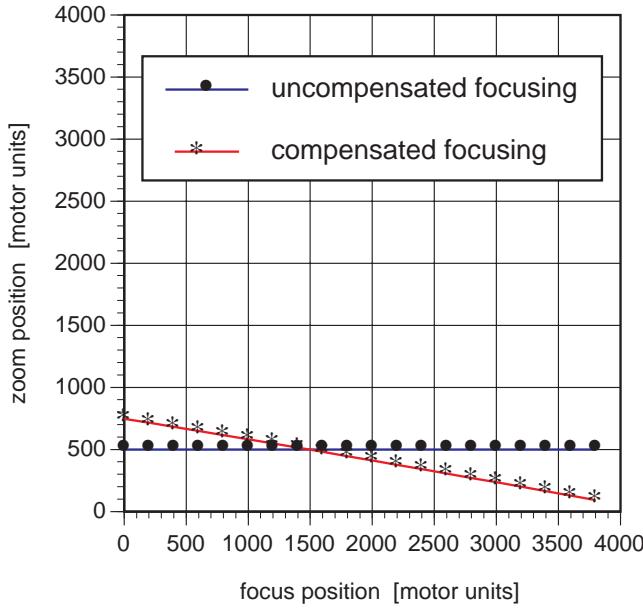


Figure 2.19: Focus versus zoom settings for both compensated and uncompensated focusing

For our Constant Magnification Focusing approach the focus and zoom motors are concurrently varied to keep the magnification constant. Figure 2.19 shows the zoom and focus settings for both compensated and uncompensated focusing. In this example the zoom motor varies from 747 motor units to 91 motor units. The compensated (i.e. Constant Magnification Focusing) curve is essentially an isomagnification contour from the magnification model shown in Fig. 2.14.

Figure 2.20 shows that with conventional focusing there is a significant bias in the box's criterion function peak position. When the focus magnification is compensated, the bias in the box's peak position is eliminated. The peak position for the bar target is unaffected. For an actual range of 1.5 m the bias in the uncompensated box target's peak position corresponds to a 6% error in the range measurement determined from a calibrated focused-distance model.

To observe the effects of gradient scaling on the position of the Tenengrad criterion function peak we have used targets similar to the two illustrated in Fig. 2.21. Target 1 is a bar whose optical density changes smoothly from white to solid black while target 2 is a solid black bar. To avoid perimeter-scaling effects both bar targets extend beyond the edges of the criterion function evaluation window. In qualitative tests we have observed significant biases in the position of the criterion function peak for the graded bar. We have no quantitative results to report.

## 2.2.4 Conclusion

In range-from-focus tasks, where the lens's focus is varied without magnification compensation, several image-scaling effects can bias the position of the sharpness criterion function's

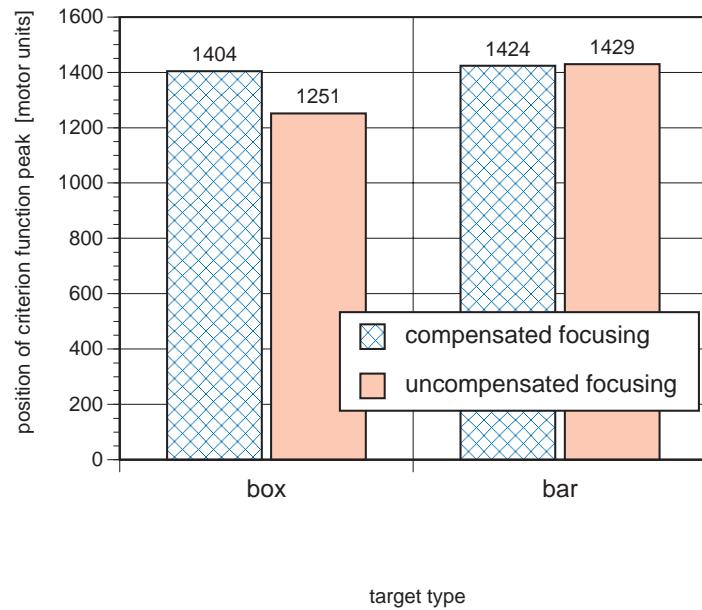


Figure 2.20: Criterion function peak position versus target type and focusing approach

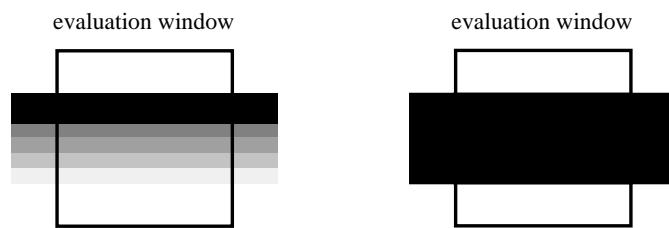


Figure 2.21: Graded bar and solid bar targets for gradient scaling experiment

peak away from the point of best focus. The amount and direction of the bias depends on the size of the aperture and on the content of the image region being evaluated. With active lens compensation, focus-magnification-induced bias problems can be efficiently and effectively eliminated.

## 2.3 Summary

The performance of vision tasks can be adversely affected when we neglect second-order lens behaviors.

In both the conventional color-imaging and focus-ranging tasks there is a first-order assumption that there is a simple relationship between the properties being measured (the color composition of a point in space and the peak position of the sharpness criterion function), and the camera parameter being varied (the color filter and the focus motor). We demonstrated that the relationships are much more complex and that more complete models of the imaging behavior allow better lens control and improve task performance.

We have shown how in color imaging chromatic aberration can cause color-band misregistration and defocus, which in turn can significantly degrade the performance of color-image analysis algorithms. To compensate for chromatic aberration we have developed a new color-imaging procedure — Active Color Imaging — which reduces color misregistration by an order of magnitude. We also developed a simple approach to measuring lateral chromatic aberration across the full field of view of a camera system. Furthermore we demonstrated how focus magnification in focus ranging can cause bias in the position of criterion function peaks, which in turn causes incorrect estimates of range. To compensate for focus magnification we implemented a method (first proposed by Darrel) — Constant Magnification Focusing — to directly eliminate the effects of image scaling.

# Chapter 3

## Image Center

To model the relationship between the positions of features in the object space and their corresponding positions in the image plane we need to know the camera's image center. First-order models of lens behavior, such as the pinhole-camera model or the thin-lens model, suggest that the image center is a single, fixed, and intrinsic parameter of the lens. On closer inspection, however, we find that there are many possible definitions for image center. Most image centers do not have the same coordinates and, moreover, move as lens parameters are changed. In this chapter we present a taxonomy that includes 15 image centers.

Camera calibration in machine vision has traditionally paid little attention to the issue of image center. Typically the image center used to model one imaging property is obtained by measuring a completely different property, if a measurement is made at all. Such approaches can reduce the overall accuracy of the camera calibration. By using the proper image center for each image property that we are trying to model and by calibrating the image centers over the appropriate ranges of lens parameters we can significantly improve the accuracy of our camera models.

### 3.1 Camera calibration and image center

Traditional camera calibration involves modeling the relationship between the positions of features in the object space and their corresponding positions in the image. The simplest model for this relationship is the pinhole camera in which the coordinates of points in the 3D object space are projected through a center of perspective projection down to the 2D image plane. The normal projection of the center of perspective projection down to the image plane is the image center.

Precise camera models have many terms that describe the properties of the imaging process, and some of these terms account for properties that vary with their distance from the center of the image. To model such properties we need to know where their image center is. Naturally, the accuracy of the model depends on the accuracy of the center.

An ideal lens would have one image center and this center would be used in modeling any of the radially varying imaging properties. In practice the manufacturing tolerances for lenses result in different imaging properties with centers in different places, as shown in Fig. 3.1 for the Fujinon/Photometrics camera system. Thus, image centers are not necessarily interchangeable. Indeed, to fully model a camera we may need several different image centers.

The situation becomes even more complex for an adjustable lens. When camera parameters, such as focus or zoom, are varied, the position of the camera's field of view and image centers will also vary. Figure 3.2 shows how the position of a fixed point at the center of the camera's field of view shifts as a function of the focus and zoom motors of the camera lens.

We start this chapter by examining why different image properties do not necessarily have the same image center in real lens systems. We also discuss why the image centers move in variable-focus and variable-focal-length camera lenses. We then present a taxonomy of image center definitions based on the number of lens settings that are required to determine the image center. Procedures for measuring 15 different image centers are given and experimental results are then presented for ten of these methods. We conclude by examining how image center and field of view move in a variable-parameter zoom lens.

## 3.2 Real lenses

Traditionally a camera's image center is considered to be the point of intersection of the lens's optical axis with the camera's sensing plane. The optical axis is defined as the line passing through the centers of the optical components in the lens. In real lenses the optical axis is not so easily defined. The type of complications that arise depend in part on whether the lens has fixed or variable parameters and on how the variable parameters are mechanically implemented.

In an ideal camera lens the components of the lens are aligned along a single axis, making the lens and all of its imaging properties radially symmetric. In real lenses things are not so easy. For a simple lens element (see Fig. 3.3) there are actually two axes of symmetry, one optical and one mechanical. The optical axis is defined as the straight line joining the centers of curvature of the two lens surfaces. The mechanical axis of the lens is determined during manufacture by the centerline of the machine used to grind the lens's edge. Ideally the optical and mechanical axes coincide, in practice, though, they don't. The tolerance between them is called *decentration* [43].

In a compound lens two or more lens elements are aligned and mounted together to form the complete lens. Ideally all of the elements are aligned along a common optical axis, but this is not always feasible given the decentration in the individual elements. The cumulative effect of the mechanical tolerances for the lens elements is that there is no "ideal" optical axis for the lens. Decentration and misalignment in the lens produce tangential lens distortion and asymmetric, radial lens distortion [11]. Thus the different imaging properties of the lens do not necessarily have a common axis of symmetry.

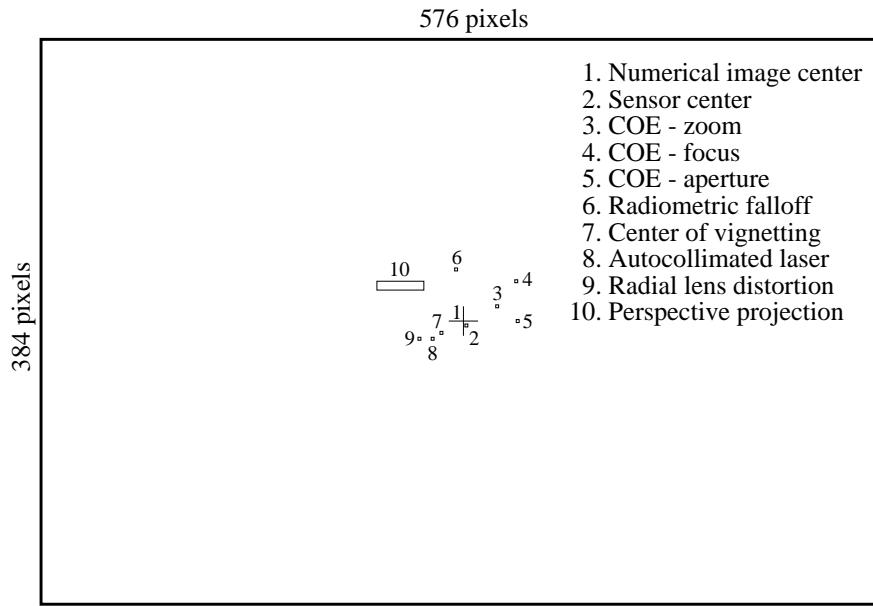


Figure 3.1: Different image centers for the Fujinon/Photometrics camera system

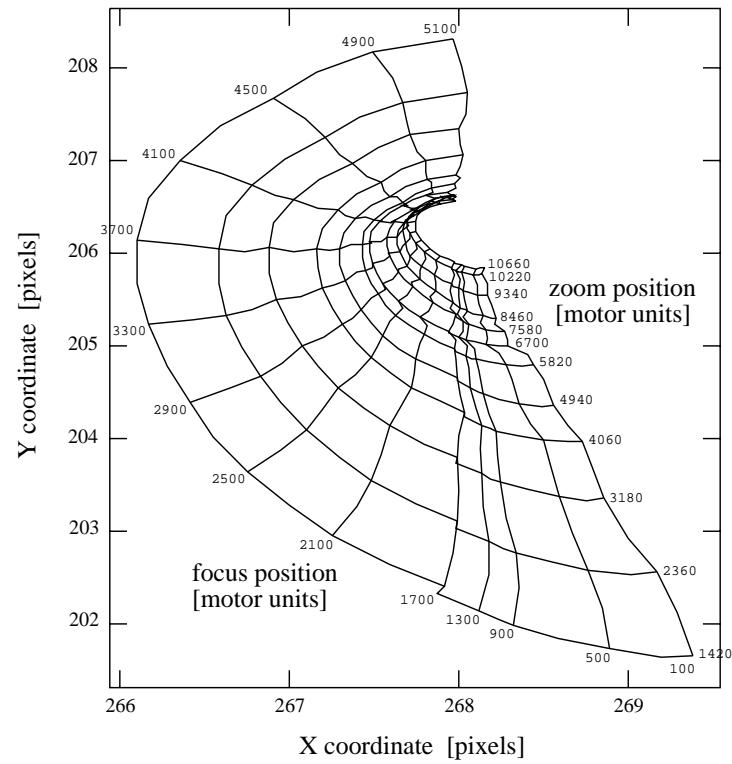


Figure 3.2: Shift in image center as a function of the focus and zoom motors

With adjustable lenses the focus and magnification are changed by varying the positions of the lens elements along the axis of the lens. Moving the lens elements is typically accomplished in one of two ways. In the first method the lens elements are mounted in a threaded section of the lens barrel, which can be rotated to move the elements along the axis of the lens. In the second method the lens elements are mounted on slides or rails, which can be translated along the axis of the lens. For both approaches the misalignments between the lenses' mechanical and optical axes change every time the positions of the lens elements are changed. The rotation of a lens group will cause a rotational drift in the position of the lens's optical axis [28], while the sliding of a lens group will cause a translational motion of the lens's optical axis in the image plane. These rotational and translational shifts in the position of the optical axis cause a corresponding rotational and translational shifting of the camera's field of view.

In lenses with variable focus and fixed focal length as illustrated in Fig. 3.4, typically all the lens elements are mounted in a single assembly. To vary the lens's focus the separation between the lens assembly and the camera sensor is changed by moving the lens assembly with either a rotational or translational mechanism. A less common focusing method found in newer 35 mm, auto-focus lens designs involves the movement of a small, lightweight element within the lens's optics to vary the focus of the image [21].

In lenses with variable focus and variable focal length (i.e. zoom lenses) (see Fig. 3.5), the focal length is changed by moving groups of lens elements relative to one another along the axis of the lens. To keep the focused distance of the lens constant as the focal length is varied, one of the lens groups must be shifted in a non-linear motion, as illustrated in Fig. 3.6. Typically this type of mechanical compensation is accomplished with a pin sliding in a rotating cam, such as the one shown in Fig. 3.7. While the cam moves with a rotating motion during zooming, the lens groups themselves move with a translational motion. The focus in zoom lenses is typically varied by using a rotational mechanism on the front lens group.

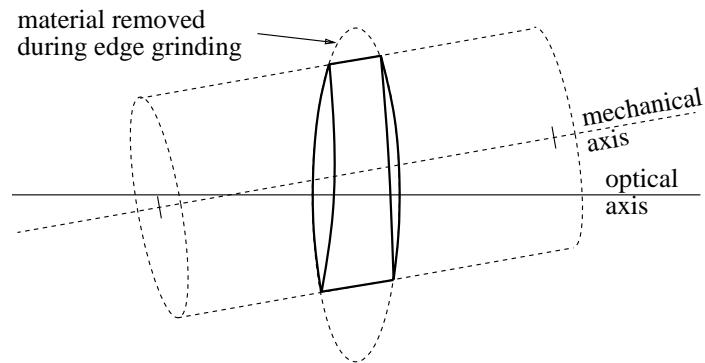


Figure 3.3: Decentration for a simple lens

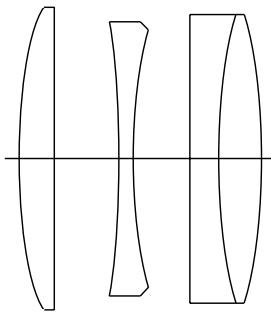


Figure 3.4: Fixed focal length lens (From [25])

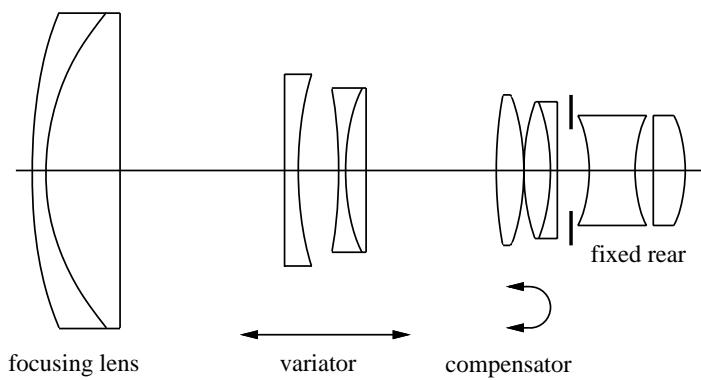


Figure 3.5: Variable focal length (zoom) lens (From [25])

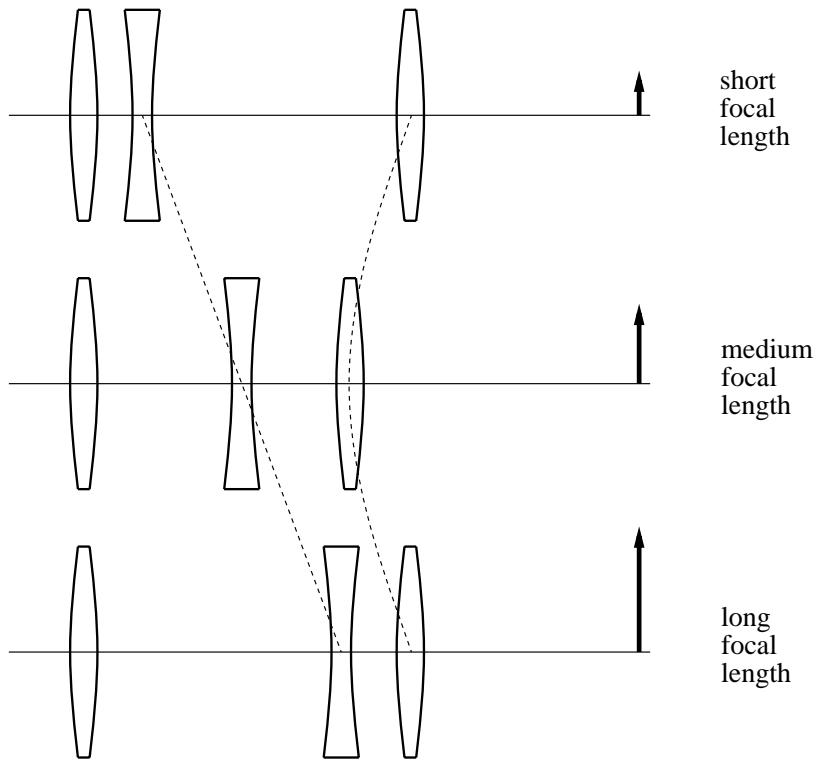


Figure 3.6: Non-linear motion of lens groups during mechanical compensation (From [43])

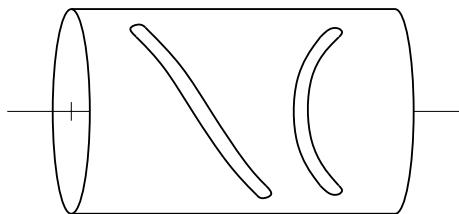


Figure 3.7: Mechanical compensation cam (From [43])

### 3.3 A taxonomy of image centers

In machine vision the image center is most commonly defined as the *focus of expansion* or the *center of perspective projection*. While the *numerical center* of the image (i.e. digitizer) coordinates is another commonly used definition, this one, unlike the other two, does not involve measurements of a system's actual imaging properties. Finding the center of image coordinates belongs to the class of techniques that we call *non-imaging* since they require no image measurements. Determining the center of perspective projection belongs to a second class that we call *single image techniques* that measure properties of images taken at a single lens setting. The focus-of-expansion approach belongs to a third class, called *multi-image techniques*, which measure properties that occur between two or more images taken at different lens settings. This approach should not be confused with simply tracking one of the single-image techniques over different lens settings.

We base our taxonomy on number of different lens settings required for establishing the image center. For techniques that make use of image measurements we further divide our taxonomy into two subcategories: *feature based* and *non-feature based*. Feature-based techniques involve the detection of feature points in the image followed by the application of a geometric interpretation of the 3D to 2D projection to yield an image center. The center of perspective projection (section 3.3.2) is an example of this type of technique. Non-feature-based techniques involve using the image sensor or some other sensing device to take direct measurements of the image formed by the lens. Taking the image of an autocollimated laser (section 3.3.2) is an example of this type of technique.

Based on these techniques we can give at least 15 different definitions of image center and divide them into the following classes:

#### Non-imaging

- Numerical Center of Image/Digitizer Coordinates (section 3.3.1)
- Center of Sensor Coordinates (section 3.3.1)

#### Single image

##### Feature based

- Center of Radial Lens Distortion (section 3.3.2)
- Center of Perspective Projection (section 3.3.2)
- Center of Lines of Interpretation (section 3.3.2)
- Center of Field of View (section 3.3.2)

##### Non-feature based

- Center of an Autocollimated Laser (section 3.3.2)
- Center of  $\cos^{4th}$  Radiometric Falloff (section 3.3.2)
- Center of Vignetting/Image Spot (section 3.3.2)
- Center of Focus/Defocus (section 3.3.2)

#### Multi-image

##### Feature based

- Center of Expansion (section 3.3.3)
- From Focus

- From Zoom
  - From Aperture
  - From Color Band
  - Focus of Expansion
- (section 3.3.3)

### 3.3.1 Non-imaging definitions

By definition non-imaging techniques do not make use of image properties to determine image center. Instead, the image center is defined in terms of the camera's sensor or digitizer properties. These properties in turn depend on the type of camera being used. Two techniques are used in modern solid state cameras to obtain digital images from a camera's sensor: video output cameras (also called closed-circuit television or CCTV cameras) and non-video digital output cameras (also called scientific, slow-scan, or pixel-coded cameras).

In video output cameras each row of the CCD is scanned off the sensor and converted to a continuous analog signal. This signal is resampled by a digitizer board to obtain a digital representation for the row. In this type of camera there is a direct relationship between the row numbers on the sensor and the row numbers on the digitizer. However, the relationship between the column numbers on the sensor and the column numbers in the digitizer is not direct: instead, it depends on the synchronization of the digitizer to the start of each row's analog signal and on the relative rates of the sensor's output clock and the digitizer's sampling clock.

In non-video digital output cameras the sensor's pixels are digitized directly as they are clocked off of the sensor, resulting in a one-to-one correspondence between the sensor's row and column pixel coordinates and the digitizer's coordinates.

#### Numerical center of image/digitizer coordinates

If the numerical center of the image coordinates is used as image center then the coordinates of the image center are trivially given by

$$\begin{aligned} C_x &= \frac{x_{max} - x_{min}}{2} \\ C_y &= \frac{y_{max} - y_{min}}{2} \end{aligned}$$

where  $x_{max}$ ,  $x_{min}$ ,  $y_{max}$ , and  $y_{min}$  are the maximum and minimum column and row numbers respectively.<sup>1</sup>

---

<sup>1</sup>Throughout this thesis we specify the image center in pixels along  $xy$  image coordinates, where  $x$  corresponds to column number in the image and  $y$  corresponds to row number.

### Center of sensor coordinates

If the numerical center of the sensor's pixel array is to be used as the image center then the coordinates of the image center are given by

$$\begin{aligned} C_x &= (C_{x \text{ sensor}} - h_x) \times \frac{f_{\text{sensor clock}}}{f_{\text{digitizer clock}}} \\ C_y &= C_{y \text{ sensor}} - h_y \end{aligned}$$

where

- $C_{x \text{ sensor}}$  is the center of the sensor in pixels in the  $x$  direction,
- $C_{y \text{ sensor}}$  is the center of the sensor in pixels in the  $y$  direction,
- $h_x$  is the number of sensor columns skipped over before digitizing starts,
- $h_y$  is the number of sensor rows skipped over before digitizing starts,
- $f_{\text{sensor clock}}$  is the frequency that sensor elements are clocked off of the CCD and
- $f_{\text{digitizer clock}}$  is the frequency at which the digitizer samples the video signal.

For non-video digital output cameras  $h_x$  and  $h_y$  are integers and  $f_{\text{sensor clock}} = f_{\text{digitizer clock}}$ .

#### 3.3.2 Single image definitions

Single-image techniques rely on the analysis of images taken at one fixed lens setting to estimate the image center. These techniques are important because in many machine-vision systems the lens parameters are not automatically adjustable; they may even be fixed.

### Center of radial lens distortion

“Lens distortion” is the displacement of an image point from the position that is predicted by a camera’s perfect perspective projection. Displacements along radial lines from the center of an image are called “radial lens distortions.” In radial lens distortion the relationship between the distorted position of a point  $(X_d, Y_d)$  on the image plane and the undistorted position of the point  $(X_u, Y_u)$  can be modeled as

$$\begin{aligned} X_u &= (X_d - C_x)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \dots) + C_x \\ Y_u &= (Y_d - C_y)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \dots) + C_y \end{aligned}$$

$$r = \sqrt{\left[ \frac{d_x}{s_x} (X_d - C_x) \right]^2 + \left[ d_y (Y_d - C_y) \right]^2}$$

where  $d_x$ ,  $d_y$ , and  $s_x$  are camera constants and  $\kappa_i$  are the distortion coefficients.

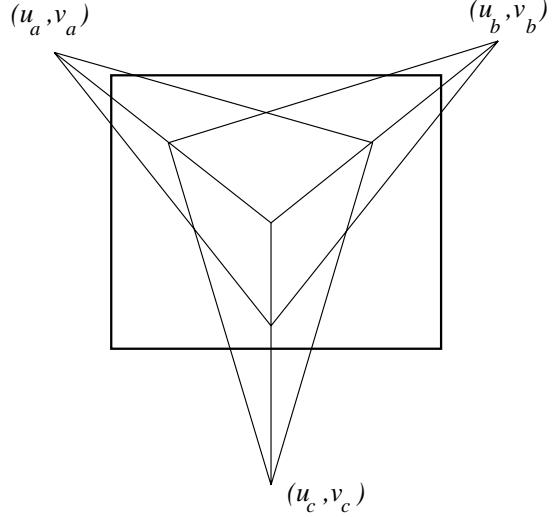


Figure 3.8: Vanishing points of a right-angled cube

To determine the center of radial lens distortion we start with a rough estimate of image center and calibrate a perspective-projection camera model for the lens using Tsai's algorithm[52]. In Tsai's algorithm the image center is assumed to be a constant that is determined before the camera model is calibrated. After calibrating we use iterative, non-linear optimization to determine the image center that produces the minimum image-plane error for the camera model. The details for this procedure are presented in Chapter 5.

### Center of perspective projection

Under perspective projection, lines that are parallel in the object space but not parallel to the camera's sensing plane will appear to intersect at a location  $(u, v)$ , called a vanishing point. With three sets of lines, where the lines within each set are parallel in object space and where each of the sets are not parallel with each other or the image plane, there will be three vanishing points  $(u_a, v_a)$ ,  $(u_b, v_b)$ , and  $(u_c, v_c)$ . Further, if the three sets of parallel lines are mutually perpendicular in object space, then the center of perspective projection for the camera can be calculated from the three vanishing points using the formula presented in [54],

$$\begin{bmatrix} C_x \\ C_y \end{bmatrix} = \begin{bmatrix} u_c - u_a & v_c - v_a \\ u_c - u_b & v_c - v_b \end{bmatrix}^{-1} \begin{bmatrix} u_b(u_c - u_a) + v_b(v_c - v_a) \\ u_a(u_c - u_b) + v_a(v_c - v_b) \end{bmatrix}$$

An image of three sets of parallel lines that are mutually orthogonal can easily be obtained by imaging the corner of a right-angled cube and using the cube's nine visible edges, as shown in Fig. 3.8.

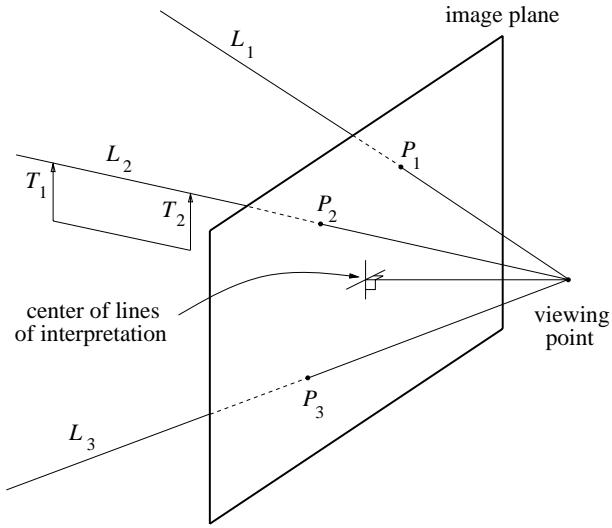


Figure 3.9: Center of lines of interpretation

### Center of lines of interpretation

In a camera each pixel lies on a “line of sight” (line of interpretation) through the object space. Theoretically all lines of interpretation should intersect behind the image plane at a location, the camera’s *viewing point*. The normal projection of the viewing point onto the imaging plane defines a center for the lines of interpretation. For this approach we require the equations of at least three non-coplanar lines of interpretation,  $L_1$ ,  $L_2$ , and  $L_3$ , and the 2D image coordinates of their intersection with the imaging plane,  $P_1$ ,  $P_2$ , and  $P_3$ . The intersection of the lines of interpretation determines the 3D coordinates of the viewing point. The relative 2D distances between the images of the lines of interpretation at  $P_1$ ,  $P_2$ , and  $P_3$  together with the equations of the lines of interpretation determine the parameters of the image plane. Finally, the normal projection of the viewpoint onto the image plane provides us with the image center, as illustrated Fig. 3.9.

To determine the equations of lines of interpretation we use a target consisting of two raised pins,  $T_1$  and  $T_2$ , mounted on the ends of a rod. The rod is manipulated manually until the two pins coincide in the camera’s image plane. A pair of surveyor’s transits are then used to determine the equation in 3D world coordinates of the line of interpretation connecting  $T_1$  and  $T_2$ . The location of the image of the two superimposed pins defines the interception point of the line of interpretation with the image plane.

As with the center of perspective projection, the technique based on lines of interpretation uses a limited number of image measurements to determine the image center, generally without regard to underlying phenomena, such as radial lens distortion. As a result the image centers established with this technique tend not to be very robust.

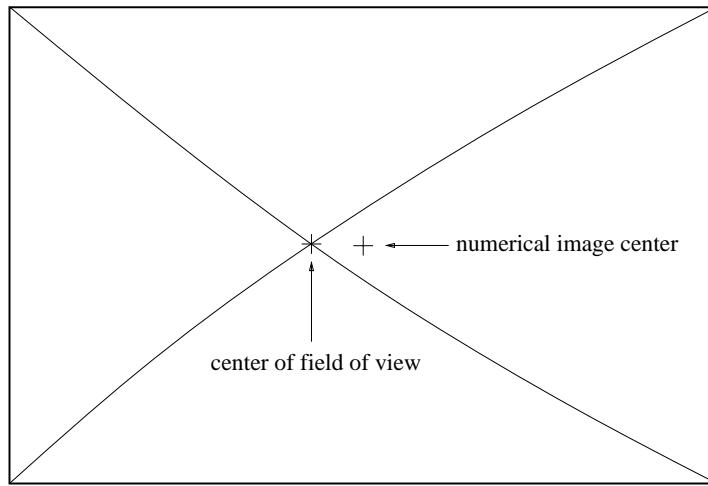


Figure 3.10: Center of field of view

### Center of field of view

A camera's four sensor corners can be used to define the extent of the camera's field of view. The field-of-view center is simply the coordinates of the image of the physical center of the field of view in object space.

To measure the field-of-view center we position the straight edge of a target such that it extends precisely from the upper right hand corner of the camera's image to the lower left hand corner. A second image is taken with the target's edge extending across to the alternate corners of the image. The field-of-view center is then determined by finding the location of the intersection of the edges in the two superimposed images, as shown in Fig. 3.10.

### Center of an autocollimated laser

In an ideal lens the centers of the lens elements' radii of curvature would all fall on a line defined as the optical axis. In this situation a ray of light traveling down the optical axis of the lens would remain unbent and would strike each lens element normal to its surface. Any light reflected back from a lens surface would travel directly back along the path of the incident ray. In a real lens the centers of the lens elements' radii of curvature do not fall on a line. Instead, due to manufacturing tolerances the lens elements are decentered and tilted relative to one another. As a result the reflected light is not returned directly along the same path; instead, it returns at various angles relative to the incident light.

In the autocollimated laser approach a low-power laser beam is passed through a hole in a white screen and into the objective of the lens under test, as illustrated in Fig. 3.11. The laser beam serves as an intense, highly collimated light ray. As the beam travels down the lens, the lens elements reflect part of the ray back out through the lens and onto the white screen. By manipulating the position and orientation of the lens, the reflections coming back

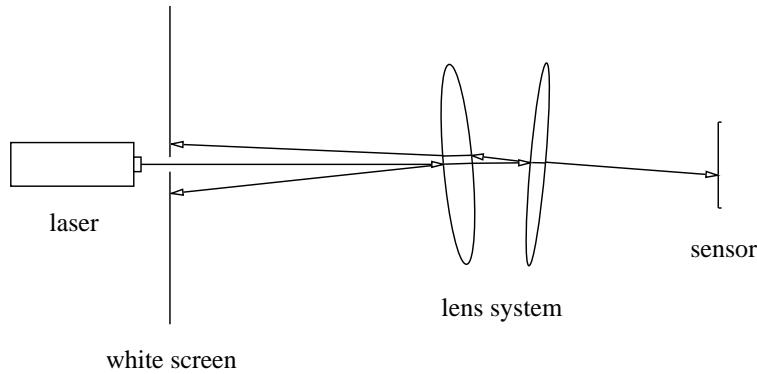


Figure 3.11: Center of an autocollimated laser

from the lens can be roughly lined up with the hole through which the laser is passed. When the reflected light is in its tightest grouping the laser is said to be “autocollimated,” meaning that the laser beam is traveling along the “best” optical axis for the lens. An image taken with the laser in this configuration yields the image center for the autocollimated laser.

### Center of $\cos^{4th}$ radiometric falloff

In a lens system the illumination of the image plane will be found to decrease away from the optical axis at least with the 4th power of the cosine of the angle of obliquity with the optical axis [25]. This falloff can be clearly seen in Fig. 3.12, which shows the profile of a scanline taken from the image of a more or less uniform white field. The abrupt drop in intensity values near the edges is due to vignetting, which will be discussed in section 3.3.2.

The most direct way to determine the center of radiometric falloff would be to take an image of a uniform white field, smooth it to remove per-pixel noise, and then find the location of the intensity peak. In practice it is nearly impossible to create a target with uniform reflectance and illumination across the full field of view. Rather than trying to measure the intensity across the full field of view at once, we instead measure the intensity of a small diffuse calibrated light source. By stepping the calibrated light source across the camera’s field of view we build up a set of intensity measurements for the entire image plane. To determine the center of the radiometric falloff we fit the simple, bivariate-quadratic polynomial

$$I(x, y) = a_{00} + a_{01}y + a_{10}x + a_{11}xy + a_{02}y^2 + a_{20}x^2$$

to the measurements. The position of the polynomial’s peak — the center of the radiometric falloff — is then given by

$$\begin{aligned} C_x &= \frac{a_{01}a_{11} - 2a_{10}a_{20}}{4a_{20}a_{02} - a_{11}^2} \\ C_y &= \frac{a_{10}a_{11} - 2a_{01}a_{02}}{4a_{20}a_{02} - a_{11}^2} \end{aligned}$$

We use a quadratic polynomial instead of a  $\cos^{4th}$  function because the fitting for the polynomial can be done in closed form.

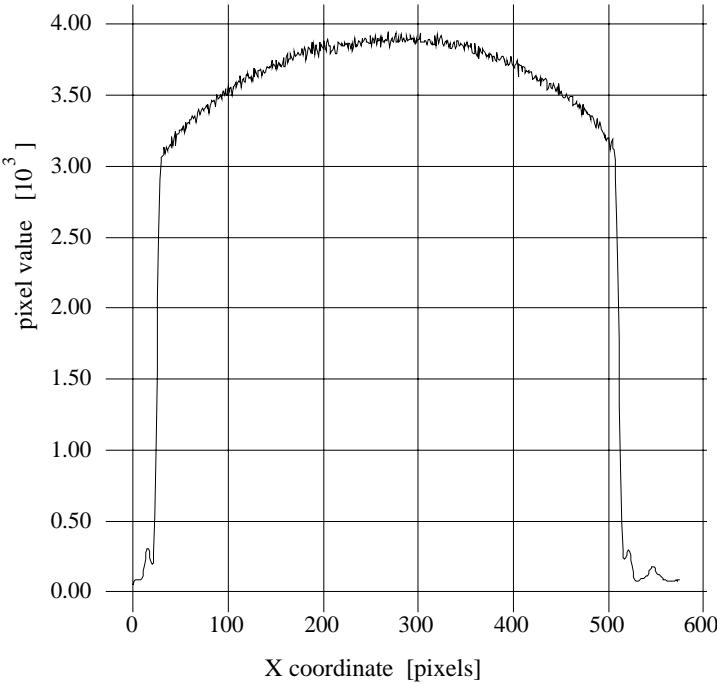


Figure 3.12: Pixel intensity profile for row 200 from Figure 3.14

### Center of vignetting/image spot

For angles nearly parallel to the optical axis the edges of the bundle of rays passing completely through the lens will usually be bounded by the diameter of the aperture stop. However, at more oblique angles of incidence the extreme rays of the bundle may be limited by the front — and rear — lens openings rather than the aperture stop, as shown in Fig. 3.13. This phenomenon is known as vignetting and leads to a reduction of the image illumination at increasing distances away from the axis [25]. Figure 3.14 shows sharply defined vignetting in an image of a uniform white field.

To determine the center of vignetting we locate the edge of the image spot along the rows and columns of the image using a standard Laplacian-of-Gaussian edge-finding technique. A circle is then fit to the spot's edge to estimate the center of the vignetting.

In virtually all commercial camera systems the size of the lens's image spot (the *image format*) is larger than the dimensions of the sensor, specifically to avoid significant vignetting effects. Thus this technique can only be used when the lens is removed from the camera system or in camera systems where the image format is smaller than the sensor size.

### Center of focus/defocus

A planar target in front of an ideal lens would produce an image of the target behind the lens that is also planar. However, with real lenses the image of a plane will not itself lie in a plane. The difference between the position of a plane's real image, illustrated in Fig. 3.15,

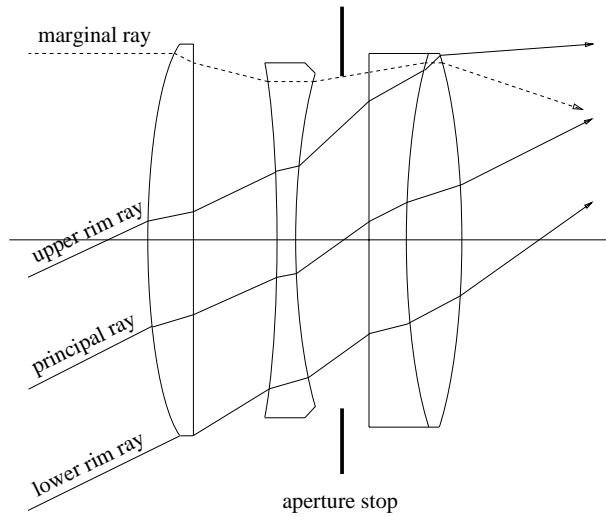


Figure 3.13: Vignetting in a lens (From [25])

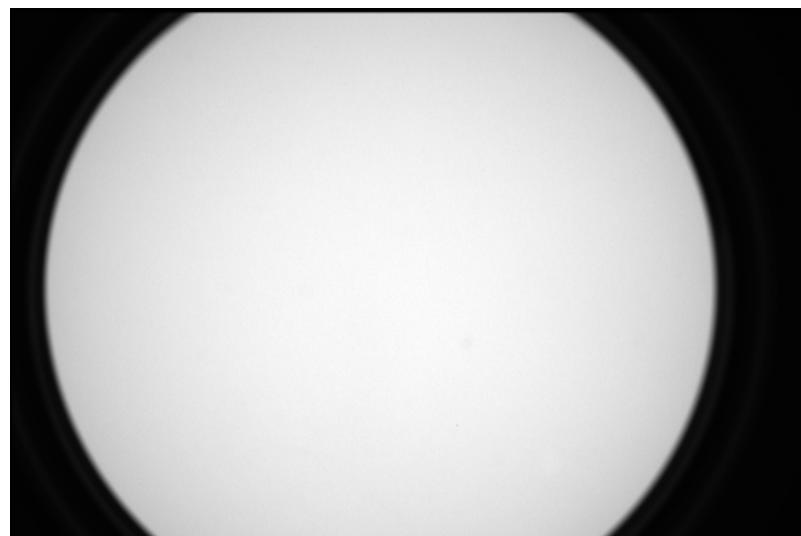


Figure 3.14: Image of a uniform white field showing sharply defined vignetting

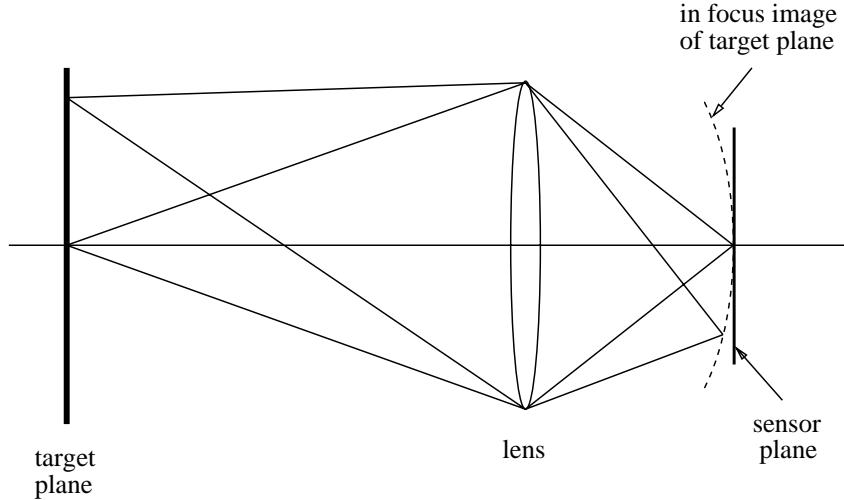


Figure 3.15: Field curvature in the image produced by a thin lens

and its ideal planar image is known as the “field curvature of the lens.” In practical terms, field curvature means that the focused distance of the lens varies across the lens’s field of view, as demonstrated by Nair [34].

To measure the center of focus (or defocus) we first image a target plane that is nearly perpendicular to the axis of the lens and parallel to the sensor plane in the camera. The field curvature of the lens introduces local defocusing in the image of the target plane. If the target plane is nearly perpendicular to the optical axis, the focus/defocus pattern will be radially symmetric. To more accurately measure the amount of defocus we use a target plane containing a uniform, high spatial frequency texture (e.g. a fine checkerboard pattern). A difference operator is run across the image to enhance the focus/defocus information contained in the image’s high frequency content and to attenuate the effect of the low frequency variations in the image intensity due to factors such as illumination and the  $\cos^{4\text{th}}$  law. The image center is then determined by fitting a radially symmetric model to the resulting pattern of focus and defocus.

### 3.3.3 Multi-image definitions

The last class in our image-center taxonomy is based on multi-image techniques. These techniques rely on the analysis of two or more images taken at different lens settings to determine an image center. Since the image center is defined in terms of the differences between images and not in terms of the properties of the individual images, multi-image techniques say more about how lens alignment and centration tolerances interact when the lens parameters are varied than they do about the image properties covered by single-image techniques.

Changing any lens parameter will cause changes in the image parameters, including, for example, the magnification, focused distance, and intensity of the image. While any of these

image properties might be used as the basis of a multi-image definition of image center, image magnification has the most apparent usefulness.

### Center of expansion for focus, zoom, aperture and color band

Given two images taken at different magnifications, exactly one position in the scene in both images will remain in the same place on the image plane. This position is called the “center of expansion” between the two images. More precisely, given two images  $I_1$  and  $I_2$  taken at two magnifications  $m_1$  and  $m_2$ , and given  $n$  reference points  $P_1 \dots P_n$  in image  $I_1$  and the corresponding points  $Q_1 \dots Q_n$  in image  $I_2$ , the center of expansion  $C$  satisfies the constraint

$$(C - P_i) = k(C - Q_i) \quad \forall i = 1 \dots n$$

where

$$k = \frac{m_1}{m_2}.$$

The relative image plane magnification  $k$  can be estimated from the change in relative separation of the points in each image by evaluating

$$\begin{aligned} k_{x_{ij}} &= \frac{q_{x_i} - q_{x_j}}{p_{x_i} - p_{x_j}}, \quad i > j, \quad |q_{x_i} - q_{x_j}| > \text{threshold} \\ k_{y_{ij}} &= \frac{q_{y_i} - q_{y_j}}{p_{y_i} - p_{y_j}}, \quad i > j, \quad |q_{y_i} - q_{y_j}| > \text{threshold} \\ k &= \frac{\Sigma k_{x_{ij}} + \Sigma k_{y_{ij}}}{n_x + n_y} \end{aligned}$$

where  $n_x$  and  $n_y$  are the number of points in the  $x$  and  $y$  directions passing the threshold test. The threshold test is necessary to minimize the effects of the measurement noise in coordinates of the reference points. Typically we use a value that is two to three orders of magnitude greater than the uncertainty in the measurement of the reference point coordinates. If  $k$  is close to unity then the relative positions of the reference points do not move significantly between the two images and the effects of radial lens distortion can be ignored.

To find the center of expansion we first define the squared error for the center as

$$\begin{aligned} e_{x_i} &= (C_x - p_{x_i}) - k(C_x - q_{x_i}) \\ e_{y_i} &= (C_y - p_{y_i}) - k(C_y - q_{y_i}) \\ e &= \sum_{i=1}^n (e_{x_i}^2 + e_{y_i}^2) \end{aligned}$$

To find the  $C_x$  and  $C_y$  that minimize the squared error we differentiate  $e$  with respect to  $C_x$  and  $C_y$  and set the results equal to zero, which yields

$$\begin{aligned} C_x &= \frac{\sum_{i=1}^n (kq_{x_i} - p_{x_i})}{n(k-1)} \\ C_y &= \frac{\sum_{i=1}^n (kq_{y_i} - p_{y_i})}{n(k-1)} \end{aligned}$$

Normally image magnification is changed by varying a lens's zoom, however, magnification can also be changed by varying the focus, aperture, and color band of the lens [56]. Thus, centers of expansion can be defined for all four lens parameters.

### Focus of expansion

In what is known as “focus-of-expansion technique” the trajectories of a number of feature points are tracked across several images taken over a range of zoom settings. The intersection of these trajectories yields an image center called the *focus of expansion*. Since the intersection of the trajectories for any pair of images will yield a center of expansion, the focus of expansion is effectively just the average center of expansion for zoom over a particular range of zoom settings. The equations for the focus of expansion are straightforward generalizations of the equations for the center of expansion.

### 3.3.4 Experimental results

To illustrate the importance of an accurate image center we calibrated our Fujinon lens and Photometrics camera using Tsai's camera calibration technique [52]. The non-coplanar data used in the calibration was obtained using a planar target containing a total of 225 uniformly spaced reference points (a  $15 \times 15$  grid) mounted on a precision motion platform.

In Tsai's technique the image center is considered to be a fixed camera parameter generally determined separately from the calibration of the camera model. Figure 3.16 shows the mean image-plane error for a range of different image centers used in a Tsai calibration on one dataset. For an image center equal to the numerical center of the image at [288, 192] (point 1 in Fig. 3.1) the mean and standard deviation of the image-plane error are 0.553 pixels and 0.413 pixels. However, for our camera and lens the image center that yields the minimum average image-plane error occurs at [258.1, 203.9] (point 9 in Fig. 3.1), where the mean and standard deviation of the error drop to 0.084 pixels and 0.046 pixels.

To illustrate the variation in the position of image center between different definitions we measured ten different image centers for our automated Fujinon lens.<sup>2</sup> The results, drawn to scale in Fig. 3.1 and listed in Table 3.1, show variations of over 90 pixels in the  $x$  direction and over 40 pixels in the  $y$  direction (image size is  $576 \times 384$  pixels). The range of values for the center-of-perspective-projection represents the results of several trials with slightly different orientations of the right-angled cube used to obtain the vanishing points.

---

<sup>2</sup>The first nine measurements were made with a focused distance of 2.16 m, an effective focal length of 98 mm, and an aperture of  $f/8.1$ . The perspective-projection measurements were made with the focused distance varying from 1.2 – 2.0 m, an effective focal length of 20 mm, and an aperture of  $f/12.5$ .

Table 3.1: Different image centers for the same camera system.

Definition	$C_x$ [pixels]	$C_y$ [pixels]
Numerical center of image/digitizer coordinates	288	192
Center of sensor coordinates	290.0	195.5
Center of expansion (zoom)	310.7	182.3
Center of expansion (focus)	324.2	164.8
Center of expansion (aperture)	324.7	191.9
Center of $\cos^{4th}$ falloff	283.1	156.7
Center of vignetting/image spot	273.2	200.1
Center of an autocollimated laser	267.0	204.0
Center of radial lens distortion	258.1	203.9
Center of perspective projection	229-261	165-171

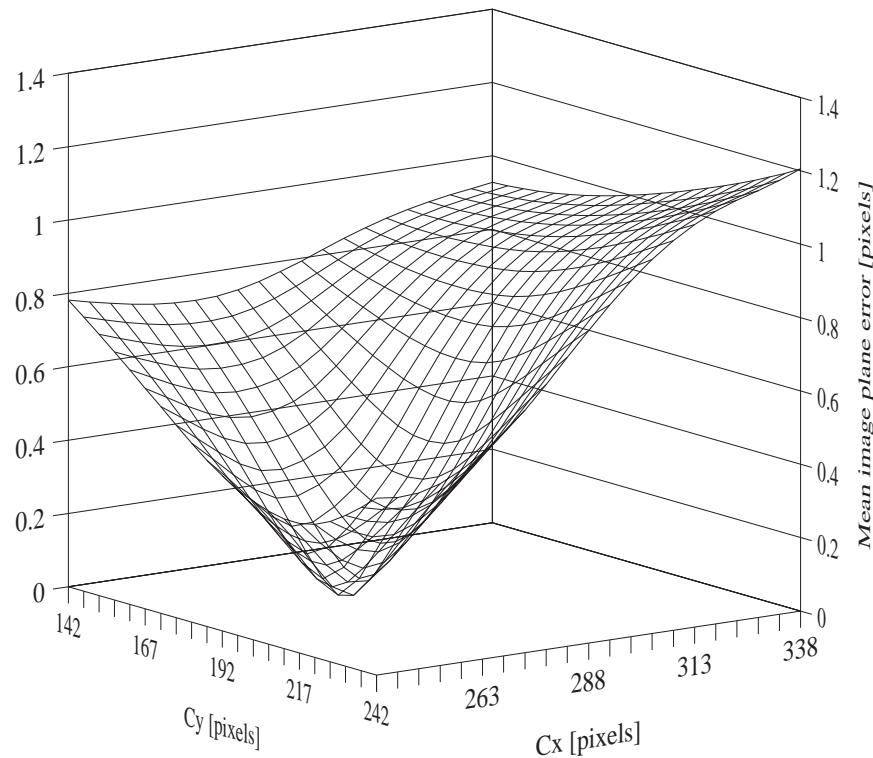


Figure 3.16: Mean image plane error as a function of image center used during calibration

### 3.4 Image center in variable-parameter lenses

Varying the focus and zoom of a lens changes the alignment of the lens components causing the camera's image centers (and its field of view) to shift. As we have shown, knowing the position of the image center is necessary to accurately model radially symmetric image properties. Thus, for our camera models to be calibrated at different lens settings we need to model how the image centers vary with lens parameters.

To see how image centers move in a variable-parameter lens we use the autocollimated laser approach because of its accuracy, repeatability, and robustness over the full range of lens settings. For the first experiment we started by autocollimating the lens at one lens setting. We then stepped through the full range of focus and zoom settings while the centroid of the image of the laser is recorded. The results, plotted in Fig. 3.2, show the laser's image moving across 3.2 pixels in the  $x$  direction and 6.6 pixels in the  $y$  direction over the full range of focus and zoom positions. Two observations are worth noting here. First, the motion of the image center is clearly rotational as a function of focus, as we would expect from the focus mechanism for our lens. Second, the motion as a function of zoom is clearly translational, again as we would expect for our lens.

To determine the mechanical repeatability of the lens we measure the position of the laser as the focus and zoom parameters are stepped through twice.<sup>3</sup> Figures 3.17 and 3.18 show that the lens has good mechanical repeatability. Figures 3.19 and 3.20 show the motion of the laser's image as either the focus or zoom parameter is held constant and the lens is stepped back and forth through the full range of the other parameter. The double curves indicate that there is an appreciable amount of mechanical hysteresis in the lens system, but this can be easily overcome by consistently approaching a given lens setting from one direction.

The discontinuities position of the laser's image in Figs. 3.18 and 3.20 are due to play in mechanical compensation used in the zoom lens. The zoom settings corresponding to the discontinuity mark the point where the direction of the focus compensation group is reversed.

---

<sup>3</sup>The automation for our lens is provided by highly repeatable digital microstepping motors (See Appendix A), thus any error is due primarily to the mechanical and optical properties of the lens itself.

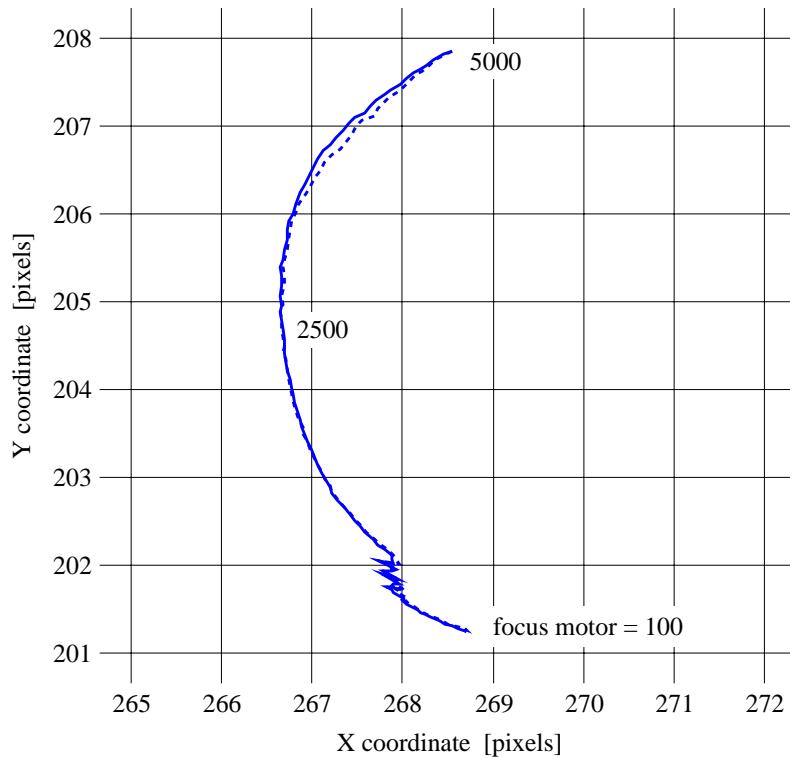


Figure 3.17: Mechanical repeatability of shift in laser image due to focus motor

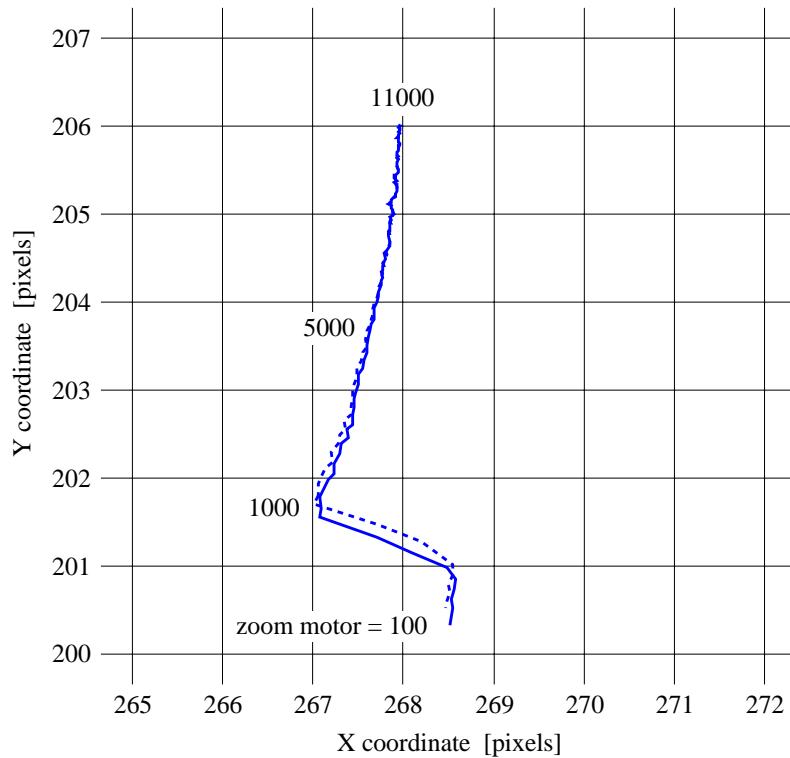


Figure 3.18: Mechanical repeatability of shift in laser image due to zoom motor

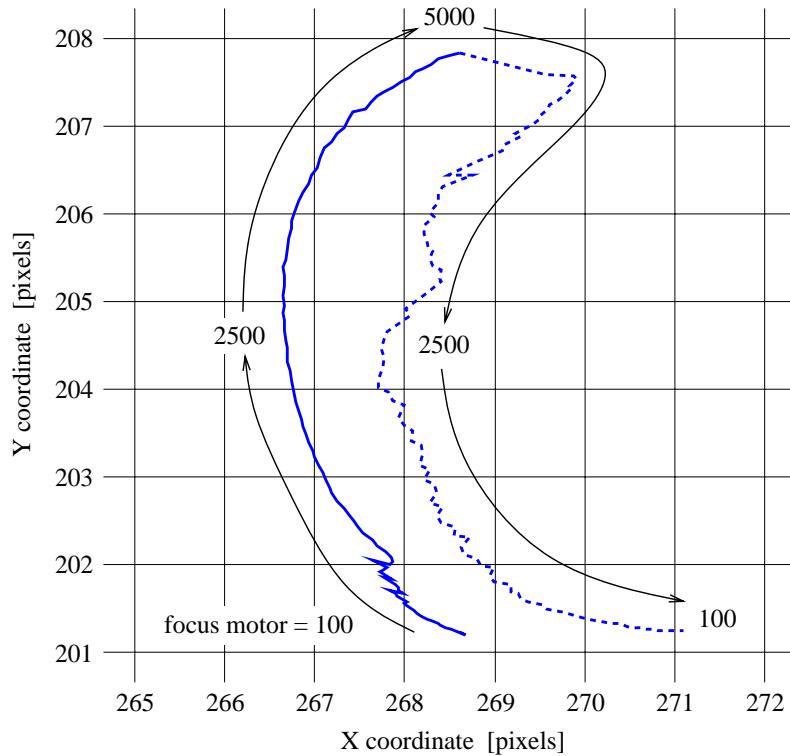


Figure 3.19: Mechanical hysteresis in shift in laser image due to focus motor

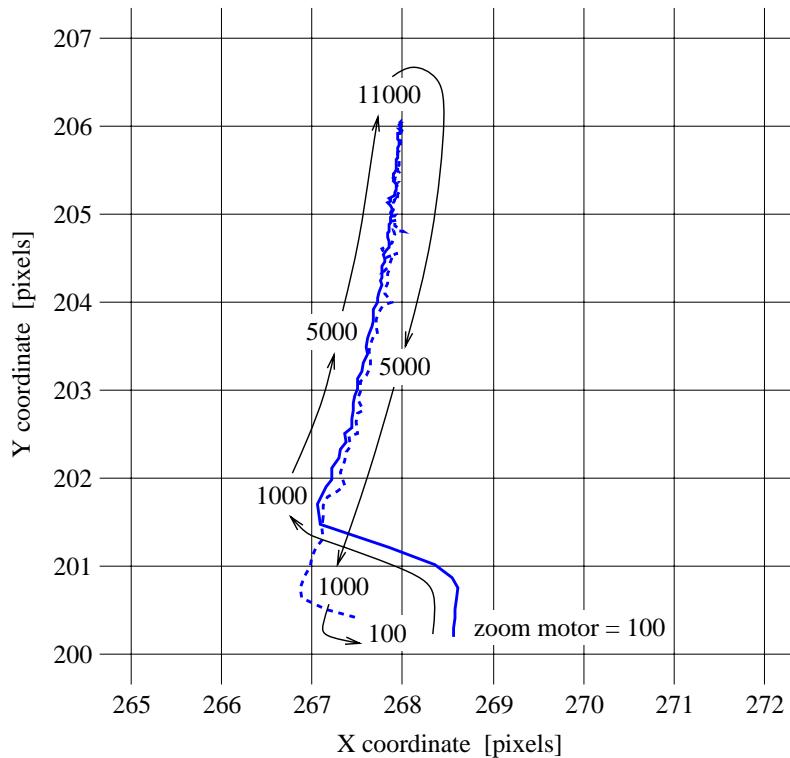


Figure 3.20: Mechanical hysteresis in shift in laser image due to zoom motor

## 3.5 Summary

Calibrating an ideal lens would simply involve modeling a 3D to 2D projection through a single center of perspective projection. Unfortunately, models for real lenses need to take into account additional imaging properties that vary radially in the distance from the center of the image. To capture these properties we need to know their center. As we have demonstrated, an inaccurate image center can have a significant effect on the accuracy of the final, calibrated model.

Still, if lenses could be manufactured perfectly they would have perfect radial symmetry around one well-defined optical axis, which could easily be determined by any one of the 15 methods we have described in this chapter. In practice, however, manufacturing tolerances produce wide variations in the locations of image centers for different image properties. Thus different measurements of image centers are not interchangeable.

The image center calibration problem becomes even more complex in variable-parameter lenses, where manufacturing tolerances can cause image centers to move significantly as the parameters are changed. However, this motion is usually regular and repeatable and can be modeled and compensated.

By using the correct image center for each image property we want to model and by calibrating the image centers over the appropriate ranges of lens parameters, we can significantly improve the calibration accuracy of our camera models.



# Chapter 4

## Modeling and Calibrating Automated Zoom Lenses

In the previous two chapters we demonstrated that the imaging behavior of automated lenses is very complex and that it cannot be predicted with traditional abstract models of the image-formation process. With little *a priori* knowledge about the underlying lens mechanisms we proposed that the only way to capture these behaviors is by collecting large amounts of data and modeling them empirically. This chapter describes our rationale and approach for empirically building adjustable camera models.

### 4.1 Fixed- versus variable-parameter lenses

In modeling and calibrating automated zoom lenses our end objective is to capture the net relationship between the lens control parameters and some aspect of the image-formation process. As illustrated in Fig. 4.1, we can conceptually subdivide this relationship into two parts. The first part,  $R_1$ , represents the relationship between the image-formation process and the hardware configuration of the lens. The hardware configuration is specified by the composition, dimensions, and positions of the optical elements of the lens. The second part,  $R_2$ , represents the relationship between the hardware configuration of the lens and lens's control parameters (if any). In fixed-parameter camera systems the lens's hardware configuration is fixed and we need to consider only  $R_1$  for modeling and calibrating the lens. In variable-parameter camera systems the lens's hardware configuration is adjustable and we must consider both  $R_1$  and  $R_2$ .

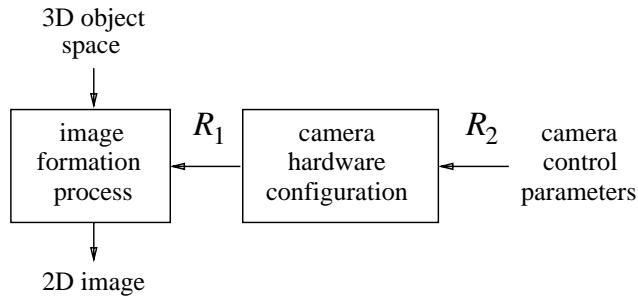


Figure 4.1: Lens control parameters and the image-formation process

#### 4.1.1 $R_1$ — Image-formation process and the hardware configuration

The low-level optics relating a real lens's hardware configuration to the actual image-formation process is generally too complex to be expressed in closed-form equations, even for a simple fixed-parameter camera lens. Lens designers deal with this problem by resorting to simulations of the image-formation process using ray tracing [28]. In ray tracing the paths of individual light rays are traced as they refract at each optical surface in the lens. With enough rays the designer can characterize the lens's image-formation process sufficiently to evaluate the lens's design. While they are explicitly related to the hardware configuration of the lens, the equations used in ray tracing cannot be used to build parameterized models of the imaging properties that we are interested in.

In machine vision we are interested in higher-level aggregate properties of the image-formation process. These range from simple image properties, such as magnification and focused distance, to more complex image properties, such as perspective projection and image defocus. In order to have computationally efficient, closed-form equations for these properties the models must be based on simplifications or abstractions of the actual image-formation process. The two most common abstract models are those of the pinhole camera model, used to explain perspective projection, and the thin-lens camera model, used to explain image defocus. The price for using an abstract model is a loss in accuracy and a disconnection between the equations of the model and the hardware configuration of the lens.

#### 4.1.2 $R_2$ — Hardware configuration and the control parameters

The relationship between the lens's control parameters and the actual hardware configuration of the lens is essentially an arbitrary design choice made by the manufacturer. Typically this relationship is hidden from the user. Worse still, as shown in Section 3.4, the mechanical nature of this relationship introduces hysteresis and discontinuities in the lens's imaging properties, problems that must be dealt with in the adjustable camera model.

### 4.1.3 Formulation issues

For fixed-parameter lenses the image-formation process is static, and thus the terms in the camera model are constants. In variable-parameter lenses the image-formation process is an adjustable function of the lens control parameters, and thus the terms of the camera model must also be variable. The question is, “How do the terms of the camera model vary with the control parameters?” This question is difficult to answer for two reasons: First, the two traditional models of the image-formation process — the pinhole camera and the thin lens — are idealized, high-level abstractions of the real image-formation process, and the connection between the lens’s physical configuration and the model terms is not direct. Second, as we’ve seen, the relationship between the lens’s physical configuration and the control parameters is complex and typically we have very little a priori knowledge about the underlying mechanisms involved. We have no good theoretical basis for the relationships. Since every model term is potentially a function of every lens control parameter, the actual relationships between the terms of our camera models and the lens control parameters must be determined empirically.

### 4.1.4 Calibration issues

Unlike the calibration of fixed-parameter lenses, the calibration of variable-parameter lenses requires measurements over ranges of lens hardware configurations. This raises several challenges. First, the dimensionality of the data is the same as the number of control parameters that are to be concurrently modeled. Even if we just took 10 measurements across the ranges of focus, zoom, and aperture controls, 1000 hardware configurations would have to be calibrated for, compared to just one for a fixed-parameter lens system.

A second challenge are certain imaging situations that cause problems for taking measurements. For example, Figs. 4.2 and 4.3 show the change in size of a checkerboard calibration target for a  $13\times$  zoom lens. As the lens is zoomed in (i.e. the focal length is increased) the number of feature points in the camera’s field of view may decrease below the number necessary to perform an accurate calibration. Conversely, as the lens is zoomed out the features on the target may become too small and/or crowded to be accurately measured. As a result, several targets with different scales may be required to cover the full range of zoom. Taking measurements over wide ranges of focus/defocus or image intensity can also be problematic.

## 4.2 Adjustable camera models

The approach we use to model a variable-parameter camera system is simply to characterize how the parameters of the fixed camera model vary with lens settings. The approach has three steps:

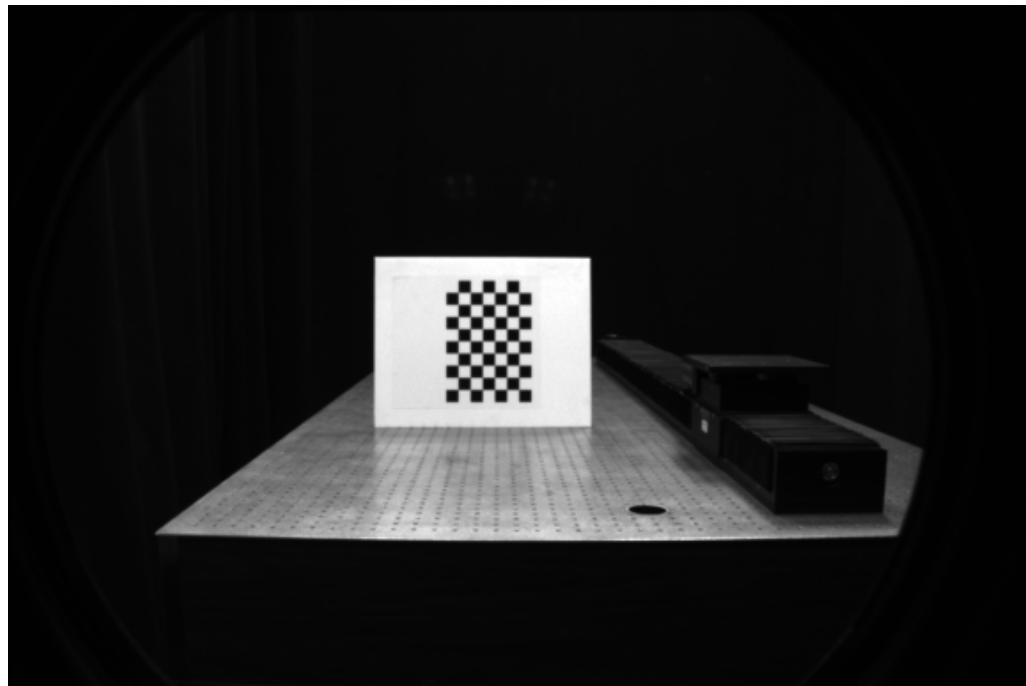


Figure 4.2: Checkerboard target at 130 mm focal length and 2 m range



Figure 4.3: Checkerboard target at 10 mm focal length and 2 m range

1. Collection of calibration data for the fixed camera model across a range of lens settings
2. Estimation of the fixed camera model's parameter values at each lens setting
3. Characterization of the relationship between the fixed camera model's parameters and the lens settings.

### 4.2.1 Notation

Before we continue we introduce the following notation.

**Lens setting:** A three tuple containing the control settings for the focus, zoom, and aperture motors on the lens.

$$S = \{m_f, m_z, m_a\}$$

**Fixed camera model:** A set of parameter values  $P_1, \dots, P_n$  that instantiate a set of equations describing the behavior of a lens at one lens setting.

$$M_f = \{P_1, \dots, P_n\}$$

**Calibration data:** A set of measurements  $D$  or  $d_i$  taken at one lens setting that can be used to estimate the fixed camera model parameters at that lens setting.

$$D = \{d_0, \dots, d_n\}$$

**Parameter model:** A function with coefficients  $a_0, \dots, a_n$  that describes the relationship between a fixed model parameter  $P$  and a lens setting  $S$ .

$$g_P(S) = \text{genericfunction}(S; a_0, \dots, a_n)$$

**Adjustable camera model:** A set of parameter models that describe all the parameter values for a fixed camera model at lens setting  $S$ .

$$M_a(S) = \{g_{P_1}(S), \dots, g_{P_n}(S)\}$$

**Fixed model error (FME):** A measure of how well a fixed camera model  $M_f$  explains a given set of calibration data  $D$  for a single lens setting.

$$\text{FME}(M_f, D) = \sum_{i=1}^n \text{error}(M_f, d_i)$$

**Adjustable model error (AME):** A measure of how well an adjustable camera model  $M_a$  explains several sets of calibration data  $D_1, \dots, D_n$  for lens settings  $S_1, \dots, S_n$ .

$$\text{AME}(M_a, D_1, \dots, D_n, S_1, \dots, S_n) = \sum_{i=1}^n \text{FME}(M_a(S_i), D_i)$$

### 4.2.2 Collecting calibration data

#### Operating ranges

The first step in building an adjustable camera model is determining the range of lens settings the model is to be calibrated for. Physical ranges for the lens settings can be expressed quite directly (e.g.  $1000 \leq m_f \leq 3000$  motor units). However, often we would like to express the operational limits for the model in terms of imaging properties such as focused distance, or depth of field, or effective focal length. Unfortunately the models relating the lens's physical control space to the image properties are often the models we are trying to build. When we have no models the only approach left is to conduct experimental surveys of the control space to find approximate limits. The choice of a “useful” operating space in terms of the imaging properties is a subject which we will defer for another thesis.

#### Sampling strategy

To formulate and calibrate the adjustable camera model we need to take measurements of the camera system at various points throughout its physical operating space. In sampling the physical operating space the sampling frequency must be sufficiently high along each control parameter so that the underlying variations in the parameters of the fixed camera model can be accurately characterized. Since we start with little or no a priori information about the relationships between the control parameters and the parameters of the camera model, the sampling strategy must be determined empirically.

In most cameras the imaging properties are designed to vary smoothly with the lens's control motors. However, two situations may cause sudden changes (discontinuities) in model parameters: when the lens has different operating modes it can switch between (e.g. shifting into and out of macro modes), and when mechanical tolerances allow play in the lens mechanisms. To accurately model rapid changes in the imaging properties we need to densely sample in the control space around them. To avoid over-sampling the entire operating space we need to have a good idea where the discontinuities are in the lens's operating space.

Generally, mechanical discontinuities in the camera's optical hardware will cause simultaneous discontinuities in a broad range of imaging parameters. We can take advantage of this by using “cheap” measurement techniques to map out the locations of the optical discontinuities in the camera's control space in place of potentially more “expensive,” complete camera calibrations. A cheap measurement technique is the center of an autocollimated laser described in Section 3.3.2. The autocollimated laser requires only one static target (the laser) and only one image of the target at each lens setting. Parameter estimation for the centroid ( $c_x, c_y$ ) is also very efficient and the target is easily measurable across the full range of zoom and focus for the lens (since the laser acts as a single ray passing through the lens, focus, zoom and aperture have little effect on its profile).

### 4.2.3 Characterizing variations in fixed model parameters

After having determined the parameters of the fixed camera model for a range of lens settings we must characterize how they vary with the lens settings.

#### Formulation of parameter models

If we estimate the parameter values from the calibration data and just store them in lookup tables we need to make no assumptions about how they vary with the lens settings. However, if we want to use an algebraic form for the parameter models or interpolate between the sampled lens settings, we must make some decisions about the underlying variations in the data. These decisions include which lens controls specifically are factors in the variation and what the order of the variation is. In some cases these decisions may be based on design objectives, such as requiring a particular parameter to be a constant for all lens settings. However, in many cases these decisions must be made by examining the data. This approach is termed adaptive model building by Box [8][9], and trial-and-error by Tarantola [50].

As we've suggested, typically we know neither the relationship between the lens motors and the hardware configuration ( $R_1$ ), nor the relationship between the hardware configuration and the parameters of the abstract camera model ( $R_2$ ). With no better basis to go on we must use very generic formulations for the parameter models. Generic functional relationships between a dependent variable  $u$  and an independent variable  $x$  can be constructed using a linear combination of  $m$  arbitrary functions of  $x$  (for a more complete discussion see Press[38]). The general form of the generic model is

$$u(x) = \sum_{i=1}^m a_i X_i(x) \quad (4.1)$$

where  $X_1(x), \dots, X_m(x)$  are arbitrary (possibly non-linear) functions of  $x$ , called a *basis*. Possible basis functions include regular polynomials, Chebyshev polynomials, Legendre polynomials, and trigonometric functions.

Given  $n$  pairs of  $u$  and  $x$  we can construct the linear system

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} X_1(x_1) & \cdots & X_m(x_1) \\ \vdots & & \vdots \\ X_1(x_n) & \cdots & X_m(x_n) \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_m \end{bmatrix} \quad (4.2)$$

$$U = WA$$

For  $n = m$  (and with no two  $x_i$ 's alike) the coefficients of (4.1) can be solved for directly using  $A = W^{-1}U$ . For  $n > m$  the system is over-constrained, and no unique solution exists for  $A$ . The least-squared-error solution can be found using the pseudo-inverse approach  $A = (W^T W)^{-1} W^T U$ . Note that the pseudo-inverse approach squares the condition number for the system reducing the precision of the coefficients by one half[57]. If the system is ill-conditioned a more robust approach, such as singular value decomposition, may be required to find an accurate solution for  $A$ .

The functional relationships that we are required to model frequently have two independent variables, the focus motor  $m_f$  and the zoom motor  $m_z$ . Empirically we have found that bivariate polynomials describe these functional relationships well and decided use them as our basis functions for our parameter models<sup>1</sup>.

The general formula for a  $q$ th order bivariate polynomial is

$$g(m_f, m_z) = \sum_{i=0}^q \sum_{j=0}^{q-i} a_{ij} m_f^i m_z^j$$

The number of coefficients required by the polynomial is

$$k = \sum_{i=0}^q \sum_{j=0}^{q-i} 1 = \frac{(q+1)(q+2)}{2}$$

Given  $n$  data points  $(u_1, \dots, u_n)$ ,  $n \geq k$ , the  $k$  coefficients  $(a_{00}, \dots, a_{q0}, a_{0q})$  for a  $q$ th order bivariate polynomial can be found by solving the  $n$  linear equations

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} 1 & m_{f_1} & m_{z_1} & m_{f_1}m_{z_1} & \cdots & m_{f_1}^q & m_{z_1}^q \\ 1 & m_{f_2} & m_{z_2} & m_{f_2}m_{z_2} & \cdots & m_{f_2}^q & m_{z_2}^q \\ 1 & m_{f_3} & m_{z_3} & m_{f_3}m_{z_3} & \cdots & m_{f_3}^q & m_{z_3}^q \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & m_{f_n} & m_{z_n} & m_{f_n}m_{z_n} & \cdots & m_{f_n}^q & m_{z_n}^q \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \\ \vdots \\ a_{0q} \end{bmatrix} \quad (4.3)$$

Since the number of data points must be greater than or equal to the number of coefficients being estimated, the highest order bivariate polynomial that can be fit to a set of  $n$  data points is

$$q_{max} = \left[ \frac{\sqrt{8n+1} - 3}{2} \right]_{INT} \quad (4.4)$$

$q_{max}$	minimum $n$
0	0
1	3
2	6
3	10
...	...

where  $[ ]_{INT}$  returns the integer part.

The above equations can be generalized for larger numbers of independent variables. Also, in the most general form the polynomial orders need not be the same for each of the independent variables.

Having chosen a form for the parameter models we must now determine what orders to use. The higher the order of the polynomial model the better the polynomial model is able to

---

<sup>1</sup>In Appendix D an examination of the calibration data leads us to choose an exponential formulation for a relative aperture model.

follow variations in the data. Beyond a certain point though these variations are due to noise rather than the underlying data. Higher polynomial orders also have more coefficients and as a result require more data (i.e. sampled lens settings) to calibrate. If the variation in the parameter values is due to zero-mean Gaussian noise then a standard Chi-squared test can be used to determine the best (in the maximum likelihood sense) polynomial order to use for modeling the data. If the noise is not zero-mean Gaussian, we must use a more empirical approach to determine the model order.

One such approach is the Minimum Description Length (MDL) principle developed by Rissanen [39]. The competing factors in this approach are the complexity of the final adjustable camera model (i.e. the set of parameter models) and the error between the adjustable camera model and the calibration data. One good metric for the adjustable model's complexity is the total number of coefficients required of the parameter models. Note that the error metric here is the error between between the adjustable camera model and the calibration data and not the error between the individual parameter models and the parameter values estimated from the calibration data.

### Calibration of parameter models

Having chosen formulations for the parameter models we now want to find the best fit between the adjustable camera model  $M_a$  (as represented by the parameter models) and the calibration data. If each fixed model parameter  $P$  is estimated from the calibration data independently then we can fit the parameter models to the estimated fixed model values one parameter at a time. If the variation in the parameter values is zero-mean Gaussian, then the least-squared-error fitting approach described in (4.3) or (4.3) will provide the coefficients for the maximum likelihood model for the parameter. If the noise is not normally distributed we can resort to local M-Estimates [38] to fit the parameter models.

If the fixed model parameters are jointly estimated the calibration problem becomes very complex. In the most direct approach we would start by replacing the parameters in the fixed camera model with the formulations for the parameter models<sup>2</sup>. Then we would have to estimate the values for all of the coefficients for all of the parameter models using the calibration data from all of the lens settings. In fixed models with any complexity this approach would be computationally prohibitive.

Instead of trying to fit all of the parameter models to the calibration data at the same time, we can work with the parameters one at a time. In our approach we fit one parameter model to the estimated fixed model values, set it aside, and then reestimate the remaining fixed camera model parameters from the calibration data. This process is repeated until all the parameter models have been fit.

Naturally, as each freely estimated parameter in the fixed camera model is replaced with a parameter model, the error between the camera model and the calibration data increases.

---

<sup>2</sup>If the fixed camera model was not non-linear before this step it would quite likely be non-linear afterwards.

For a given set of parameter models the final level of error generally depends on the sequence in which the models are fit to the calibration data.

The most obvious sequencing algorithm is a purely greedy one. In it we conduct a series of trials by separately fitting the parameter models to the as yet unmodeled parameters. We then choose the fit that increases the the error between the camera model and the calibration data (the AME) by the smallest amount. The process is repeated until all the parameter models have been fit.

In practice this approach works poorly since higher order parameter models tend to get fit first. As the fitting of the parameter models progresses, any noise-induced variation in the estimated values of the fixed camera model's parameters gets concentrated into the last parameters to be fit. The lower polynomial order of these parameter models makes for a poor fit with the data, causing the error between the camera model and the calibration data to blow up.

A better approach is to fit the parameter models from lowest polynomial order to highest order, using a greedy algorithm whenever two or more parameter models have the same polynomial order. We call this algorithm the ascending polynomial order, greedy within order sequencing.

After the parameter models have been fit to the estimated parameter values we cycle through the parameters, reestimating and then refitting the parameter models to improve the fit between the adjustable camera model and the calibration data. This process continues until no further improvement is seen in the AME.

The sequencing problem could also be cast as another MDL problem. In this case the competing factors would be the total number of coefficients for the parameter models that have been fit and the error between the camera model and the calibration data. We have not explored this approach further.

## 4.3 Summary

The methodology we presented for empirically building camera models for systems with variable-parameter lenses involves first calibrating a conventional fixed camera model at a number of lens settings spanning the adjustable model's physical operating space. We then characterize how the parameters of the fixed model vary with individual lens settings by alternately fitting polynomials to individual model parameters and reestimating the as yet unfitted parameters using the calibration data. This process is repeated until all of the fixed camera model's terms have been replaced with polynomial functions of the lens control parameters. The resulting predictive camera model can interpolate between the original sampled lens settings to produce — for any lens setting — a set of values for the parameters in the fixed camera model. As part of our methodology we use a preliminary measurement survey of the physical operating space to identify regions where rapid changes in imaging properties may require denser sampling for calibration data.

Unlike previously developed modeling approaches, ours makes no a priori assumptions about the dependencies between the parameters of the fixed camera model and the lens settings. Our approach handles imaging behavior that cannot be explained using the abstract models of the image-formation process, is general and can be applied to produce — for any image property — an adjustable camera model from the fixed one and allows any number of independent lens control variables to be incorporated. Furthermore, the degree of accuracy and complexity, and consequently the required calibration effort, can be chosen arbitrarily.



# Chapter 5

## A Fixed Perspective-Projection Camera Model

To build an adjustable camera model for an imaging property we are interested in requires a fixed model of that property. In this chapter and the next we describe the approach we use to model 3D to 2D perspective-projection. In the following sections we explain the formulation for the fixed model along with the algorithms and techniques for calibrating it.

### 5.1 Perspective-projection camera models

Perspective-projection camera models map the coordinates of points in 3D object space into 2D image coordinates. Alternately, an inverted perspective-projection model can be used to determine the corresponding line-of-sight through the 3D object space based on a 2D image coordinate. Perspective-projection models are used in viewpoint planning and for measurement applications, such as stereo.

Many different models have been used for perspective projection [2], [12], [58], [22], [20], [14], [5]. We use the model described by Tsai [53].

Tsai’s camera model consists of 11 parameters: six extrinsic, “exterior-orientation” parameters ( $R_x, R_y, R_z, T_x, T_y, T_z$ ) that describe the position and orientation of the camera’s coordinate frame with respect to the world-coordinate frame, and five intrinsic, “interior-orientation” parameters ( $f, C_x, C_y, s_x, \kappa_1$ ) that describe the camera’s image-formation process. All 11 camera parameters are constants estimated from calibration data taken from a single camera view (i.e. the exterior and interior orientation of the camera is fixed). Whenever the camera is moved in the world-coordinate system its exterior orientation must be recomputed while its interior orientation remains unchanged.

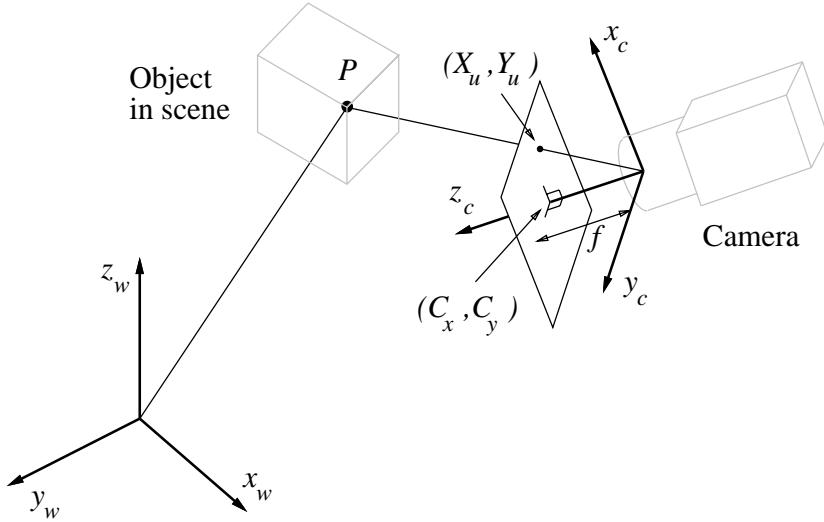


Figure 5.1: Fixed perspective-projection camera model geometry

## 5.2 Formulation of Tsai's model

In Tsai's model, illustrated in Fig. 5.1, the origin of the camera-centered coordinate system  $(x_c, y_c, z_c)$  coincides with the front nodal point of the camera, the  $z_c$  axis coincides with the camera's optical axis. The image plane is assumed to be parallel to the  $(x_c, y_c)$  plane and at a distance  $f$  from the origin, where  $f$  is the effective focal length of the pinhole camera.

The relationship between the position of a point  $P$  within the world coordinates  $(x_w, y_w, z_w)$  and the point's image in the camera's frame buffer  $(X_f, Y_f)$  is defined by a sequence of coordinate transformations. The first transformation is a rigid body rotation and translation from the world-coordinate system  $(x_w, y_w, z_w)$  to the camera-centered coordinate system  $(x_c, y_c, z_c)$ . This is described by

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (5.1)$$

where

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (5.2)$$

is the  $3 \times 3$  rotation matrix describing the orientation of the camera in the world-coordinate system.  $R$  can also be expressed as

$$R = \text{Rot}(R_x)\text{Rot}(R_y)\text{Rot}(R_z) \quad (5.3)$$

the product of three rotations around the  $x$ ,  $y$ , and  $z$  axes of the world-coordinate system.

The second transformation is a perspective projection (using an ideal pinhole-camera model) of the point in the camera coordinates to the position of its image in undistorted sensor-plane

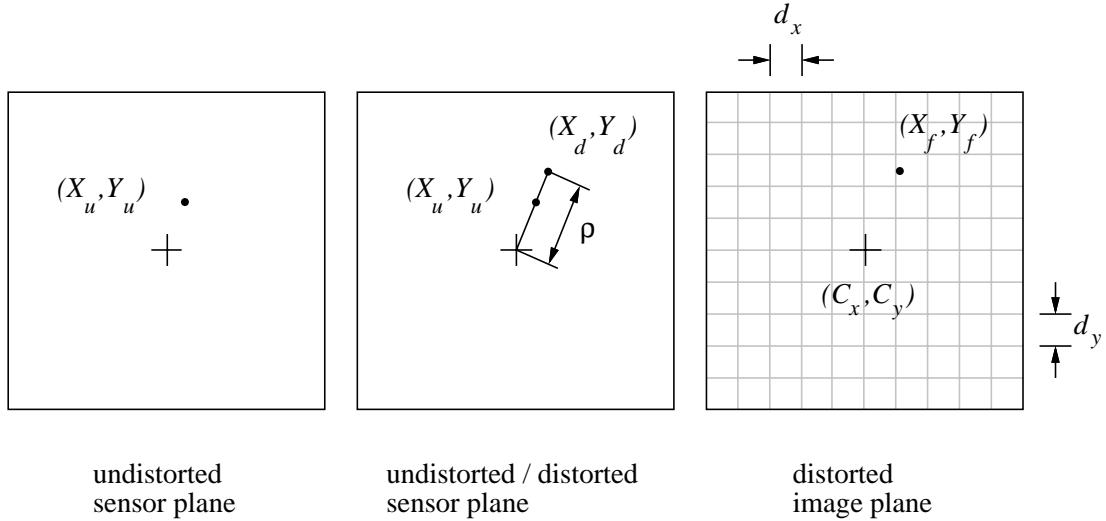


Figure 5.2: Transformation from undistorted sensor to distorted frame coordinates

coordinates,  $(X_u, Y_u)$ . This transformation is described by

$$X_u = f \frac{x_c}{z_c} \quad (5.4)$$

and

$$Y_u = f \frac{y_c}{z_c} \quad (5.5)$$

where  $f$  is the effective focal length of the pinhole camera.

The third transformation (illustrated in Fig. 5.2) is from the undistorted (ideal) position of the point's image in the sensor plane to the true position of the point's image,  $(X_d, Y_d)$ , which results from geometric lens distortion. This is described by

$$X_u = X_d(1 + \kappa_1 \rho^2), \quad (5.6)$$

$$Y_u = Y_d(1 + \kappa_1 \rho^2) \quad (5.7)$$

and

$$\rho = \sqrt{X_d^2 + Y_d^2} \quad (5.8)$$

where  $\kappa_1$  is the coefficient of radial lens distortion. While a more complex model describing both radial and tangential geometric lens distortion could have been used, the accuracy provided by this model is sufficient to demonstrate the development of the adjustable camera model.

The final transformation is between the true position of the point's image on the sensor plane and its coordinates in the camera's frame buffer,  $(X_f, Y_f)$ . This is described by

$$X_f = d_x^{-1} X_d s_x + C_x \quad (5.9)$$

and

$$Y_f = d_y^{-1} Y_d + C_y \quad (5.10)$$

where  $C_x$  and  $C_y$  are the coordinates (in pixels) of the intersection of the  $z_c$  axis and the camera's sensor plane;  $d_x$  and  $d_y$  are the effective center-to-center distances between the camera's sensor elements in the  $x_c$  and  $y_c$  directions; and  $s_x$  is a scaling factor to compensate for any uncertainty in the ratio between the number of sensor elements on the CCD and the number of pixels in the camera's frame buffer in the  $x$  direction.

### 5.3 Performance metrics

One of the first questions we have about any camera model is how accurately it captures the imaging behavior. This information is necessary both for measuring progress during model calibration and estimating the performance or accuracy of any application the model is used in.

Given the measured coordinates of a point in the object space  $(x_w, y_w, z_w)$  and the measured position of the point's image in the frame buffer  $(X_f, Y_f)$  we can define an error metric for the model anywhere along the model's chain of coordinate transformations. One obvious error metric is the difference between the position of a point's image we measure and the position the camera model predicts. If we use the difference in positions following the last coordinate transformation (i.e. after the lens distortion effects have been added to the point's projection through the camera model) we can define the distorted image plane error (DIPE) as

$$\text{DIPE} = \sqrt{(X_f - X'_f)^2 + (Y_f - Y'_f)^2}$$

where  $(X_f, Y_f)$  is the measured position of the point's image and  $(X'_f, Y'_f)$  is the position of the point's 3D coordinates  $(x_w, y_w, z_w)$  projected through the camera model.

In many applications it is desirable to operate in a virtual, undistorted image plane in the camera. In fact, Tsai's fixed camera model is designed to allow converting directly from distorted sensor coordinates  $(X_d, Y_d)$  into undistorted sensor coordinates  $(X_u, Y_u)$ , while going in the opposite direction requires significantly more computation. We define the undistorted image plane error (UIPE) as

$$\text{UIPE} = \sqrt{(\Delta X_{f_u})^2 + (\Delta Y_{f_u})^2} \quad (5.11)$$

where

$$\begin{aligned}\Delta X_{f_u} &= d_x^{-1}(X_{u_2} - X_{u_1})s_x, \\ \Delta Y_{f_u} &= d_y^{-1}(Y_{u_2} - Y_{u_1}).\end{aligned}$$

$(X_{u_2}, Y_{u_2})$  are calculated from the measured position of the point's image  $(X_f, Y_f)$  using equations (5.6), (5.7), and (5.8), while  $(X_{u_1}, Y_{u_1})$  are calculated from the 3D coordinates of the point  $(x_w, y_w, z_w)$  using (5.1), (5.4), and (5.5). The algorithms that we use to calibrate the camera model (and that we will describe later on) minimize the sum-of-squared error in the undistorted image plane for the calibration data.

In inverse perspective-projection problems it is often helpful to know what level of accuracy the camera model has in the object space. By projecting an image point  $(X_f, Y_f)$  back through the camera model we can calculate the closest distance of approach between the image point's line-of-sight and the point in 3D object space  $(x_w, y_w, z_w)$  that was supposed to have cast the image. This object-space error (OSE) can be calculated as

$$\text{OSE} = \sqrt{(x_c - X_u t)^2 + (y_c - Y_u t)^2 + (z_c - f t)^2}$$

where

$$t = \frac{x_c X_u + y_c Y_u + z_c f}{X_u^2 + Y_u^2 + f^2},$$

$f$  is the camera's effective focal length,  $(x_c, y_c, z_c)$  are the 3D coordinates  $(x_w, y_w, z_w)$  rotated and translated into the camera's coordinate frame, and  $(X_u, Y_u)$  are  $(X_f, Y_f)$  transformed into undistorted sensor coordinates.

The above three error metrics are measurements of how well the camera model captures the perspective projection imaging property of the camera system. Another possible error metric is how well the camera model performs in a particular type of application. Tsai [53] gives a theoretical upper bound for a 3D measurement error in the case where two calibrated cameras have been used in a stereo pair. Das [18] describes a set of equations that can be used to convert uncertainty in the image plane (UIPE) into uncertainty in the range error for a general stereo configuration.

## 5.4 Calibration data

Our estimation of the unknown parameters in the fixed camera model is based on calibration data consisting of 3D object space coordinates and corresponding 2D image coordinates. For the experiments described in this thesis we used a planar calibration target mounted on a translation stage (see Fig. 5.3). The normal of the calibration target is exactly parallel to the stage's direction of travel. The calibration target itself contains 1/8-inch-diameter, black reference points precisely spaced out on a regular, 1-inch grid.

For any set of images of the calibration target the relative 3D coordinates  $(x_w, y_w, z_w)$  of the reference points are known from their grid position in the target plane and from the position of the target plane along the translation stage. The  $(X_f, Y_f)$  positions of the dots in the image plane are measured to sub-pixel, accuracy using the procedure described in Appendix C.

The accuracy of the parameter estimation for the fixed camera model depends in part on the distribution of data points across both the 3D object space and the 2D image plane. To be able to accurately estimate the  $f$  and  $T_z$  parameters the calibration data must have some variation in depth in the camera coordinate frame's  $z$  axis. As a general rule the calibration data should cover a range from the closest to the farthest extent of the volume that the camera model is to be used for. Also, to accurately measure the radial lens-distortion coefficient ( $\kappa_1$ ) and the image center  $(C_x, C_y)$  the data's radial positions should vary across the camera's field of view.

## 5.5 Model calibration

In the calibration of our fixed camera models we assume that the six exterior orientation parameters ( $R_x, R_y, R_z, T_x, T_y, T_z$ ) and five interior orientation parameters ( $f, \kappa_1, C_x, C_y, s_x$ ) are all unknown and must be estimated from the calibration data. We calibrate the fixed camera model in two steps. First we use a quick algorithm to obtain approximate estimates for nine of the camera's 11 parameters and then we use iterative, non-linear optimization to refine all 11 parameters.

### 5.5.1 Initial parameter estimation

For our first calibration step we use Tsai's [53] non-coplanar calibration algorithm to determine values for the  $T_x, T_y, T_z, f, \kappa_1$ , and  $s_x$  parameters and for the  $3 \times 3$  rotation matrix  $R$ . Tsai's algorithm begins by setting the radial lens distortion coefficient ( $\kappa_1$ ) to zero and then calculating  $R, T_x, T_y$ , and  $s_x$  directly from the calibration data. Next, with  $\kappa_1$  still zero an initial estimate is calculated for  $f$  and  $T_z$ . Finally,  $f, T_z$ , and  $\kappa_1$  are refined using an iterative, non-linear optimization.

In Tsai's published algorithm the non-linear optimization of the camera model minimizes the sum of the error

$$\text{error} = (Y_u - Y'_u)^2$$

where  $Y_u$  is the  $y$  component of the measured position of the point's image  $(X_f, Y_f)$  transformed into sensor coordinates and  $Y'_u$  is the  $y$  component of point's world coordinates projected through the camera model into the sensor plane. Presumably this was done for speed and/or to reduce the influence of timing jitter noise that can be present in the  $x$  components of images taken using video cameras[4]. In any event this approach makes almost no use of the calibration information in the  $x$  direction. In our camera systems timing jitter is either not present or can be averaged out of the image. Thus for our implementation of Tsai's algorithm, and for the full non-linear optimization approach described next, we make full use of the calibration data and minimize the sum-of-squared error in both the  $x$  and  $y$  directions, i.e.

$$\text{error} = (X_u - X'_u)^2 + (Y_u - Y'_u)^2 \quad (5.12)$$

In Tsai's algorithm  $C_x$  and  $C_y$  are assumed to be known a priori. Three methods for measuring  $C_x$  and  $C_y$  are outlined by Lenz and Tsai in [30]. As we have demonstrated in Chapter 3, the validity of these types of measurement is questionable. However, since the output of Tsai's algorithm is only serving as an initial estimate for the second step of our camera calibration a rough estimate for  $C_x$  and  $C_y$  is sufficient for our purposes. We use the center of an autocollimated laser (described in Section 3.3.2) to give us a rough estimate.

For closed circuit TV (CCTV) cameras the one-to-one correspondence between the CCD's rows and the camera's frame buffer rows allows the value of  $d_y$  to be obtained directly from the CCD's specification sheets. The value for  $d_x$  depends on the relative frequencies of the clock used to shift data from the CCD onto the video signal and the clock used to sample the

video signal at the frame buffer. Several methods exist for accurately measuring  $d_x$  [30][3]. In Tsai's camera model the  $s_x$  parameter is used to compensate for any error in  $d_x$ . Since we include  $s_x$  in the camera model's non-linear optimization only a rough estimate of  $d_x$  is required.

For digital output cameras, such as the Photometrics camera that we use in some of the following experiments, we can obtain exact values for  $d_x$  and  $d_y$  directly from the specifications for the camera's CCD.

### 5.5.2 Full non-linear optimization

In the second step of calibrating the fixed camera model we take the output of Tsai's non-coplanar algorithm and use iterative non-linear optimization to refine the values of all 11 variable model parameters ( $R_x, R_y, R_z, T_x, T_y, T_z, f, \kappa_1, C_x, C_y, s_x$ ). To perform the non-linear optimization we use the IMSL routine DUNLSF which employs a modified Levenberg-Marquardt algorithm and a finite-difference Jacobian<sup>1</sup> to minimize the error function (5.12) for all points in the calibration data.

Since we cannot easily enforce the orthonormality constraint on the rotation matrix  $R$  in (5.1) during the non-linear optimization we use (5.3) and optimize  $R_x$ ,  $R_y$ , and  $R_z$  instead. The rotation angles  $R_x$ ,  $R_y$ , and  $R_z$  can be determined from  $R$  using the equations

$$R_z = \arctan\left(\frac{r_4}{r_1}\right) \quad (5.13)$$

$$R_y = \arctan\left(\frac{r_7}{r_1 \cos(R_z) + r_4 \sin(R_z)}\right) \quad (5.14)$$

$$R_x = \arctan\left(\frac{r_3 \sin(R_z) - r_6 \cos(R_z)}{r_5 \cos(R_z) - r_2 \sin(R_z)}\right) \quad (5.15)$$

While the  $R$  provided by Tsai's algorithm is not guaranteed to be orthonormal and thus the calculated rotation angles are not necessarily that good, their accuracy is sufficient for them to serve a starting point for the non-linear optimization.

### 5.5.3 Public archive for code

An implementation of the above algorithms, written in the C programming language, can be found in the Vision List Archive SHAREWARE subdirectory available via anonymous ftp from FTP.TELEOS.COM. The code requires two non-linear optimization routines that can be found in either of the IMSL or the NAG commercial software packages.

---

<sup>1</sup>The complexity of the fixed camera model makes the direct calculation of the Jacobian computationally prohibitive.

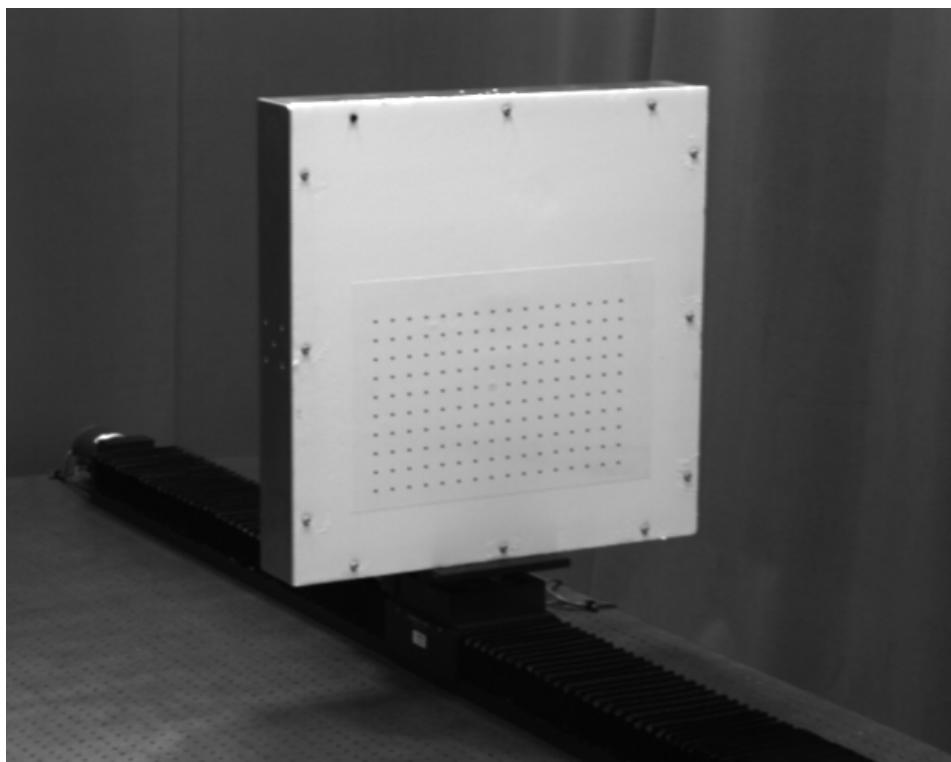


Figure 5.3: Calibration target and translation stage

## 5.6 Recalibrating exterior orientation

Whenever the camera is moved to a new pose the camera's exterior orientation must be recomputed from calibration data taken at the new position. The camera's interior orientation remains unchanged.

To compute just the exterior orientation we use a modification of Tsai's algorithm. We start by using the previously calibrated intrinsic parameters  $C_x$ ,  $C_y$ , and  $s_x$  to transform  $(X_f, Y_f)$  from the new calibration data into undistorted sensor coordinates  $(X_u, Y_u)$ . These coordinates are used in the first stage of Tsai's algorithm to determine  $R$  (and subsequently  $R_x$ ,  $R_y$ , and  $R_z$ ). Given  $(X_u, Y_u)$  and  $R$  we then estimate  $T_x$ ,  $T_y$ , and  $T_z$  using the following approach. Rewriting (5.1), (5.2), (5.4), and (5.5) we obtain two equations,

$$\begin{aligned} X_{u_i} &= f \frac{x_{w_i} r_1 + y_{w_i} r_2 + z_{w_i} r_3 + T_x}{x_{w_i} r_7 + y_{w_i} r_8 + z_{w_i} r_9 + T_z} \\ &= f \frac{x_{k_i} + T_x}{z_{k_i} + T_z} \end{aligned} \quad (5.16)$$

and

$$\begin{aligned} Y_{u_i} &= f \frac{x_{w_i} r_4 + y_{w_i} r_5 + z_{w_i} r_6 + T_y}{x_{w_i} r_7 + y_{w_i} r_8 + z_{w_i} r_9 + T_z} \\ &= f \frac{y_{k_i} + T_y}{z_{k_i} + T_z}, \end{aligned} \quad (5.17)$$

for each point in the new calibration data. Using (5.16) and (5.17) and the new calibration data we then form an over-determined set of linear equations

$$\left[ \begin{array}{ccc} f & 0 & -X_{u_i} \\ \vdots & \vdots & \vdots \\ \hline 0 & f & -Y_{u_i} \\ \vdots & \vdots & \vdots \end{array} \right] \left[ \begin{array}{c} T_x \\ T_y \\ T_z \end{array} \right] = \left[ \begin{array}{c} X_{u_i} z_{k_i} - f x_{k_i} \\ \vdots \\ Y_{u_i} z_{k_i} - f y_{k_i} \\ \vdots \end{array} \right]$$

which can be solved to obtain estimates for  $T_x$ ,  $T_y$ , and  $T_z$ .

Finally,  $R_x$ ,  $R_y$ ,  $R_z$ ,  $T_x$ ,  $T_y$ , and  $T_z$  are refined using the iterative non-linear optimization routine described in Section 5.5.2.

Parameter	Value	Units
$f$	60.013	mm
$C_x$	267.198	pixels
$C_y$	255.040	pixels
$\kappa_1$	-0.000103	1/mm <sup>2</sup>
$s_x$	1.079	
$R_x$	-0.084	degrees
$R_y$	0.589	degrees
$R_z$	0.182	degrees
$T_x$	-521.238	mm
$T_y$	-527.935	mm
$T_z$	1581.238	mm
mean UIPE	0.064	pixels
standard deviation UIPE	0.033	pixels
maximum UIPE	0.182	pixels
mean OSE	0.042	mm
standard deviation OSE	0.024	mm
maximum OSE	0.135	mm

Table 5.1: Example of a calibrated fixed camera model

## 5.7 Calibration example

To demonstrate the calibration of a fixed camera model we calibrated the Cosmicar/Panasonic camera system for the lens setting ( $m_f = 2000$ ,  $m_z = 1000$ ,  $m_a = 1500$ ). The calibration data for the model came from two images of the calibration target taken with sensor-to-target ranges of 1.5 m (Fig. 5.4) and 2.5 m (Fig. 5.5). The absolute position of the origin for the world-coordinate system was arbitrarily assigned to be in the target plane at 1.5 m range, approximately 520 mm up and 520 mm to the left of the center of the camera's field of view. The two images provided 186 data points.

Table 5.1 shows the calibrated fixed camera model after the final non-linear optimization step. The small values for the mean UIPE and maximum UIPE indicate that the calibrated camera model does a good job of capturing the lens's 3D to 2D imaging behavior.

Figures 5.6 and 5.7 show plots of the actual error between the measured positions of the reference points from the calibration data and the positions predicted by the calibrated model. In the plots the measured positions are marked with small square boxes. The vectors from each box point to the position of the point's image predicted by the camera model. In both figures the error vectors are magnified 100 $\times$ . The non-random distribution of the direction vectors indicates limitations in the dimensional accuracy of the target and translation stage used to obtain the calibration data.

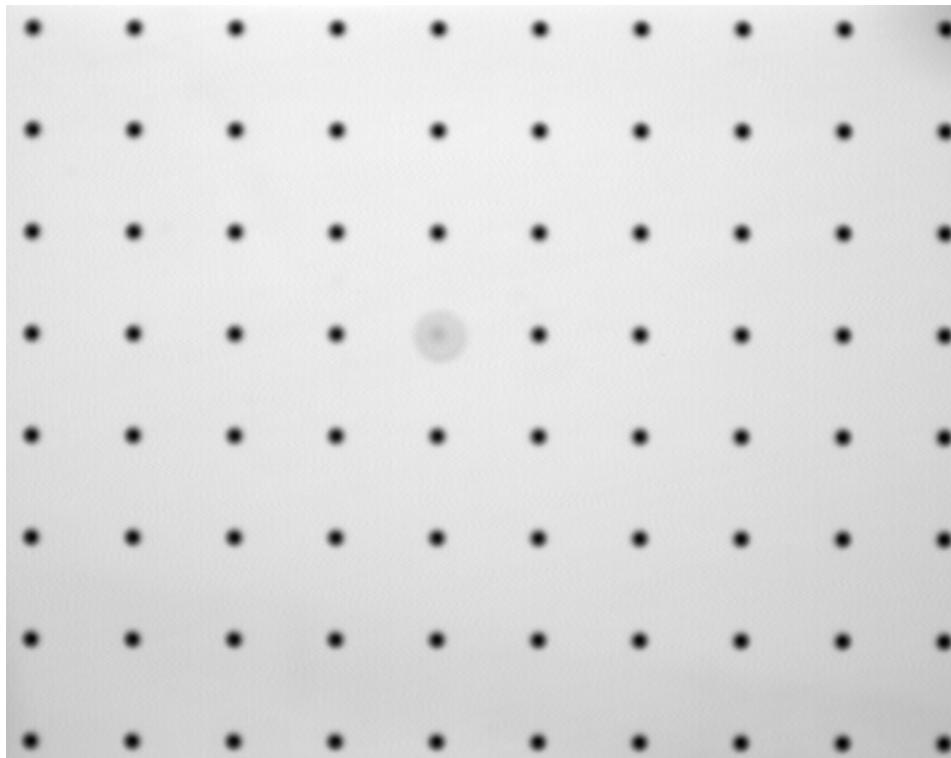


Figure 5.4: Calibration target at 1.5 m

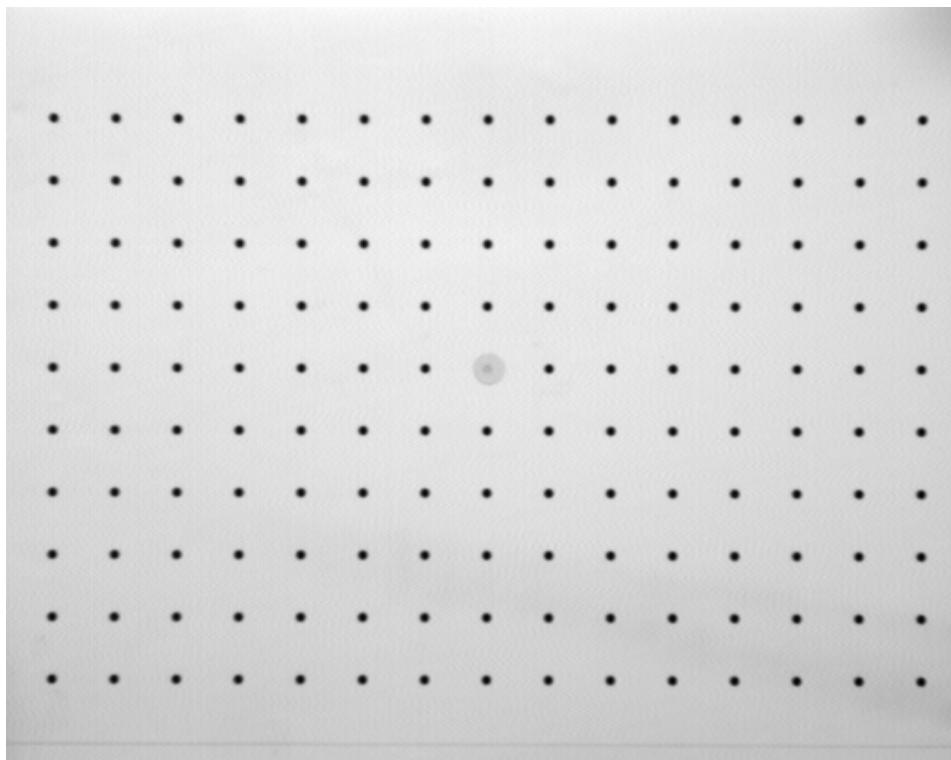


Figure 5.5: Calibration target at 2.5 m

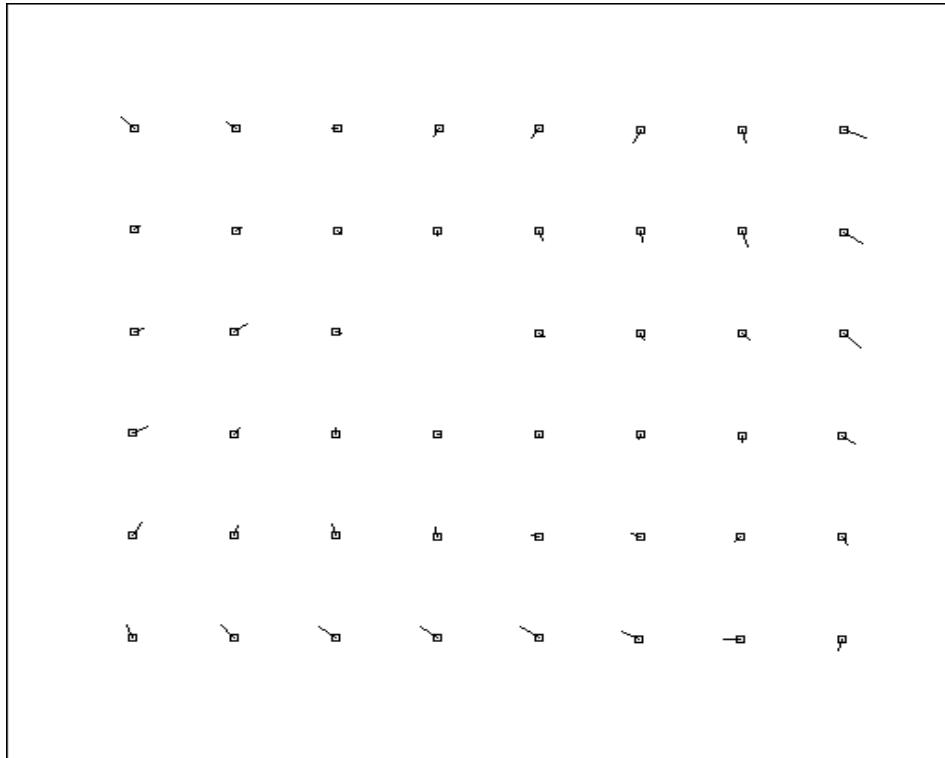


Figure 5.6: Residual error for calibration target at 1.5 m (magnified 100 $\times$ )

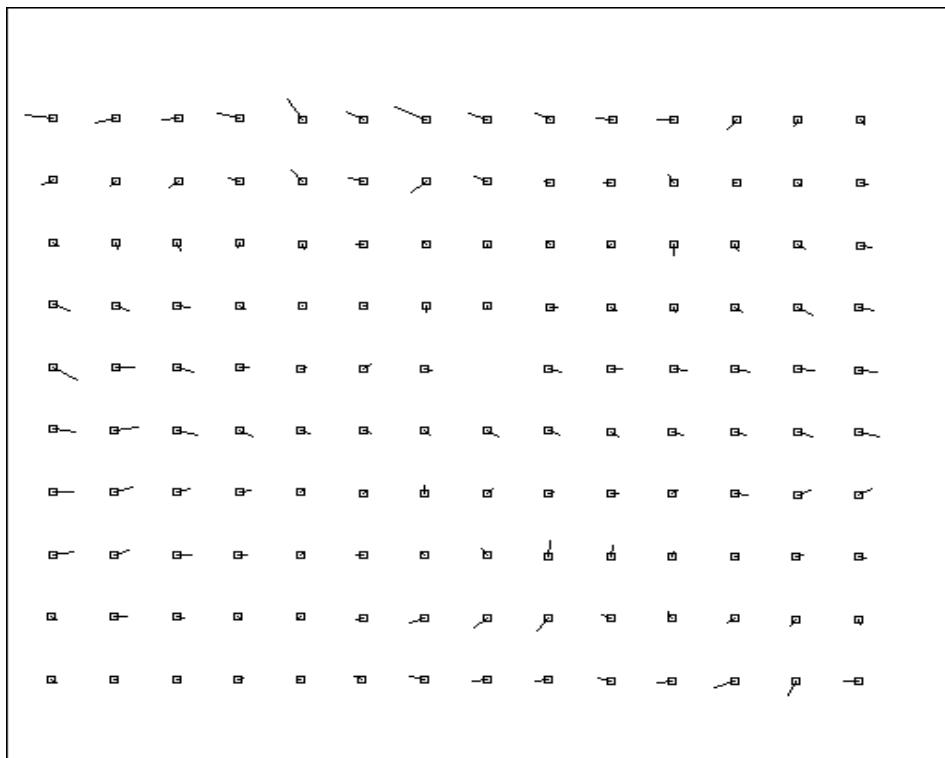


Figure 5.7: Residual error for calibration target at 2.5 m (magnified 100 $\times$ )

## 5.8 Exterior orientation recalibration example

To test our ability to reacquire the camera's extrinsic parameter models after the camera has been moved we took one calibration dataset, changed the camera's exterior orientation (pose), and then took a second dataset. Using the first dataset we obtained a fully calibrated fixed camera model. The intrinsic parameters from this model are then used in a partial calibration with the second dataset to estimate the camera's new extrinsic parameters. The second set of data was then used to obtain a fully calibrated fixed camera model. The test is illustrated graphically in Fig. 5.10.

For the first dataset we actually used the data from Section 5.7. Before taking the second dataset the camera was shifted -100 mm in the camera  $x_w$  coordinate, -100 mm in the camera  $y_w$  coordinate and -100 mm in the camera  $z_w$  coordinate, and then rotated until the center of the field of view was located roughly in the middle of the two target plane positions. The new dataset was obtained from two images of the calibration target taken at ranges of 1.5 m and 2.5 m (Figs. 5.8 and 5.9). The world-coordinate frame for the new dataset was the same as that used for the first dataset. The second dataset of data contains 203 data points.

Column 2 of Table 5.2 shows the results for the full calibration on the data from the first pose. Column 4 shows the results for the full calibration for the second pose. The 20% increase in the mean UIPE for the second set of data is due in part to the lack of flatness of the calibration target. With the first set of data the target is viewed head on and target flatness is less critical. For the second set the viewing angle is more oblique.

Column 3 of Table 5.2 shows the calibration results when the intrinsic parameters from the first set of data are used in a partial calibration to obtain the extrinsic parameters from the second set of data. Comparing the results for the second set of data (columns 3 and 4) we see that the exterior orientation and mean UIPE are virtually the same for the partial and full calibration. This demonstrates that the algorithm in Section 5.6 can accurately reacquire the camera's exterior orientation after the camera has been moved. It also demonstrates that the intrinsic model can be carried between poses.

## 5.9 Summary

In this chapter we presented the fixed camera model that will be the basis of the adjustable camera models developed in the next chapter. We have discussed the formulation of the fixed model along with the algorithms and techniques required to calibrate it for a fixed-parameter camera system. We have also presented performance metrics and validation techniques for the fixed camera model.

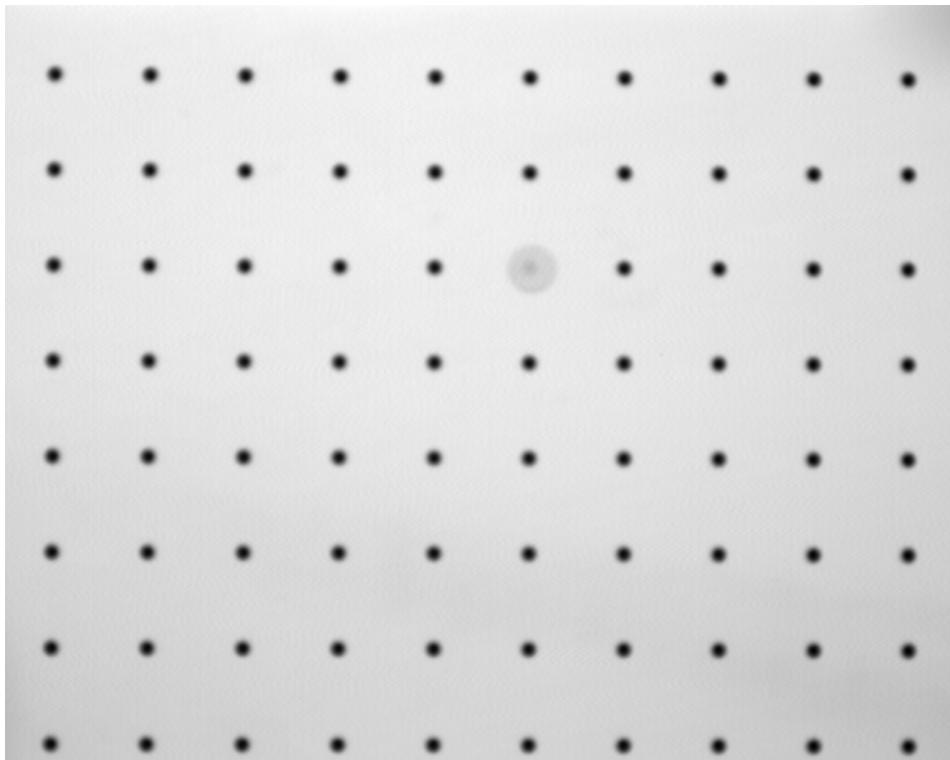


Figure 5.8: Calibration target at 1.5 m with camera shifted

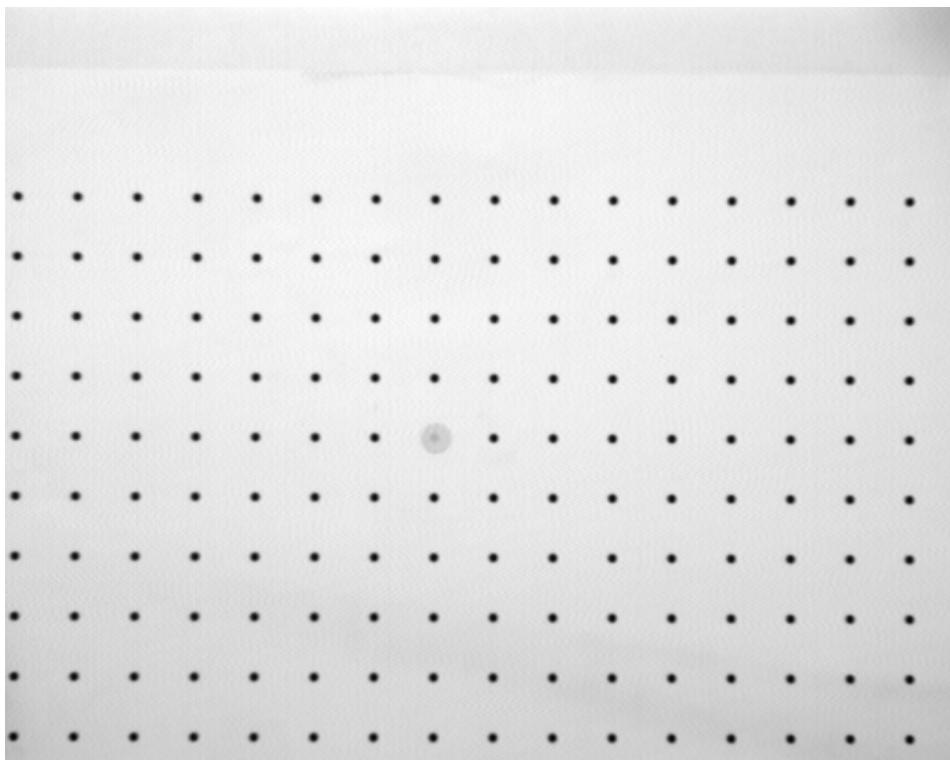


Figure 5.9: Calibration target at 2.5 m with camera shifted

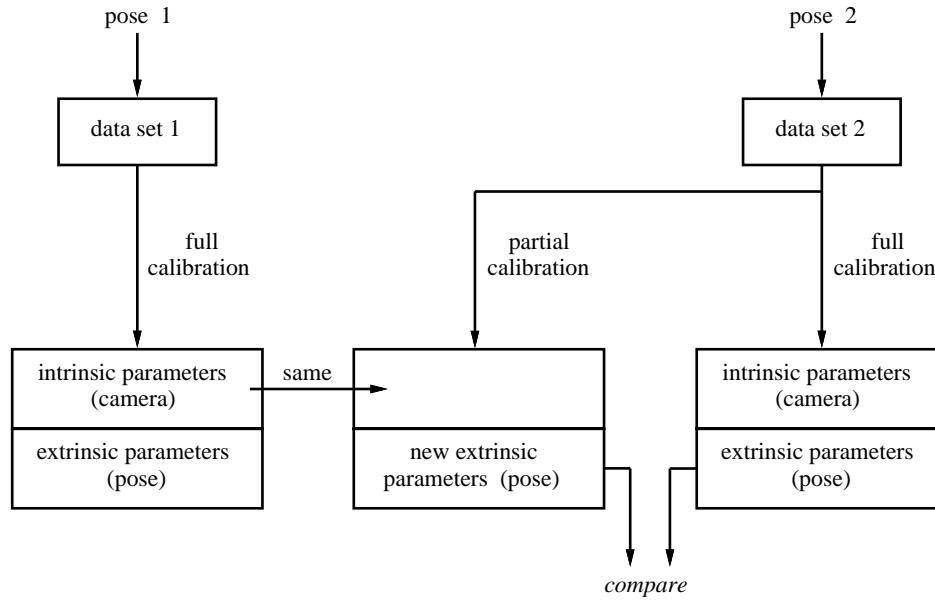


Figure 5.10: Experiment in carrying intrinsic parameters to new pose

Parameter	Pose 1	Pose 2	Pose 2	Units
	full calibration	partial calibration	full calibration	
$f$	60.013	60.013	60.058	mm
$C_x$	267.198	267.198	267.296	pixels
$C_y$	255.040	255.040	254.626	pixels
$\kappa_1$	-0.000103	-0.000103	-0.000096	1/mm <sup>2</sup>
$s_x$	1.079	1.079	1.078	
$R_x$	-0.084	-2.832	-2.838	degrees
$R_y$	0.589	-2.042	-2.044	degrees
$R_z$	0.182	0.303	0.303	degrees
$T_x$	-521.238	-497.003	-497.045	mm
$T_y$	-527.935	-547.358	-547.186	mm
$T_z$	1581.238	1689.919	1690.989	mm
mean UIPE	0.064	0.077	0.076	pixels
standard deviation UIPE	0.033	0.045	0.040	pixels
maximum UIPE	0.182	0.223	0.202	pixels
mean OSE	0.042	0.050	0.049	mm
standard deviation OSE	0.024	0.031	0.029	mm
maximum OSE	0.135	0.153	0.157	mm

Table 5.2: Calibration results from carrying intrinsic parameters to new pose



# Chapter 6

## Adjustable Perspective-Projection Camera Models

In the previous chapter we described Tsai's 11-parameter perspective-projection camera model for a fixed camera system. Based on this model we constructed the adjustable camera model that will be described in this chapter. Two variants of the model are developed. The first is calibrated and tested on both of our camera systems for continuous ranges of focus and zoom. The second is calibrated and tested on one of the systems for a continuous range of aperture.

### 6.1 Definitions

Before we proceed it will be helpful to redefine some notation and introduce some new notation specific to the perspective-projection camera model.

**Lens setting:** A three-tuple containing the control settings for the focus, zoom, and aperture motors on the lens.

$$S = \{m_f, m_z, m_a\}$$

**Calibration data point:** A five-tuple containing the 3D world coordinates of a point and its 2D frame-buffer coordinates.

$$d = \{x_w, y_w, z_w, X_f, Y_f\}$$

**Calibration data set:** A set of calibration data points,  $d_i$ , taken at one lens setting, in one world coordinate system, from one fixed camera position and orientation.

$$D = \{d_0, \dots, d_n\}$$

**Fixed camera model:** An 11-tuple containing the intrinsic and extrinsic parameters for the fixed perspective-projection camera model.

$$M_f = \{f, C_x, C_y, \kappa_1, s_x, R_x, R_y, R_z, T_x, T_y, T_z\}$$

**Adjustable camera model:** A set of 11 parameter models that describe the values of the intrinsic and extrinsic parameters for the fixed perspective-projection camera model at any given lens setting  $S$ .

$$\begin{aligned} M_a(S) = & \{g_f(S), g_{C_x}(S), g_{C_y}(S), g_{\kappa_1}(S), g_{s_x}(S), \\ & g_{R_x}(S), g_{R_y}(S), g_{R_z}(S), g_{T_x}(S), g_{T_y}(S), g_{T_z}(S)\} \end{aligned}$$

**Undistorted image plane error (UIPE):** The magnitude (in pixels) of the difference between the measured position of a point  $d_i$  in the image and the position predicted by a camera model  $M$  (assuming an undistorted image in both cases).

$$\text{UIPE}(M, D) = \sqrt{(\Delta X_{f_u})^2 + (\Delta Y_{f_u})^2}$$

where

$$\begin{aligned} \Delta X_{f_u} &= d_x^{-1}(X_{u_2} - X_{u_1})s_x, \\ \Delta Y_{f_u} &= d_y^{-1}(Y_{u_2} - Y_{u_1}), \end{aligned}$$

$(X_{u_2}, Y_{u_2})$  are calculated from measured position of the point's image  $(X_f, Y_f)$  using (5.6), (5.7) and (5.8);  $(X_{u_1}, Y_{u_1})$  are calculated from the 3D coordinates of the point  $(x_w, y_w, z_w)$  using (5.1), (5.4) and (5.5).

**Mean undistorted image plane error (M\_UIPE):** The average value of the UIPE for model  $M$  and all points  $d_i$  in a dataset  $D$ .

$$\text{M\_UIPE}(M, D) = \frac{1}{n} \sum_{i=1}^n \text{UIPE}(M, d_i)$$

**Sum of the squared undistorted image plane error (SS\_UIPE):** The sum of the square of the UIPE for model  $M$  and all points in a dataset  $D$ .

$$\text{SS\_UIPE}(M, D) = \sum_{i=1}^n [\text{UIPE}(M, d_i)]^2$$

## 6.2 Performance metrics

Our objective was to develop a model of the camera's imaging behavior that "holds calibration" across ranges of lens parameters. By "holds calibration" we mean that the model maintains an acceptable level of accuracy at any lens setting<sup>1</sup>.

---

<sup>1</sup>Since our "ground truth" is limited to the set of calibration data used to develop the model, the best we can do is have a model that "holds calibration" across the available calibration data.

Given calibration data for a particular lens setting, the performance of the adjustable camera model can be expressed using any of the fixed camera model metrics presented in Section 5.3. To be able to compare different adjustable camera models we require an aggregate measure of the model’s performance for all of the calibrated lens settings.

There are many ways to combine the adjustable camera model’s performance statistics at each calibrated lens setting into a set of statistics for all of the settings. If we are interested in the total fit between the adjustable model  $M_a$  and the calibration datasets  $D_i$  at each and every data point, then “per point error” metrics can be used, such as the sum of the sum of the squared undistorted image plane error,

$$\text{SSS\_UIPE} = \sum_{i=1}^m \text{SS\_UIPE}(M_a(S_i), D_i)$$

A drawback with per point error metrics is that the number of data points in each set of calibration data  $D_i$  may vary with lens setting  $S_i$  so that different lens settings receive different weightings in the performance metric.

If we are more concerned with the performance of the adjustable model  $M_a$  at each lens setting  $S_i$ , then we need a metric that is invariant to the number of data points involved, for example the M\_UIPE. One useful performance metric of this type is the mean of the mean undistorted image plane error,

$$\text{MM\_UIPE} = \frac{1}{m} \sum_{i=1}^m \text{M\_UIPE}(M_a(S_i), D_i)$$

For the following adjustable models we base calibration decisions (i.e. initial fitting sequence and iterative refinement) on the SSS\_UIPE metric because it gives the same weight to every data point.

For displays of the adjustable model’s performance we use the MM\_UIPE metric because it has a more direct (and intuitive) relationship with the model’s accuracy in a given application.

## 6.3 Fujinon/Photometrics model for focus and zoom

In this section we describe the adjustable perspective-projection camera model we developed for the Fujinon/Photometrics camera system for ranges of focus and zoom.

### 6.3.1 Operating range

For the operating range for this model we chose a focus range of  $1500 \leq m_f \leq 4000$  motor units, which corresponds roughly to a focused distance of 1.5 m to 2.5 m. The correspondence is not exact as the lens’s focused distance is also affected by the zoom and aperture controls.

For the zoom we chose a range of  $1500 \leq m_z \leq 4000$  motor units, which corresponds to focal lengths from approximately 130 mm down to 45 mm. For the aperture we used a fixed setting of 380 motor units, which corresponds roughly to  $f/16$ .

### 6.3.2 Sampling strategy

Figures 6.1 and 6.2 show the  $x$  and  $y$  image coordinates of an autocollimated laser plotted against the focus and zoom settings for the lens. The plots show a relatively smooth variation in the laser's position across the full operating space. Thus, for the sampling strategy for this lens we arbitrarily chose a regular  $11 \times 11$  sampling of focus and zoom settings for a total of 121 separate settings ( $S_1, \dots, S_{121}$ ) across the operating space for our camera model. From (4.4) the largest bivariate polynomial that can be fit to the 121 data points is 14th order (120 coefficients).

### 6.3.3 Calibration data

Calibration data for the adjustable perspective-projection camera model was obtained using the target described in Section 5.4. At each sample position in the camera operating space three images of the target were taken at ranges of 1.5 m, 2.0 m and 2.5 m between the target and the camera's sensor plane. For each set of images the relative 3D coordinates  $(x_w, y_w, z_w)$  of the reference points on the target are known from their grid position in the target plane and from the position of the target plane along the translation stage. The world coordinate system's origin was chosen to be in the target plane at the 2.5 m range, approximately 340 mm up and 320 mm to the left of the center of the camera's field of view. The  $(X_f, Y_f)$  positions of the dots in each image were measured to sub-pixel accuracy using the procedure described in Appendix C. For the 121 different lens settings ( $S_1, \dots, S_{121}$ ) we obtained 121 sets of calibration data ( $D_1, \dots, D_{121}$ ). Each set contained between 110 and 429 calibration data points.

### 6.3.4 Initial fixed camera model calibration

Using the procedure described in Section 5.5 we calibrated fixed camera models  $(M_{f_1}, \dots, M_{f_{121}})$  for each set of calibration data  $(D_1, \dots, D_{121})$ . Figures 6.3 through 6.14 show the 11 fixed model parameters and the M\_UIPE plotted against the focus and zoom motor positions. Despite the apparent noise in many of the model terms, the M\_UIPE for the lens lies between 0.090 pixels and 0.123 pixels across the full operating space chosen for the camera model. The MM\_UIPE over the operating space is 0.099 pixels.

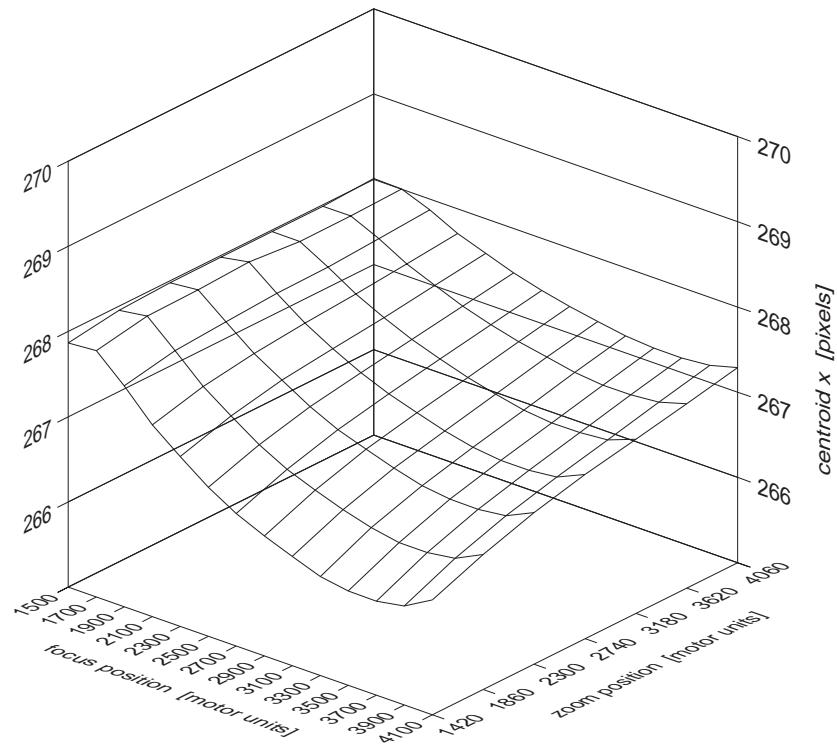


Figure 6.1: Variation in X coordinate of autocollimated laser's image with focus and zoom

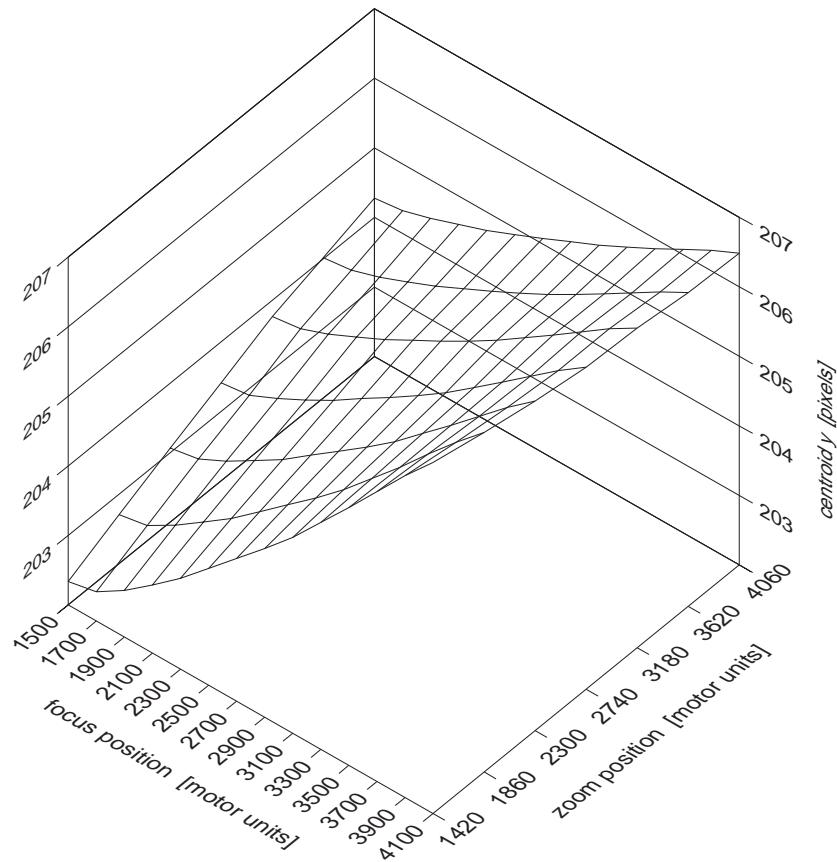


Figure 6.2: Variation in Y coordinate of autocollimated laser's image with focus and zoom

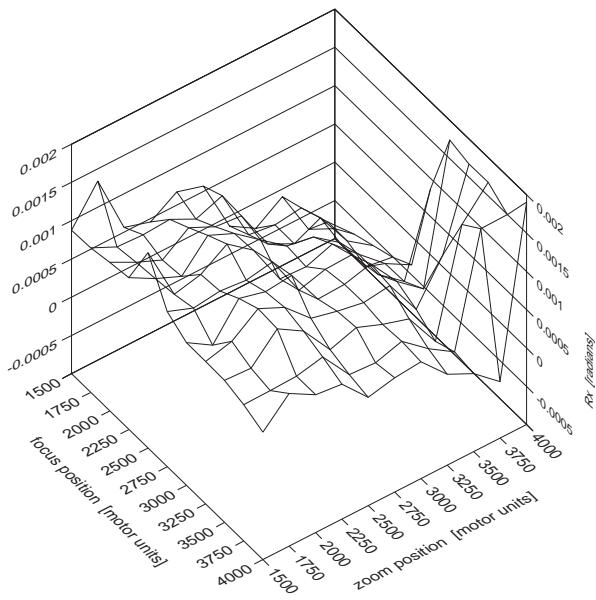


Figure 6.3: Fujinon fixed camera model  $R_x$  versus focus and zoom motors

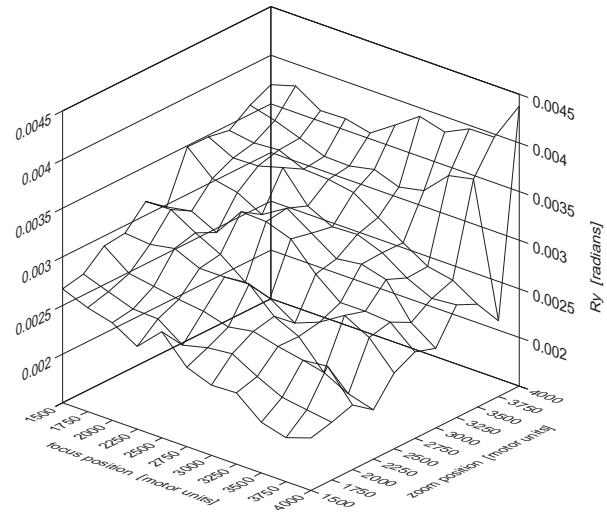


Figure 6.4: Fujinon fixed camera model  $R_y$  versus focus and zoom motors

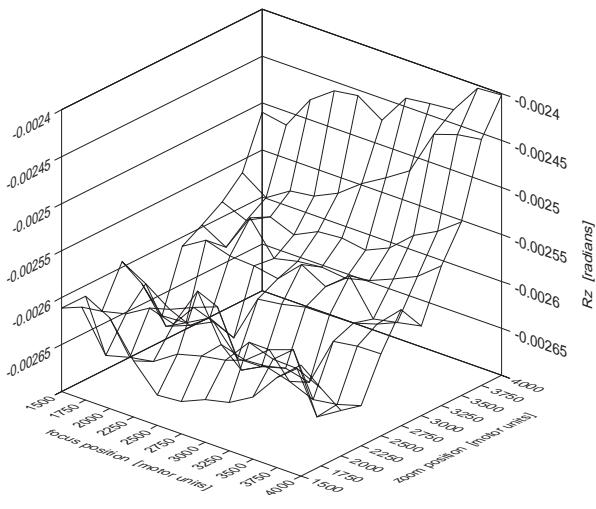


Figure 6.5: Fujinon fixed camera model  $R_z$  versus focus and zoom motors

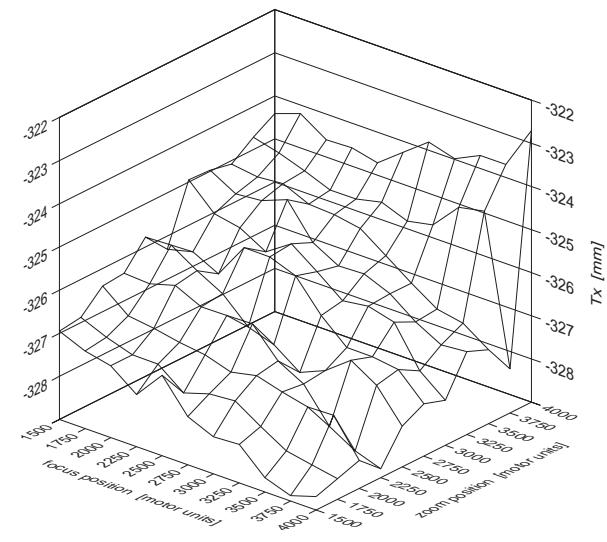


Figure 6.6: Fujinon fixed camera model  $T_x$  versus focus and zoom motors

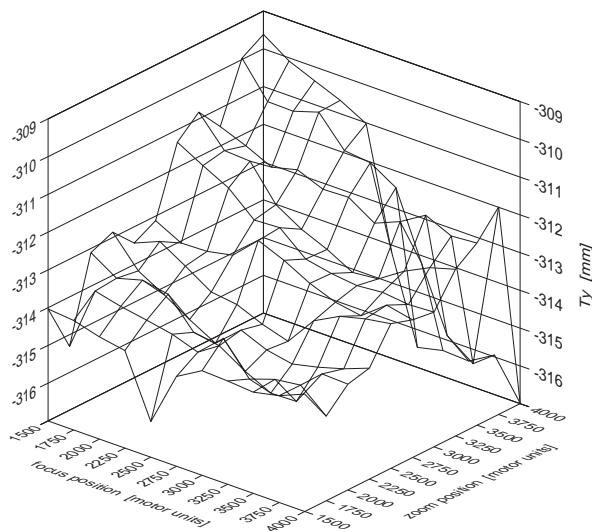


Figure 6.7: Fujinon fixed camera model  $T_y$  versus focus and zoom motors

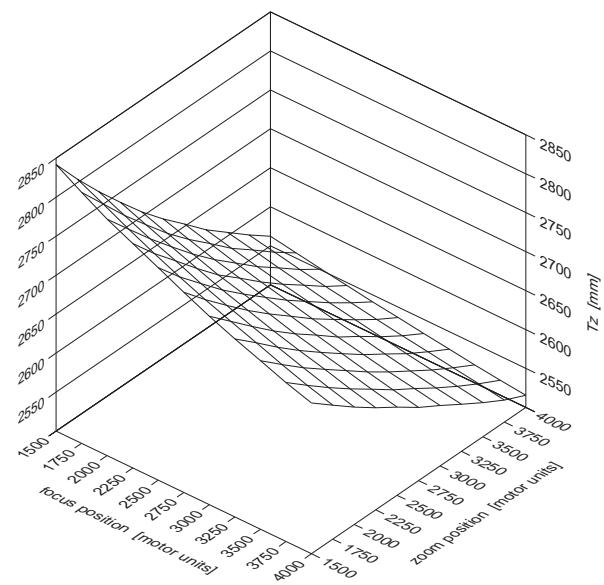


Figure 6.8: Fujinon fixed camera model  $T_z$  versus focus and zoom motors

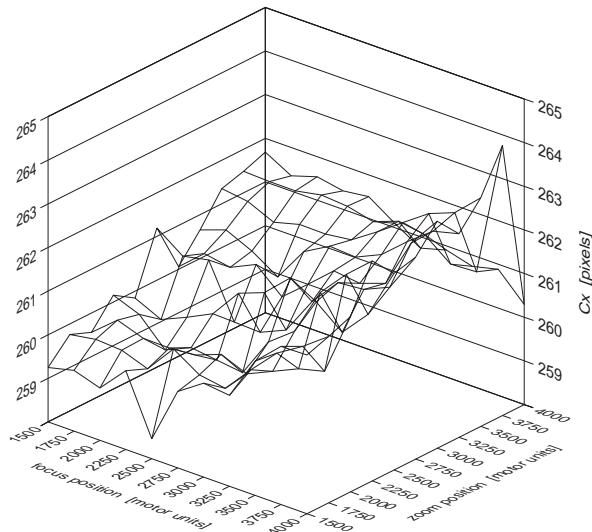


Figure 6.9: Fujinon fixed camera model  $C_x$  versus focus and zoom motors

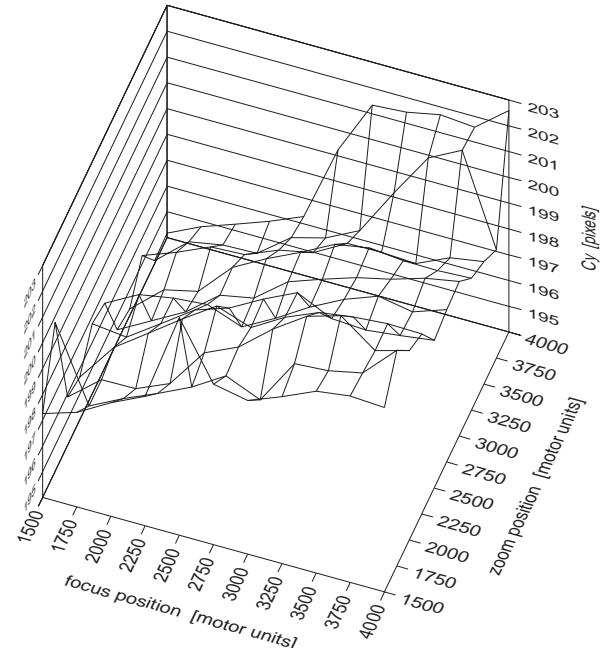


Figure 6.10: Fujinon fixed camera model  $C_y$  versus focus and zoom motors

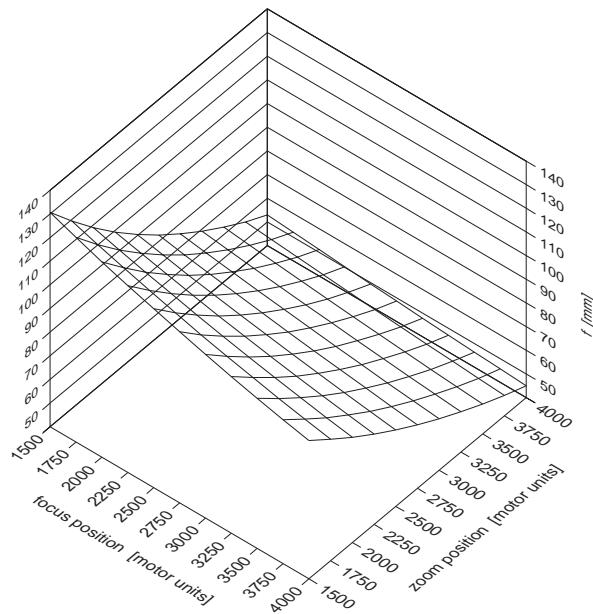


Figure 6.11: Fujinon fixed camera model  $f$  versus focus and zoom motors

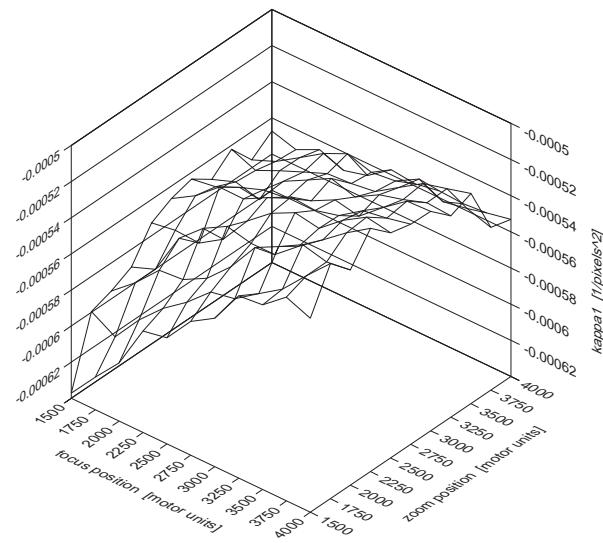


Figure 6.12: Fujinon fixed camera model  $\kappa_1$  versus focus and zoom motors

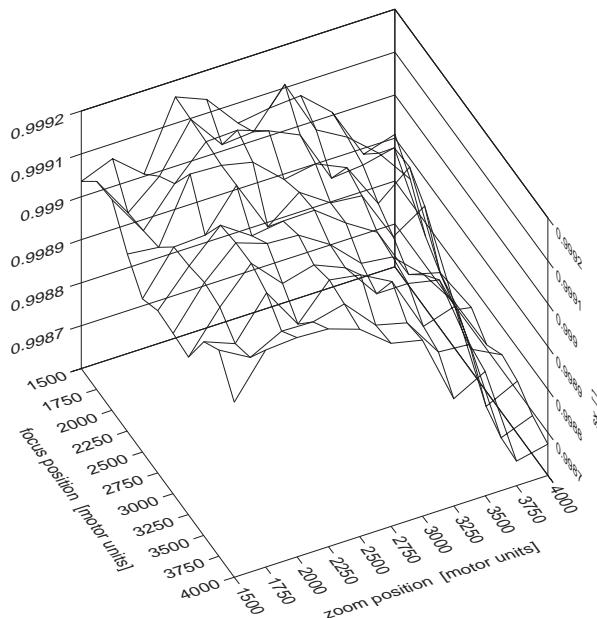


Figure 6.13: Fujinon fixed camera model  $s_x$  versus focus and zoom motors

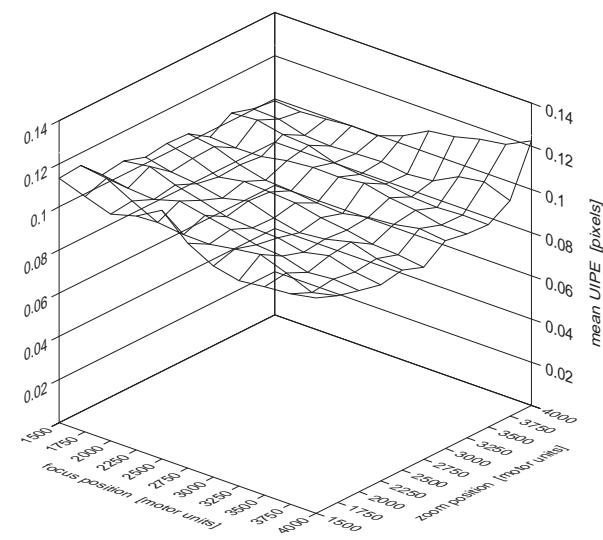


Figure 6.14: Fujinon fixed camera model M\_UIPE versus focus and zoom motors

### 6.3.5 Formulating the parameter models

For the parameter models we used bivariate polynomial functions. Ideally the noise in the parameter values would be zero mean Gaussian. Unfortunately the parameter values are determined using an iterative non-linear optimization on a criterion surface that, in practice, has many local minima. As a result, the fixed camera models are multi-valued. That is, for any given set of calibration  $D$ , the fixed camera calibration can potentially produce several different sets of fixed camera model parameters. The set that is found depends on the noise in the data and on the initial conditions used in the non-linear optimization. Thus, the variation in the values of the fixed model parameters is not due to Gaussian, zero mean, constant standard deviation noise. This has two implications for fitting the parameter models. The first is that least-squares error fitting is not a maximum likelihood estimator for the surface. Even so, since the least-squares error fitting can be accomplished with a direct non-iterative approach this is our preferred fitting method. Tests using much slower but more robust surface fitting techniques using local M-Estimates[38] showed no significant improvement in the performance of the final adjustable camera model.

The second implication of the non-Gaussian noise in the parameter values is that we cannot use a Chi-Squared test to determine how high a polynomial order to use for each parameter model. Instead we chose the model order based on design requirements for the adjustable camera model and on empirical results. To make the final adjustable model easier to use the  $R_x$ ,  $R_y$ ,  $R_z$ ,  $T_x$ , and  $T_y$  parameters were modeled with zero order polynomials. Since the  $s_x$  parameter was related to uncertainties in fixed elements in the camera system, it was also modeled with a zero order polynomial. We tried a wide range of polynomial orders for the remaining parameter models. The final values we used represented an arbitrary tradeoff between increased complexity and improved performance for the final adjustable model. In section 6.7 we discuss alternate strategies for choosing parameter model orders.

Table 6.1 lists the parameters, the orders chosen for their parameter models and the rationale for the choice of order. The Fujinon/Photometrics adjustable camera model required a total of  $(6 \times 1) + (1 \times 6) + (4 \times 21) = 96$  coefficients for the parameter models.

### 6.3.6 Fitting the parameter models

To fit the parameter models to the calibration data we used the ascending order, greedy within order sequencing described in Section 4.2.3. Table 6.2 shows the sequence in which the parameter models were fit, along with the MM\_UIPE, maximum UIPE, and SSS\_UIPE statistics for the adjustable camera models at each stage. The first entry in the table is for the original unfitted fixed camera model parameters. Steps 1 to 11 are for the initial surface fitting sequence. Steps 12 and 13 are for iterative refinement. The last entry is for the final adjustable camera model.

Figures 6.15 through 6.19 show the final adjustable camera model surfaces for the parameters having second- and fifth-order polynomials. While the final  $f$ ,  $T_z$ , and  $\kappa_1$  models are all similar in shape to the original unfitted parameters, the remaining models are all rather

Parameter	Polynomial Order	Reason
$s_x$	0	Changing the camera's image formation process should not change the relative scale factor between the $x$ and $y$ axes so we only permit a constant for this parameter.
$R_x \ R_y \ R_z \ T_x \ T_y$	0	For ease of use of the adjustable camera model we would like the position and orientation of the camera's coordinate frame relative to the world coordinate frame to remain unchanged as the lens parameters are varied so we only permit constants for these parameters.
$T_z$	5	Changing the lens's hardware configuration redistributes the optical components along the camera's optical axis shifting the positions of the lens's front and rear nodal points. This in turn changes the separation of the origins of the world and camera coordinate systems. Empirically we find that a fifth-order polynomial works well.
$f$	5	While primarily a function of the zoom actuator, $f$ is also a function of the focus, aperture and image band. Empirically we find that a fifth-order polynomial works well.
$C_x \ C_y$	5	Changing the lens's hardware configuration changes the alignment of the lens's optical components causing the camera's field of view to shift. Empirically we find that a fifth-order polynomial works well.
$\kappa_1$	2	Changing the optical configuration of the lens changes the factors causing radial lens distortion. Empirically we find that a second-order polynomial works well.

Table 6.1: Choice of polynomial orders for Fujinon/Photometrics parameter models

Fitting Step	Parameter	Polynomial Order	MM_UIPE [pixels]	max UIPE [pixels]	SSS_UIPE [pixels <sup>2</sup> ]
unfitted fixed models			0.099457	0.707898	404.268858
1	$R_z$	0	0.099341	0.719303	403.835679
2	$s_x$	0	0.099357	0.735674	403.760237
3	$T_y$	0	0.099868	0.732523	408.106362
4	$T_x$	0	0.101027	0.736067	417.610304
5	$R_y$	0	0.102501	0.735245	426.124820
6	$R_x$	0	0.109440	0.770797	467.337702
7	$\kappa_1$	2	0.109468	0.770712	467.920357
8	$f$	5	0.109449	0.771996	467.958315
9	$T_z$	5	0.109530	0.772310	469.284819
10	$C_y$	5	0.109681	0.774827	470.673135
11	$C_x$	5	0.110829	0.776854	480.119280
12	$C_x$	5	0.109681	0.774828	470.673046
13	$R_x$	0	0.107671	0.776371	458.842669
final adjustable model			0.107671	0.776371	458.842669

Table 6.2: Fitting sequence for Fujinon/Photometrics parameter models for focus and zoom

different than their unfitted data.

Figure 6.20 shows the final M\_UIPE for the adjustable camera model. The final MM\_UIPE across the full range of lens settings is 0.108 pixels, which is a 9% increase over the average of 0.099 pixels for the unfitted fixed camera models. Figure 6.21 shows the difference between the M\_UIPE for the final adjustable camera model and the M\_UIPE for the unfitted fixed camera models. For  $m_z > 3750$  and  $m_f > 2000$  the adjustable camera model's M\_UIPE is actually better than that of the unfitted fixed camera models.

If we were to calibrate an adjustable camera model for another copy of the same lens the shapes of the  $C_x$  and  $C_y$  surfaces would be different due to the different optical misalignments in each lens. However, the shapes of the  $f$ ,  $T_z$ , and  $\kappa_1$  surfaces for both lenses would be similar, as would the positions of any image properties discontinuities in the lenses' control space.

### 6.3.7 Validating the adjustable model

To test the generality of the adjustable camera model and the repeatability of the camera hardware we used two independent sets of calibration data, Set 1 and Set 2. The two sets were taken one after the other with the exterior orientation of the camera unchanged between sets. Each set was used to generate an adjustable camera model. The two models were then tested on both sets of calibration data.

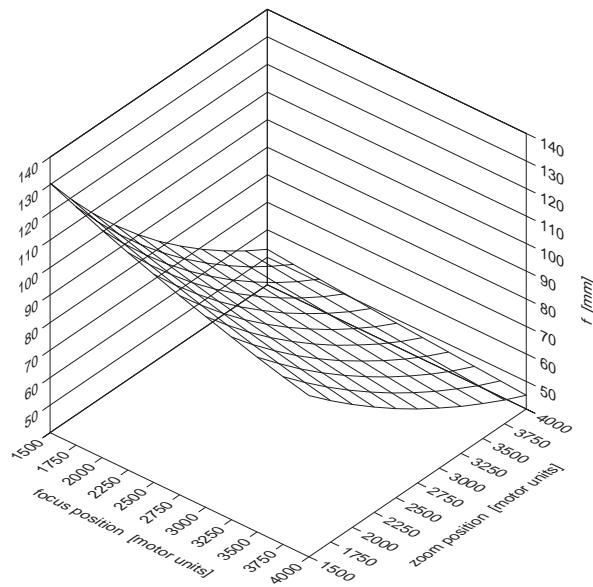


Figure 6.15: Fujinon adjustable camera model  $f$  versus focus and zoom motors

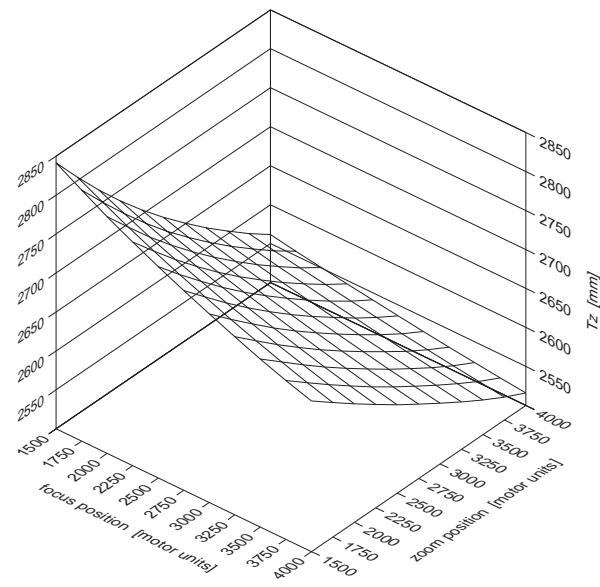


Figure 6.16: Fujinon adjustable camera model  $T_z$  versus focus and zoom motors

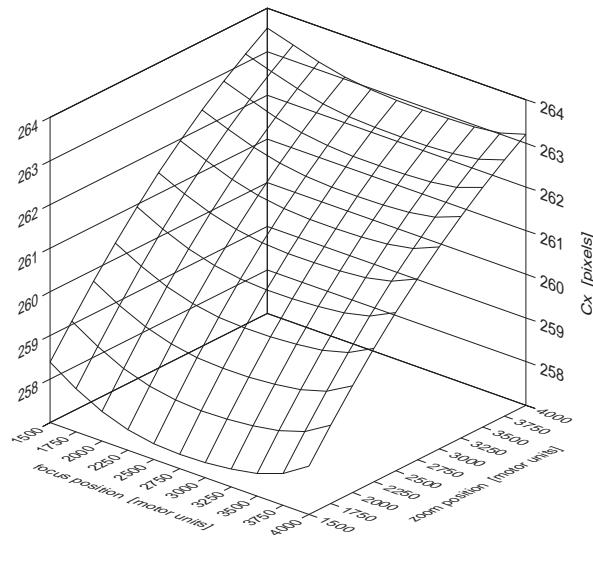


Figure 6.17: Fujinon adjustable camera model  $C_x$  versus focus and zoom motors

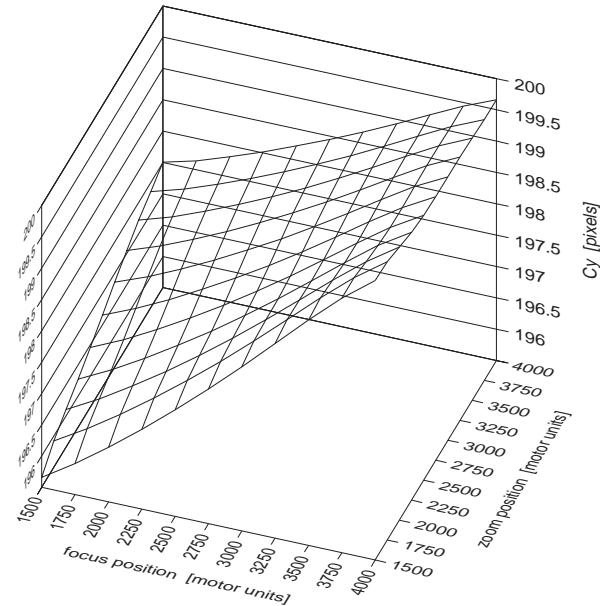


Figure 6.18: Fujinon adjustable camera model  $C_y$  versus focus and zoom motors

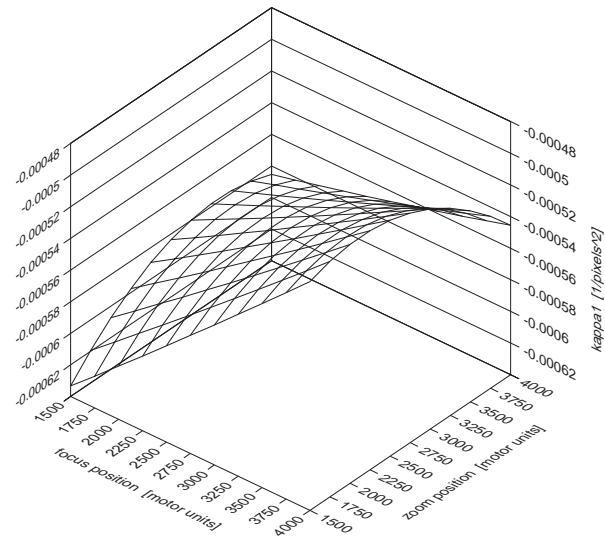


Figure 6.19: Fujinon adjustable camera model  $\kappa_1$  versus focus and zoom motors

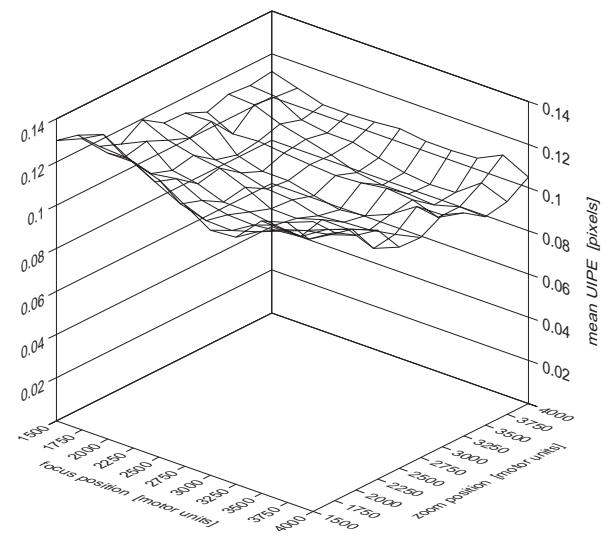


Figure 6.20: Fujinon adjustable camera model  $M_{UIPE}$  versus focus and zoom

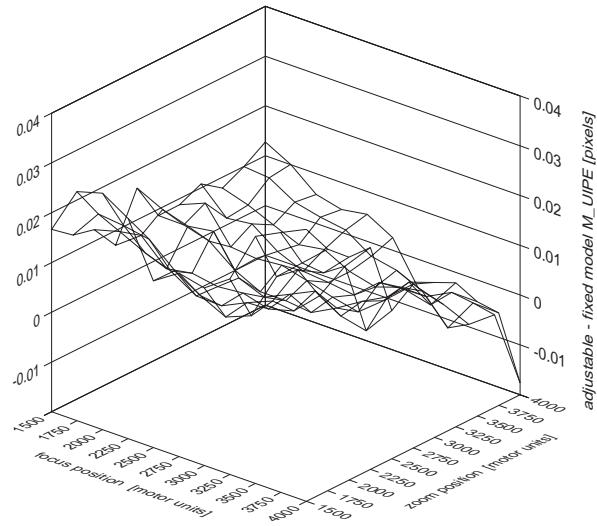


Figure 6.21: Difference between the final adjustable model  $M_{UIPE}$  and the initial fixed model  $M_{UIPE}$

Parameter Model	Set 1	Set 2	Units
$R_x$	0.029197	0.028341	degrees
$R_y$	0.163823	0.164548	degrees
$R_z$	-0.148578	-0.149301	degrees
$T_x$	-326.50	-326.48	mm
$T_y$	-313.22	-313.19	mm
$s_x$	0.998984	0.999096	

Table 6.3: Zero-order parameter models for two sets of calibration data from the same pose

	Test 1	Test 2	Test 3	Test 4	Units
data used for model	Set 1	Set 2	Set 1	Set 2	
data used for testing	Set 1	Set 2	Set 2	Set 1	
MM_UIPE	0.107671	0.107606	0.112449	0.112740	pixels
max UIPE	0.776371	0.763927	0.755688	0.785734	pixels
SSS_UIPE	458.842669	457.688740	494.742462	503.272373	pixels <sup>2</sup>

Table 6.4: Performance of two adjustable models on two sets of calibration data from the same pose

Table 6.3 shows the values of the zero-order parameter models calculated for both sets of data. Table 6.4 shows the calibration statistics for the two adjustable camera models tested against both sets of calibration data. The close correspondence between the zero-order models as well as the nearly identical statistics for the two adjustable models tested against their own calibration data, indicates that the camera hardware is repeatable. The less than 5% increase in the MM\\_UIPE when the adjustable models calibrated with one set of data are tested against the alternate set of data indicate that the adjustable models generalize across different sets of calibration data.

Figures 6.22 through 6.25 show plots of the M\\_UIPE for both camera models tested against both sets of calibration data. The shapes are not significantly different from one another, which is yet another indication that the adjustable models are general.

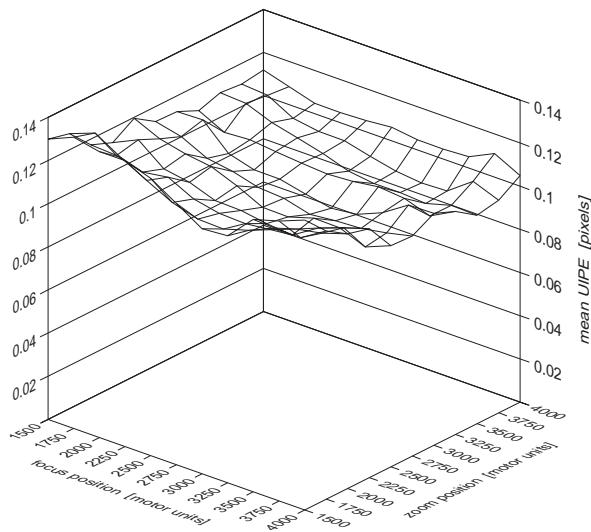


Figure 6.22: M\_UIPE for model calibrated with dataset 1 and tested on dataset 1

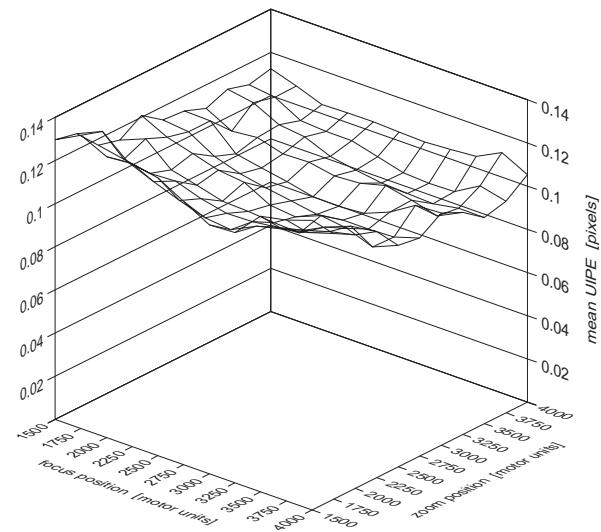


Figure 6.23: M\_UIPE for model calibrated with dataset 2 and tested on dataset 2

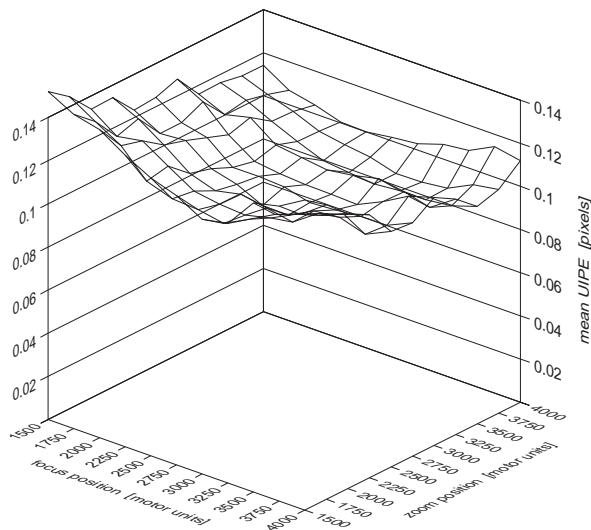


Figure 6.24: M\_UIPE for model calibrated with dataset 1 and tested on dataset 2

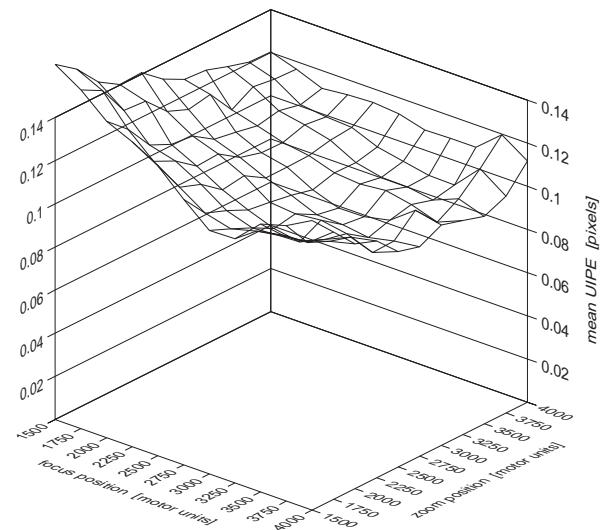


Figure 6.25: M\_UIPE for model calibrated with dataset 2 and tested on dataset 1

## 6.4 Recalibrating exterior orientation

Having spent a great deal of time and effort to produce an adjustable camera model for the lens, the next obvious question is how can it be used. As with the fixed camera model, when the camera system is moved to a new pose the adjustable model's interior orientation functions ( $g_f, g_{C_x}, g_{C_y}, g_{C_z}, g_{s_x}$ ) will be unaffected<sup>2</sup>. However, the exterior orientation of the camera system ( $R_x, R_y, R_z, T_x, T_y, T_z$ ) will have to be recomputed for the new pose. By design our adjustable camera model was built with zero-order functions for the first five exterior orientation parameters  $R_x, R_y, R_z, T_x$ , and  $T_y$ . The only interaction between the camera's exterior orientation and the lens settings is through the  $g_{T_z}(m_f, m_z)$  function. To deal with this interaction we define a new function,

$$\begin{aligned} g'_{T_z}(m_f, m_z) &= T_{z_0} + [g_{T_z}(m_f, m_z) - g_{T_z}(m_{f_0}, m_{z_0})] \\ &= T_{z_0} + \Delta T_z(m_f, m_z, m_{f_0}, m_{z_0}) \end{aligned}$$

which separates  $g_{T_z}$  into a fixed exterior orientation component,  $T_{z_0}$ , and a variable interior orientation component,  $\Delta T_z$ . The fixed component,  $T_{z_0}$ , is estimated along with the other five exterior orientation constants when the lens is set to a base setting,  $(m_{f_0}, m_{z_0})$ . For more precise estimates of the new pose additional base settings can be used. The variable component,  $\Delta T_z$ , accounts for the shift of the lens's principal point along the camera coordinate frame's  $z$  axis, relative to the base lens setting. Figure 6.26 illustrates these relationships in the 2D  $xz$  camera coordinate plane.

We defer a demonstration of the exterior orientation recalibration until we have developed an adjustable camera model for the Cosmicar/Panasonic system.

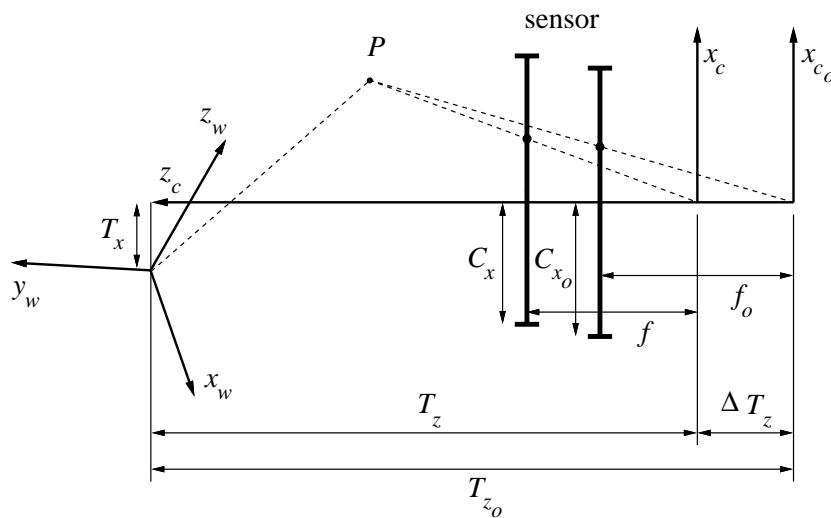


Figure 6.26: Extrinsic (and intrinsic) parameter changes with lens settings

---

<sup>2</sup>The adjustable camera model can only be guaranteed to be accurate over the range of depths and camera parameters that the calibration data covered.

## 6.5 Cosmicar/Panasonic model for focus and zoom

In this section we describe the adjustable perspective-projection camera model we developed for the Cosmicar/Panasonic camera system for ranges of focus and zoom.

### 6.5.1 Operating range

For the operating range for this model we chose a focus range of  $1000 \leq m_f \leq 3000$  motor units, which corresponds roughly to a focused distance of 1.3 m to 2.8 m. For the zoom we chose a range of  $500 \leq m_z \leq 1500$  motor units, which corresponds to focal lengths from approximately 83 mm down to 45 mm. For the aperture we used a fixed setting of 1500 motor units, which corresponds roughly to  $f/7$ .

### 6.5.2 Sampling strategy

Figures 6.27 and 6.28 show the  $x$  and  $y$  coordinates for an autocollimated laser plotted against the focus and zoom motor settings for the lens. In contrast to the plots for the Fujinon lens (Figs. 6.1 and 6.2), the plots for the Cosmicar lens show a significant discontinuity along the zoom axis at approximately  $m_z = 1200$ . To model this discontinuity properly we would have had to increase the amount of sampling around this region of the operating space. However, for this example we consciously ignored the discontinuity and chose a regular  $5 \times 5$  sampling of focus and zoom motor settings for a total of 25 separate lens settings ( $S_1, \dots, S_{25}$ ) across the camera's operating space. From (4.4) the largest bivariate polynomial that can be fit to the 25 data points is 5th order (21 coefficients).

### 6.5.3 Calibration data

Calibration data for the adjustable perspective-projection camera model was obtained using the target described in Section 5.4. At each sample position in the camera operating space two images of the target were taken at ranges of 1.5 m and 2.5 m between the target and the camera's sensor plane. For each set of images the relative 3D coordinates  $(x_w, y_w, z_w)$  of the reference points on the target are known from their grid position in the target plane and from the position of the target plane along the translation stage. The world coordinate system's origin was chosen to be in the target plane at the 1.5 m range, approximately 510 mm up and 510 mm to the left of the center of the camera's field of view. The  $(X_f, Y_f)$  positions of the dots in each image were measured to sub-pixel accuracy using the procedure described in Appendix C. For the 25 different lens settings ( $S_1, \dots, S_{25}$ ) we obtained sets of calibration data ( $D_1, \dots, D_{25}$ ). Each set contained between 122 and 272 calibration data points.

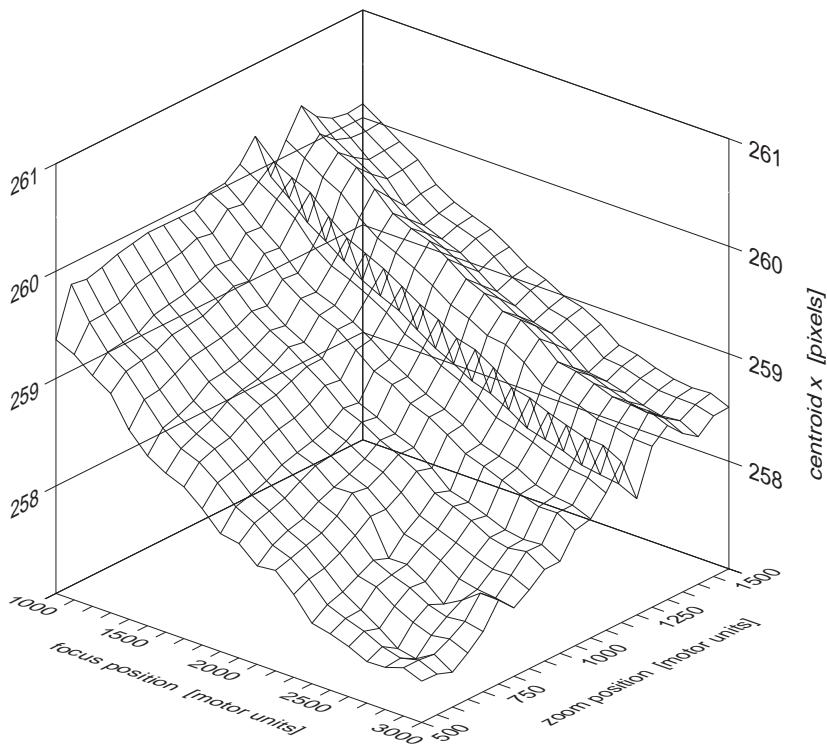


Figure 6.27: Variation in X coordinate of autocollimated laser's image with focus and zoom

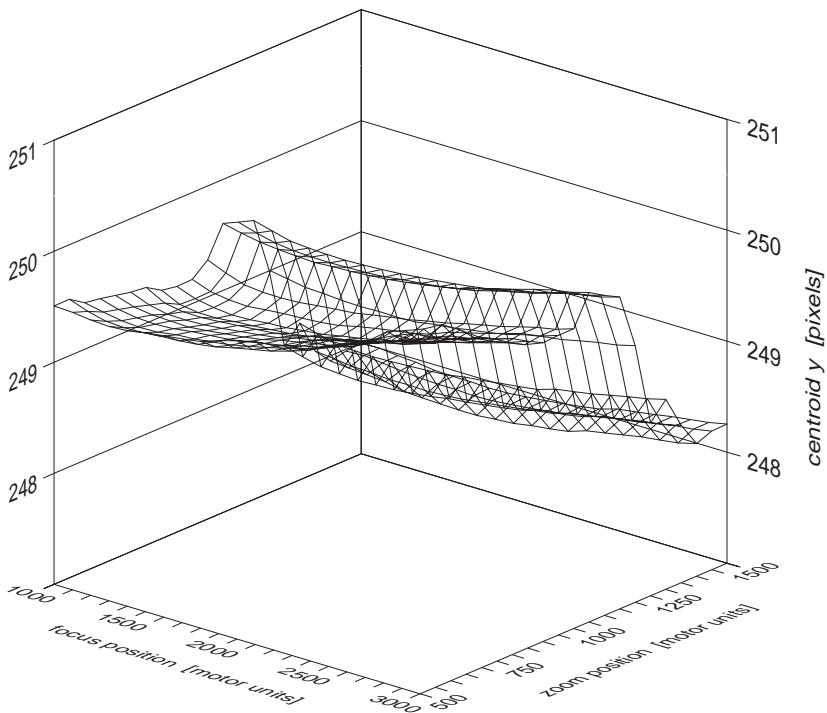


Figure 6.28: Variation in Y coordinate of autocollimated laser's image with focus and zoom

### 6.5.4 Initial fixed camera model calibration

Using the procedure described in Section 5.5 we calibrated fixed camera models ( $M_{f_1}, \dots, M_{f_{25}}$ ) for each set of calibration data ( $D_1, \dots, D_{25}$ ). Figures 6.29 through 6.40 show the 11 fixed model parameters and the M\_UIPE plotted against the focus and zoom motor positions. Despite the apparent noise in many of the model terms, the M\_UIPE for the lens lies between 0.064 pixels and 0.088 pixels across the full operating space chosen for the camera model. The MM\_UIPE over the operating space is 0.076 pixels.

### 6.5.5 Formulating the parameter models

The rationale for the parameter model formulations for the Cosmicar/Panasonic adjustable camera model was the same as that used for the Fujinon/Photometrics camera. We used zero-order polynomial models for the  $s_x$ ,  $R_x$ ,  $R_y$ ,  $R_z$ ,  $T_x$ , and  $T_y$  terms, a second-order polynomial for the  $\kappa_1$  term, and fourth-order polynomials for the  $f$ ,  $C_x$ ,  $C_y$ , and  $T_z$  terms. The Cosmicar/Panasonic adjustable camera model required a total of  $(6 \times 1) + (1 \times 6) + (4 \times 15) = 72$  coefficients for the parameter models.

### 6.5.6 Fitting the parameter models

To fit the parameter models to the calibration data we used the same ascending order, greedy within order sequencing strategy used for the Fujinon/Photometrics model. Table 6.5 shows the sequence in which the parameter models were fit, along with the MM\_UIPE, maximum UIPE, and SSS\_UIPE statistics for the adjustable camera models at each stage. The first entry in the table is for the original unfitted fixed camera model parameters. Steps 1 to 11 are for the initial surface fitting sequence. Steps 12, 13, and 14 are for iterative refinement. The last entry is for the final adjustable camera model.

Figures 6.41 through 6.45 show the final adjustable camera model surfaces for the parameters having second- and fourth- order polynomial models. As with the Fujinon/Photometrics camera the final  $f$ ,  $T_z$ , and  $\kappa_1$  parameter models for the Cosmicar/Panasonic camera are all similar in shape to the original unfitted parameters, while the remaining parameter models are all rather different than their unfitted data.

Figure 6.46 shows the final M\_UIPE for the adjustable camera model. The final MM\_UIPE over the full camera operating space is 0.078 pixels, which is a 3% increase over the average of 0.076 pixels for the unfitted fixed camera models. Figure 6.47 shows the difference between the M\_UIPE for the final adjustable camera and the M\_UIPE for the unfitted fixed camera models.

Figures 6.43 and 6.44 contain the same discontinuity in the imaging behavior that was found in the initial survey of the camera model's operating space with the autocollimated laser (Figs. 6.27 and 6.28).

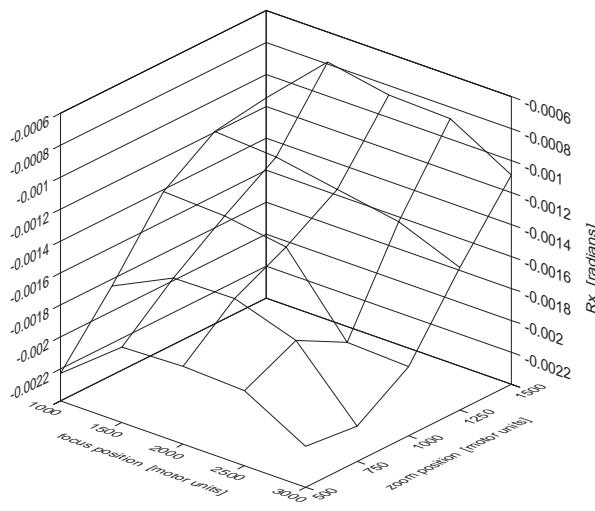


Figure 6.29: Cosmicar fixed camera model  $R_x$  versus focus and zoom motors

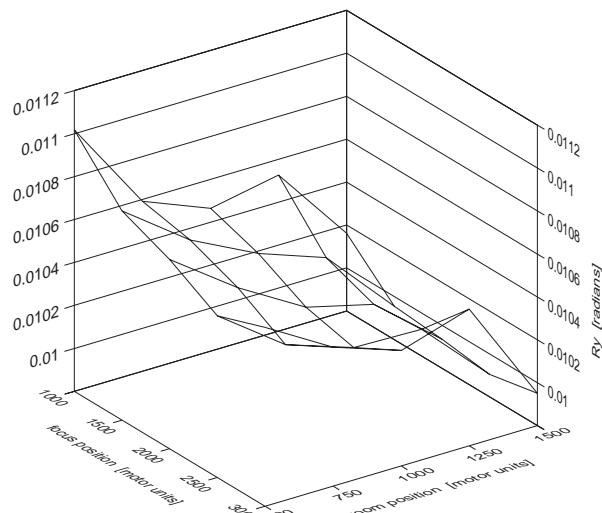


Figure 6.30: Cosmicar fixed camera model  $R_y$  versus focus and zoom motors

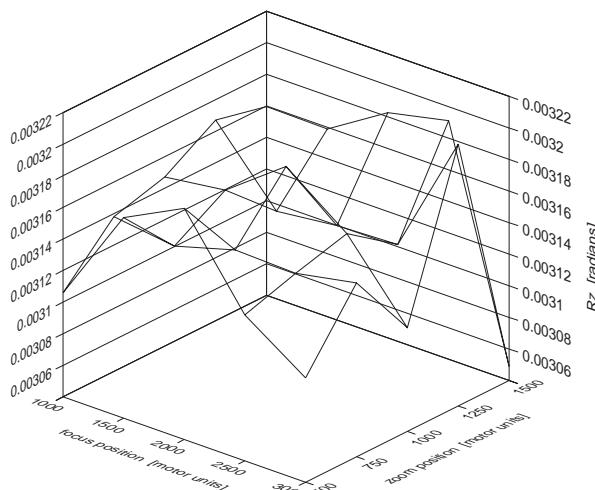


Figure 6.31: Cosmicar fixed camera model  $R_z$  versus focus and zoom motors

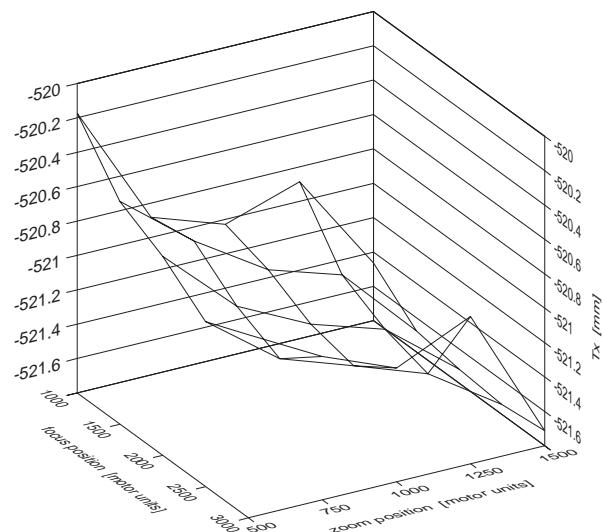


Figure 6.32: Cosmicar fixed camera model  $T_x$  versus focus and zoom motors

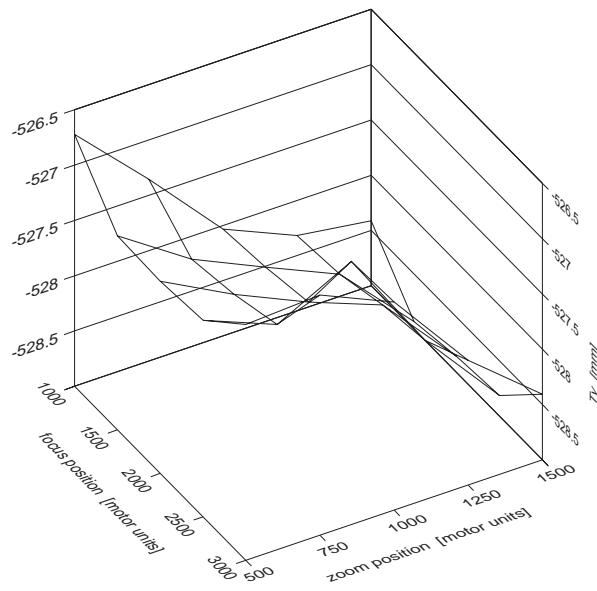


Figure 6.33: Cosmicar fixed camera model  $T_y$  versus focus and zoom motors

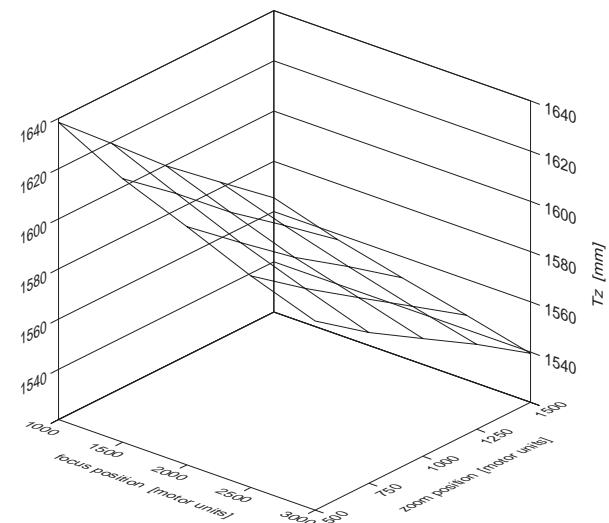


Figure 6.34: Cosmicar fixed camera model  $T_z$  versus focus and zoom motors

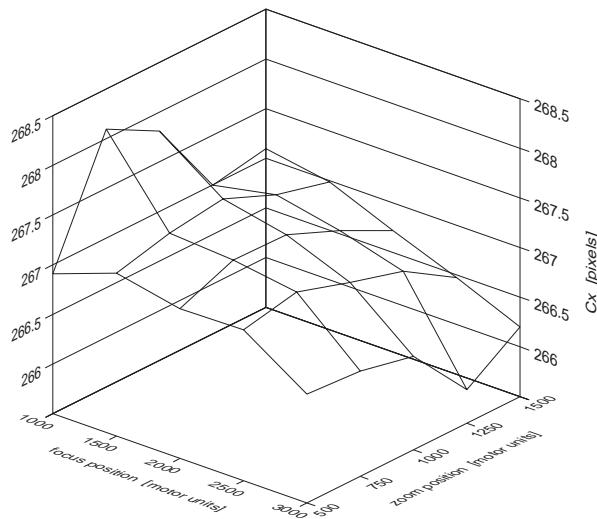


Figure 6.35: Cosmicar fixed camera model  $C_x$  versus focus and zoom motors

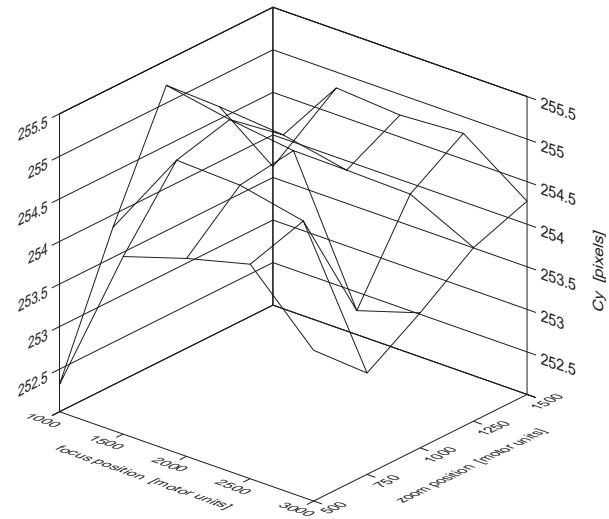


Figure 6.36: Cosmicar fixed camera model  $C_y$  versus focus and zoom motors

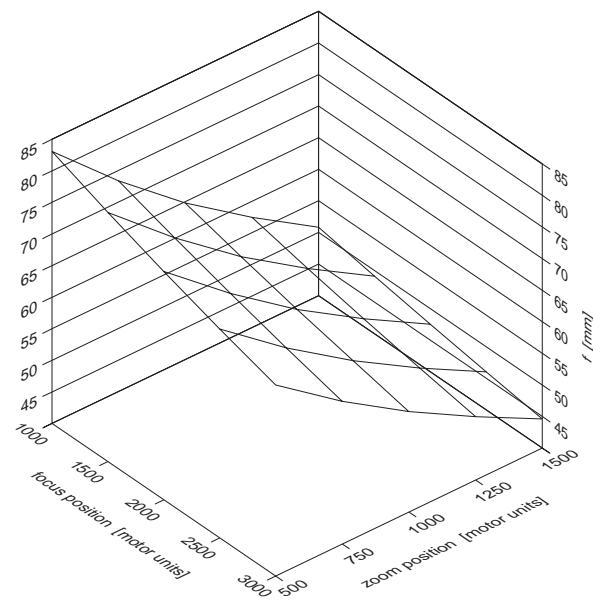


Figure 6.37: Cosmicar fixed camera model  $f$  versus focus and zoom motors

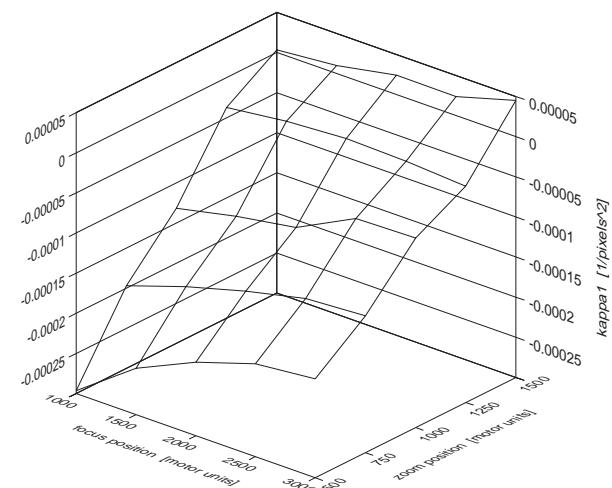


Figure 6.38: Cosmicar fixed camera model  $\kappa_1$  versus focus and zoom motors

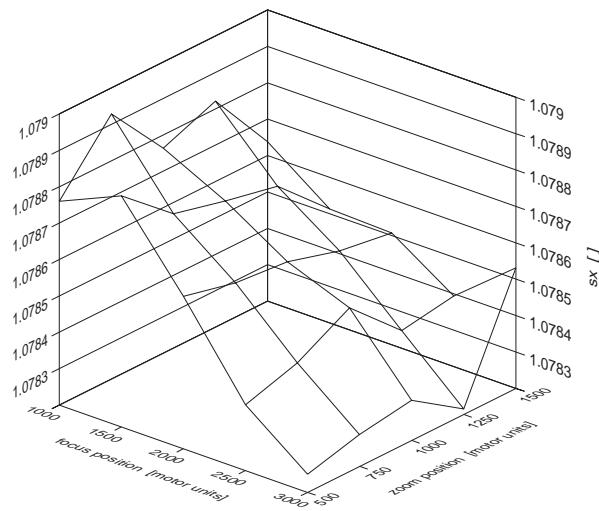


Figure 6.39: Cosmicar fixed camera model  $s_x$  versus focus and zoom motors

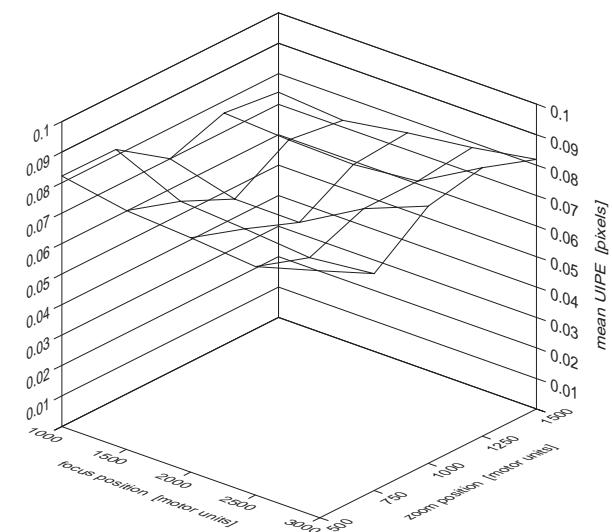


Figure 6.40: Cosmicar fixed camera model M\_UIPE versus focus and zoom motors

Fitting Step	Parameter	Polynomial Order	MM_UIPE [pixels]	max UIPE [pixels]	SSS_UIPE [pixels <sup>2</sup> ]
unfitted fixed models			0.075532	0.360894	36.702467
1	$T_x$	0	0.074999	0.361373	35.808658
2	$T_y$	0	0.074677	0.358428	35.602254
3	$R_z$	0	0.074643	0.368658	35.640737
4	$R_x$	0	0.075338	0.368822	36.268329
5	$s_x$	0	0.076347	0.394985	37.261021
6	$R_y$	0	0.078349	0.396832	39.428921
7	$\kappa_1$	2	0.078690	0.402834	39.809953
8	$T_z$	4	0.078729	0.401802	39.880417
9	$f$	4	0.078897	0.405296	40.011936
10	$C_y$	4	0.079305	0.408550	40.338094
11	$C_x$	4	0.079791	0.409626	40.783569
12	$C_x$	4	0.079305	0.408550	40.338064
13	$R_x$	0	0.079065	0.403548	40.174146
14	$R_y$	0	0.078200	0.409401	39.044316
final adjustable model			0.078200	0.409401	39.044316

Table 6.5: Fitting sequence for Cosmicar/Panasonic parameter models for focus and zoom

Another interesting feature to note in Fig. 6.45 is  $\kappa_1$ 's transition from negative values to positive values for zoom motor positions greater than 1400. This corresponds to a change from barrel distortion to pincushion distortion for the lens.

### 6.5.7 Validating the adjustable model

To test the generality of the adjustable camera model and the repeatability of the camera hardware we used two independent sets of calibration data, Set 1 and Set 2. The two sets were taken one after the other with the exterior orientation of the camera unchanged between sets. Each set was used to generate an adjustable camera model. The two models are then tested on both sets of calibration data.

Table 6.6 shows the values of the zero-order parameter models calculated for both sets of data. Table 6.7 shows the calibration statistics for the two adjustable camera models tested against both sets of calibration data. The close correspondence between the zero-order models as well as the very similar statistics for the two adjustable models tested against their own calibration data, indicates that the camera hardware is repeatable. The tests also show a 45% increase in the MM\_UIPE when the adjustable models calibrated with one set of data are tested against the alternate set of data. While not as good as the results for the Fujinon/Photometrics camera the fact that the M\_UIPE is still in the order of 1/10 of a pixel is still very reasonable.

Figures 6.48 through 6.51 show plots of the M\_UIPE for both camera models tested against

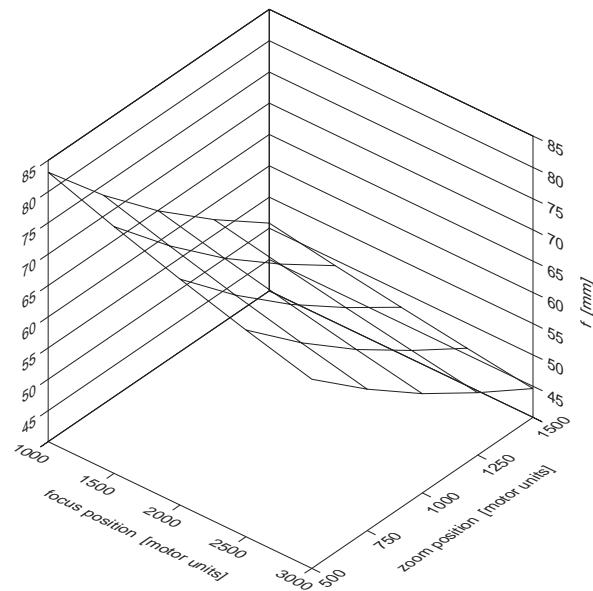


Figure 6.41: Cosmicar adjustable camera model  $f$  versus focus and zoom motors

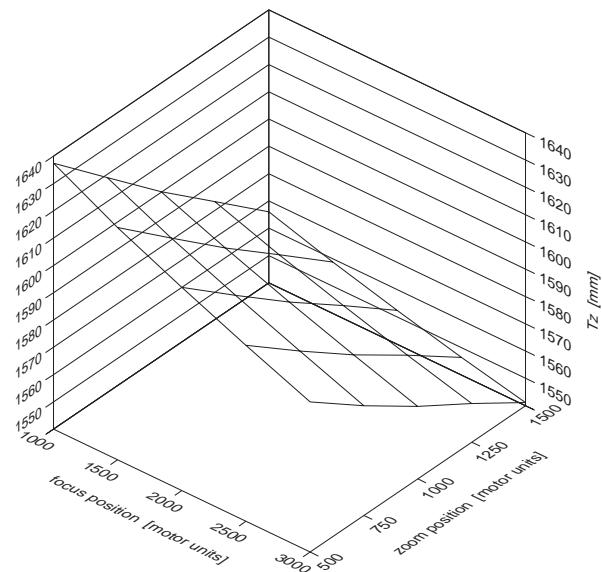


Figure 6.42: Cosmicar adjustable camera model  $T_z$  versus focus and zoom motors

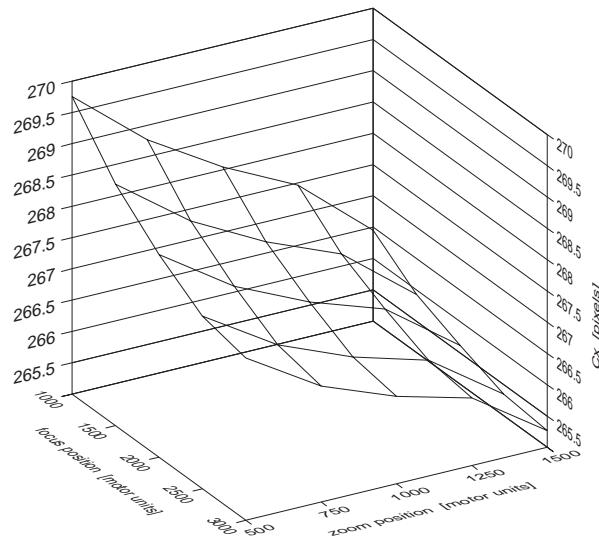


Figure 6.43: Cosmicar adjustable camera model  $C_x$  versus focus and zoom motors

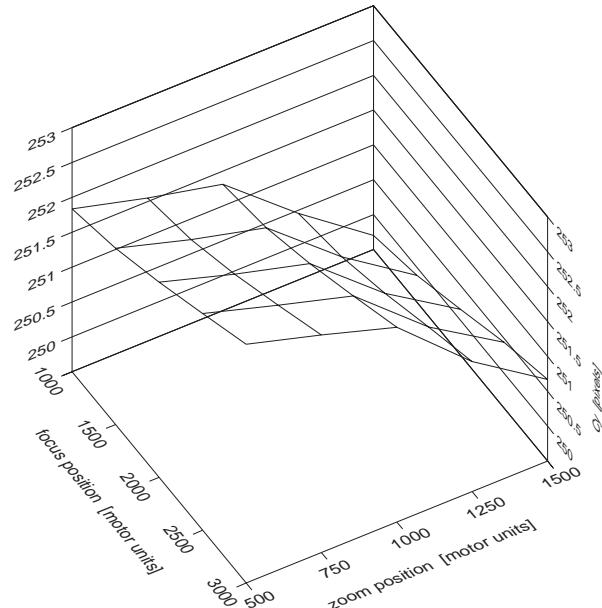


Figure 6.44: Cosmicar adjustable camera model  $C_y$  versus focus and zoom motors

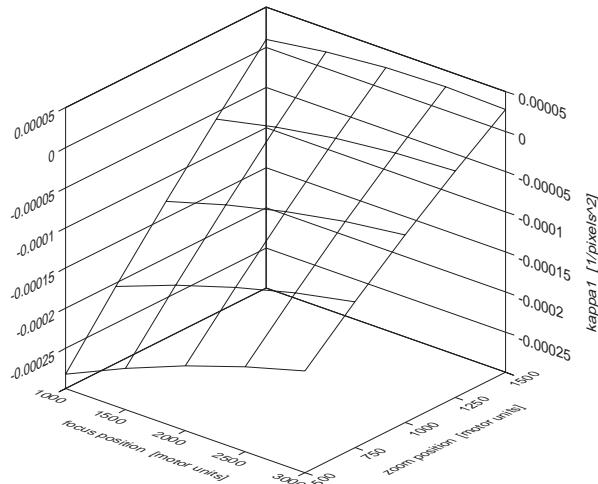


Figure 6.45: Cosmicar adjustable camera model  $\kappa_1$  versus focus and zoom motors

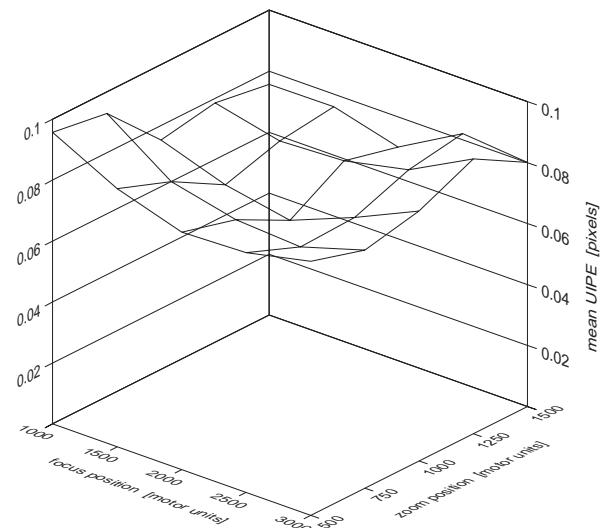


Figure 6.46: Cosmicar adjustable camera model M\_UIPE versus focus and zoom

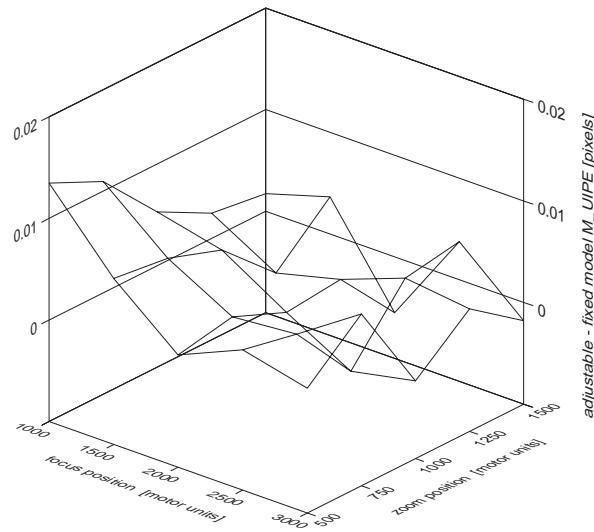


Figure 6.47: Difference between the final adjustable model M\_UIPE and the initial fixed model M\_UIPE

Parameter Model	Set 1	Set 2	Units
$R_x$	-0.132285	-0.144780	degrees
$R_y$	0.594484	0.596439	degrees
$R_z$	0.179774	0.179382	degrees
$T_x$	-521.114	-521.065	mm
$T_y$	-526.596	-526.239	mm
$s_x$	1.078538	1.078548	

Table 6.6: Zero-order parameter models for two sets of calibration data from the same pose

	Test 1	Test 2	Test 3	Test 4	Units
data used for model	Set 1	Set 2	Set 1	Set 2	
data used for testing	Set 1	Set 2	Set 2	Set 1	
MM_UIPE	0.078200	0.076728	0.113314	0.110384	pixels
max UIPE	0.409401	0.377333	0.448178	0.521263	pixels
SSS_UIPE	39.044316	37.880414	81.859555	80.331112	pixels <sup>2</sup>

Table 6.7: Performance of two adjustable models on two sets of calibration data from the same pose

both sets of calibration data. As with the Fujinon lens, the shapes are not significantly different from one another which is an indication that the two adjustable models are general.

### 6.5.8 Exterior orientation recalibration example

To test our ability to reacquire the camera's extrinsic parameter models after the camera has been moved we took one calibration dataset, changed the camera's exterior orientation (pose), and then took a second dataset. Using the first dataset we obtained a fully calibrated adjustable camera model. The intrinsic parameter models from this model are then used in a partial calibration with the second dataset to estimate the camera's new extrinsic parameter models. The partial calibration was performed first with one base setting and then with four base settings for improved precision. The second set of data was then used to obtain a fully calibrated adjustable camera model. The test is illustrated graphically in Fig. 6.52.

For the first dataset we actually used the data in Set 1 from Section 6.5.7. Before taking the second dataset the camera was shifted -100 mm in the camera  $x_w$  coordinate, -100 mm in the camera  $y_w$  coordinate and -100 mm in the camera  $z_w$  coordinate, and then rotated until the center of the field of view was located roughly in the middle of the two target plane positions. The new dataset was obtained from pairs images of the calibration target taken

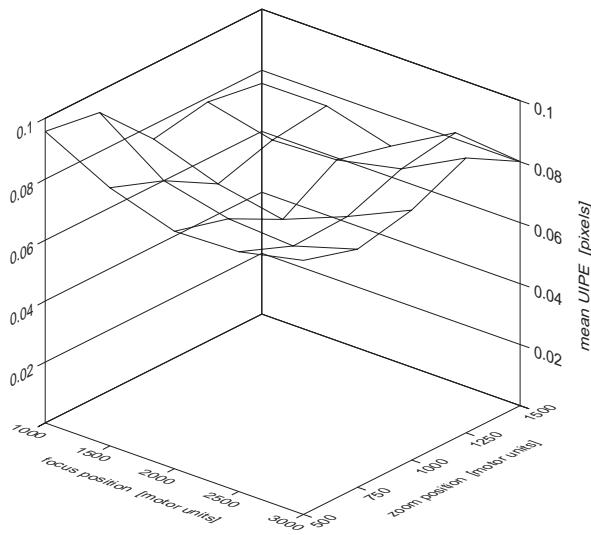


Figure 6.48: M\_UIPE for model calibrated with dataset 1 and tested on dataset 1

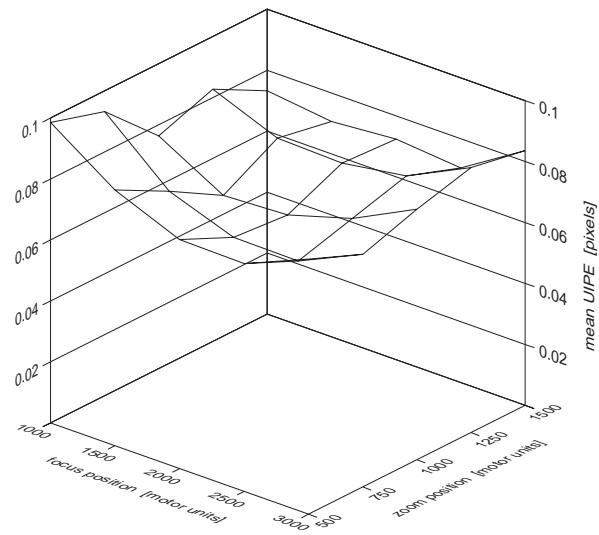


Figure 6.49: M\_UIPE for model calibrated with dataset 2 and tested on dataset 2

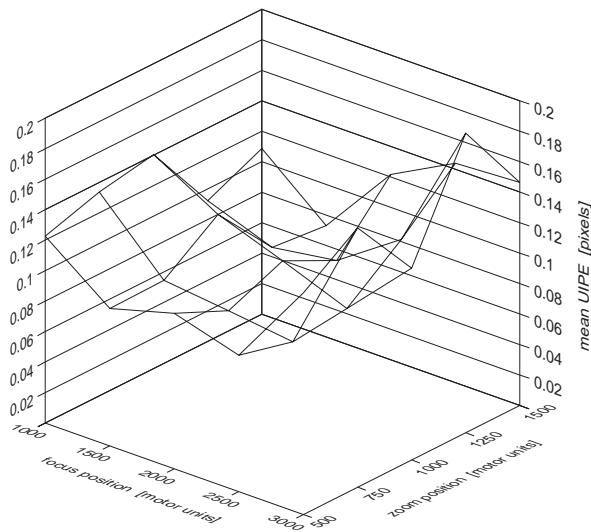


Figure 6.50: M\_UIPE for model calibrated with dataset 1 and tested on dataset 2

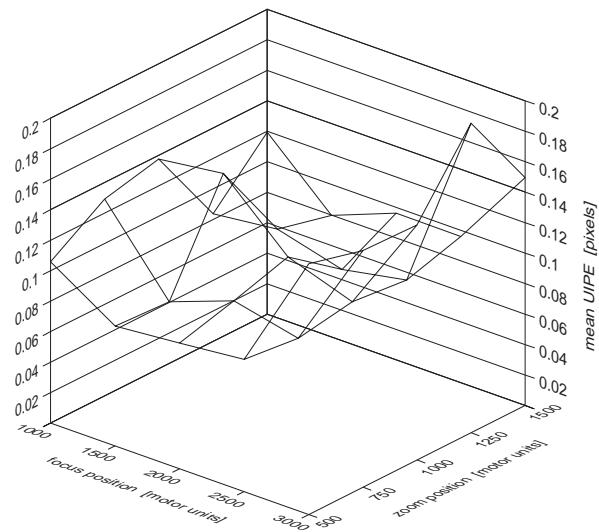


Figure 6.51: M\_UIPE for model calibrated with dataset 2 and tested on dataset 1

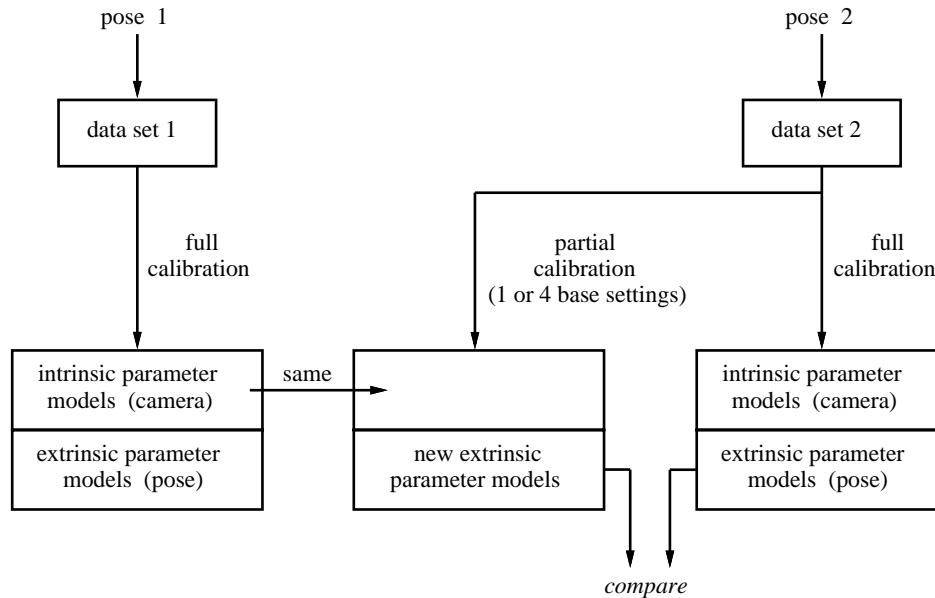


Figure 6.52: Experiment in carrying intrinsic parameter models to new pose

on the same  $5 \times 5$  sampling of lens settings used for the first dataset.

Column 2 of Table 6.8 shows the results for the full calibration on the data from the first pose. Column 5 shows the results for the full calibration for the second pose. Columns 3 and 4 of Table 5.2 shows the calibration results when the intrinsic parameter models from the first pose are used in a partial calibration with one and with four base settings to obtain the extrinsic parameter models for the second pose. Comparing the results for the second pose (columns 3, 4 and 5) we see that the zero-order extrinsic models for the partially calibrated models are reasonably close to the values for the fully calibrated adjustable model. Using four base settings reduces the mean M\_UIPE by 12% and the maximum UIPE by 35%.

Figure 6.53 shows the M\_UIPE for the full calibration at the first pose. Figure 6.54 shows the M\_UIPE for the full calibration at the second pose. Figure 6.55 shows the M\_UIPE for the partial calibration when only one base setting is used. Figure 6.56 shows the M\_UIPE for the partial calibration when four base settings are used.

While the results of this experiment are not as good as those for the fixed camera model (Section 5.8), a MM\_UIPE in the order of 1/7 of a pixel and a maximum UIPE for all of the data points of 0.54 pixels is reasonable and it still indicates that we can reacquire the camera's extrinsic parameter models after the camera has been moved. It also demonstrates that the intrinsic parameter models can be carried between poses.

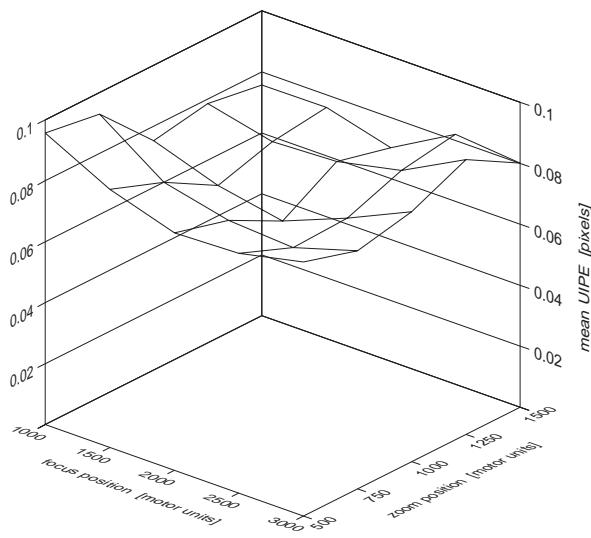


Figure 6.53: M\_UIPE for model calibrated at pose 1

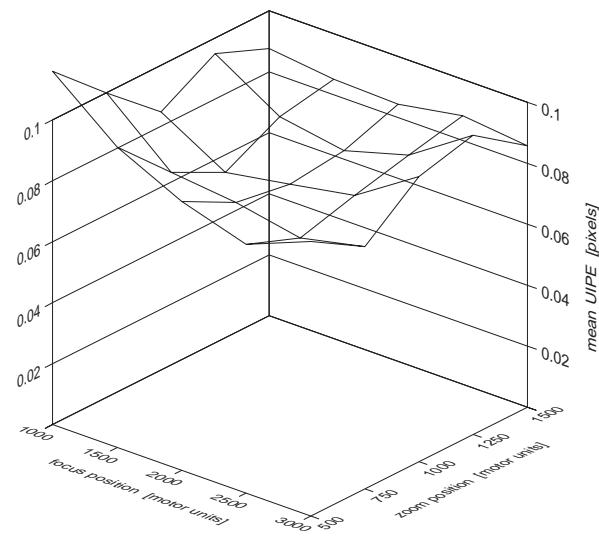


Figure 6.54: M\_UIPE for model calibrated at pose 2

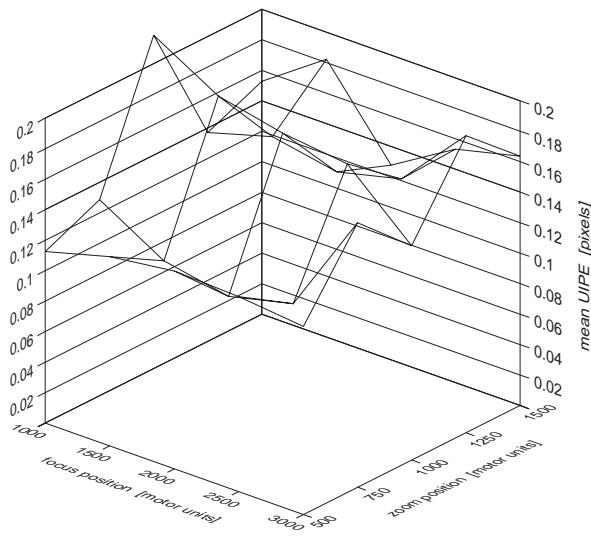


Figure 6.55: M\_UIPE for extrinsic recalibration using 1 base setting

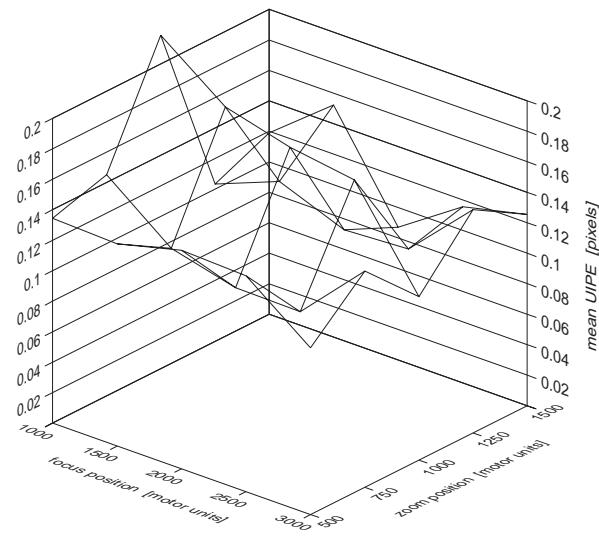


Figure 6.56: M\_UIPE for extrinsic recalibration using 4 base settings

Parameter	Pose 1	Pose 2	Pose 2	Pose 2	Units
	full calibration	partial calibration	partial calibration	full calibration	
	1 base setting	4 base settings			
$R_x$	-0.132285	-2.878022	-2.877802	-3.015983	degrees
$R_y$	0.594484	-2.032490	-2.035238	-2.032272	degrees
$R_z$	0.179774	0.308359	0.306804	0.304904	degrees
$T_x$	-521.114	-496.822	-496.797	-496.705	mm
$T_y$	-526.596	-546.035	-546.029	-545.951	mm
$s_x$	1.078538	1.078538	1.078538	1.078460	
MM_UIPE	0.078200	0.147785	0.131852	0.086522	pixels
max UIPE	0.409401	0.735700	0.543040	0.384438	pixels

Table 6.8: Calibration results when carrying intrinsic parameter models to new pose

## 6.6 Fujinon/Photometrics model for aperture

In the preceding sections we described two adjustable camera models that held calibration over ranges of focus and zoom. For these models the aperture settings were fixed. However, in many situations we might want to be able to vary the lens's aperture. To determine how strong an influence the aperture has on perspective projection we built an adjustable perspective-projection camera model for the Fujinon/Photometrics camera for a range of aperture settings.

For our new model the focus was fixed at 2500 motor units and the zoom at 2000 motor units. The aperture was varied in 50 motor unit steps from 350 to 2000 motor units. The 35 lens settings sampled correspond to a range of relative apertures from  $f/16.5$  to  $f/2.8$ . Calibration data was obtained from images taken with 1.5 m, 2.0 m and 2.5 m between the camera's sensor and the calibration target. Below the lower setting the aperture was closed while above the upper setting the target features were too blurry to reliably measure all three image planes.

Figures 6.57 through 6.67 show plots of the 11 fixed model parameters and the 11 fitted parameter models versus the aperture motor. Beyond an aperture setting of 1500 motor units (approximately  $f/4.7$ ) the measurement error caused by defocus made the fixed camera calibration unstable.

To model the effects of aperture on the 11 parameters in our fixed camera model we used zero-order polynomials for the  $s_x$ ,  $R_x$ ,  $R_y$ ,  $R_z$ ,  $T_x$ , and  $T_y$  terms and first-order polynomials for the  $T_z$ ,  $f$ ,  $C_x$ ,  $C_y$ , and  $\kappa_1$  terms. As with the previous two adjustable camera models, these model orders were chosen based on design requirements for the model and on empirical results.

To fit the parameter models to the calibration data we used ascending order, greedy within order sequencing. Table 6.9 shows the sequence in which the parameter models were fit, along with the MM\_UIPE, maximum UIPE, and SSS\_UIPE statistics for the adjustable camera models at each stage. The first entry in the table is for the original unfitted fixed camera model parameters. Steps 1 to 11 are for the initial fitting sequence. Steps 12 and 13 are for iterative refinement. The last entry is for the final adjustable camera model.

Figure 6.68 shows the M\_UIPE for both the unfitted fixed camera models and the final adjustable model.

The results of this experiment indicate that the aperture has a significant effect on the perspective projection behavior of the camera. The relatively strong dependency of the  $f$ ,  $\kappa_1$ , and  $T_z$  terms on aperture can be attributed to the relationship between the first-order lens aberrations and the diameter of the aperture stop (see Appendix D).

To build an adjustable perspective-projection camera model for ranges of focus, zoom, and aperture would require collecting calibration data in all three ranges and constructing parameter models using trivariate polynomials. This experiment shows that for this focus and zoom position, the order of the variation of the fixed model's parameters with aperture is

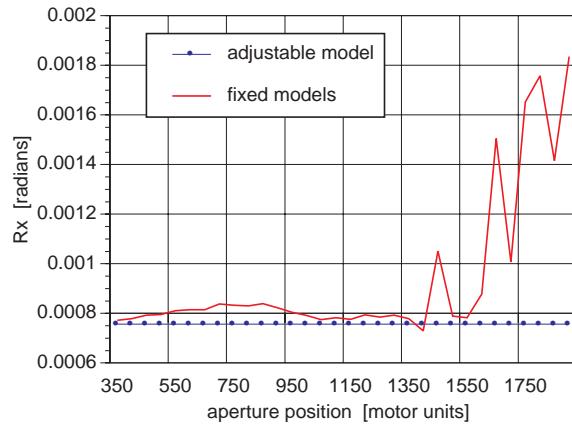


Figure 6.57: Fujinon camera model  $R_x$  versus aperture motor

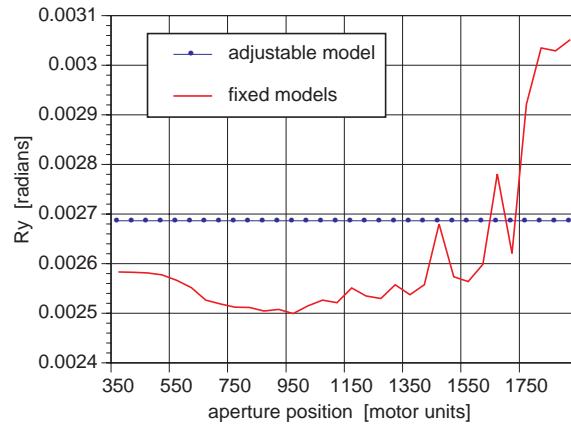


Figure 6.58: Fujinon camera model  $R_y$  versus aperture motor

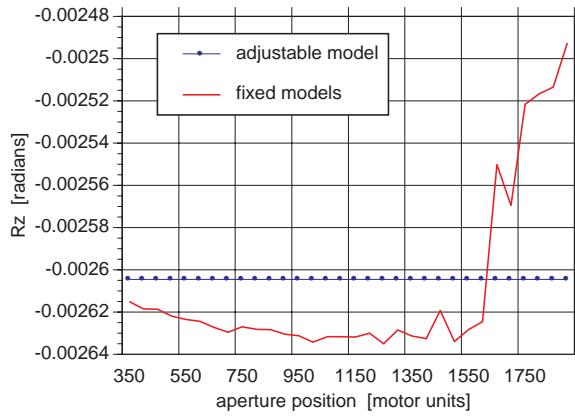


Figure 6.59: Fujinon camera model  $R_z$  versus aperture motor

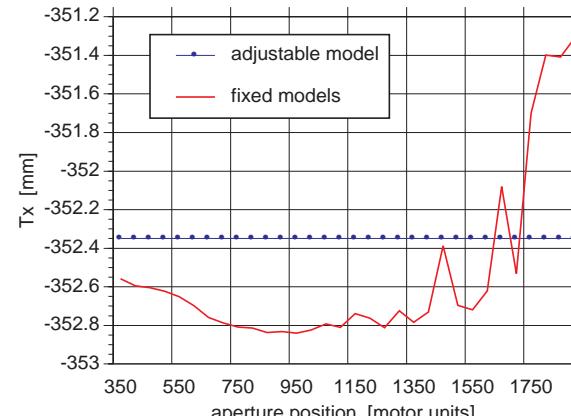


Figure 6.60: Fujinon camera model  $T_x$  versus aperture motor

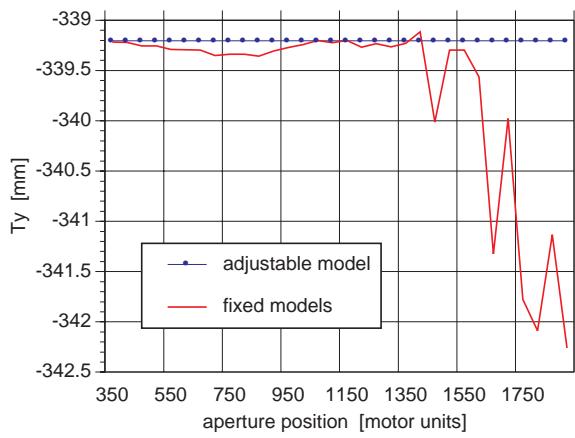


Figure 6.61: Fujinon camera model  $T_y$  versus aperture motor

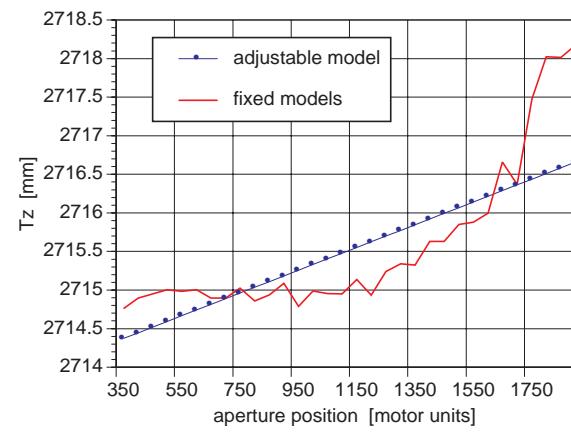


Figure 6.62: Fujinon camera model  $T_z$  versus aperture motor

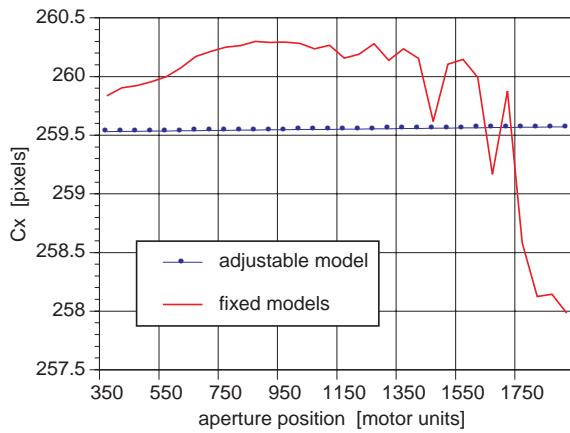


Figure 6.63: Fujinon camera model  $C_x$  versus aperture motor

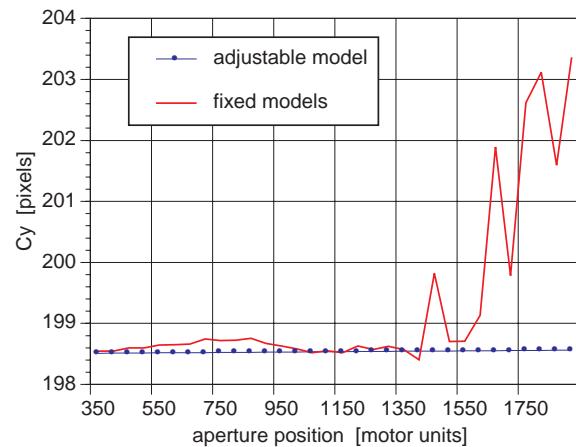


Figure 6.64: Fujinon camera model  $C_y$  versus aperture motor

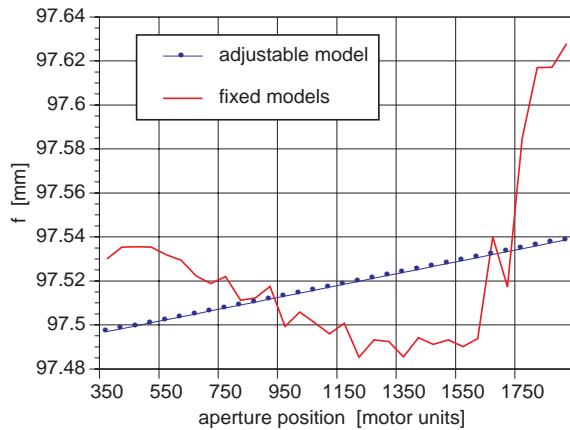


Figure 6.65: Fujinon camera model  $f$  versus aperture motor

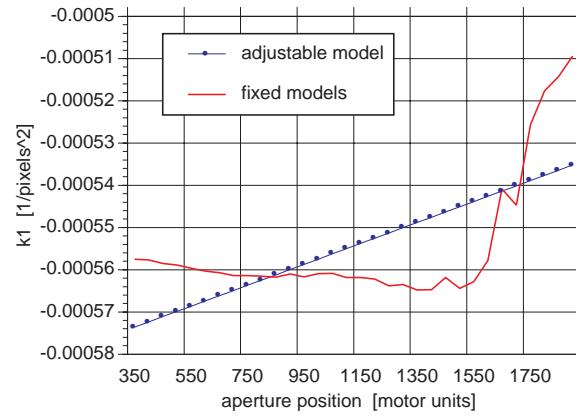


Figure 6.66: Fujinon camera model  $\kappa_1$  versus aperture motor

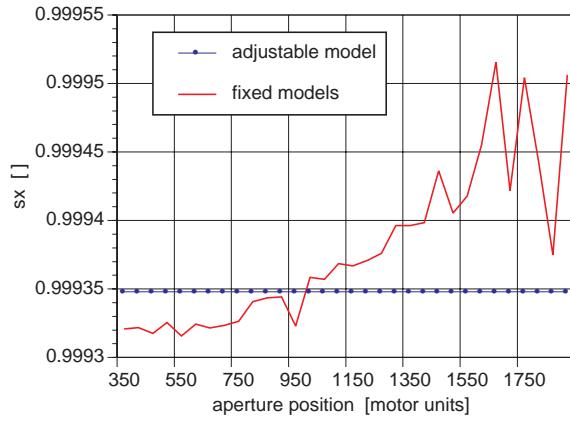


Figure 6.67: Fujinon camera model  $s_x$  versus aperture motor

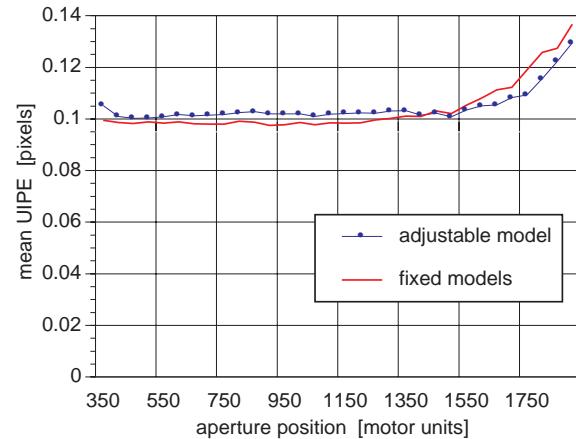


Figure 6.68: Fujinon camera model  $M_{UIPE}$  versus aperture motor

Fitting Step	Parameter	Polynomial Order	MM_UIPE [pixels]	max UIPE [pixels]	SSS_UIPE [pixels <sup>2</sup> ]
unfitted fixed models			0.103878	0.305112	64.304930
1	$T_x$	0	0.102600	0.255010	62.983127
2	$R_z$	0	0.102462	0.260005	62.873925
3	$s_x$	0	0.102356	0.265299	62.834046
4	$T_y$	0	0.102315	0.271714	62.836810
5	$R_y$	0	0.102507	0.271922	63.265362
6	$R_x$	0	0.104870	0.297507	65.638900
7	$C_x$	1	0.104910	0.297209	65.692866
8	$T_z$	1	0.104981	0.301933	65.795436
9	$f$	1	0.105154	0.307753	65.992409
10	$\kappa_1$	1	0.105234	0.295440	66.349142
11	$C_y$	1	0.105477	0.311013	66.783483
12	$C_y$	1	0.105234	0.295440	66.349137
13	$R_x$	0	0.104503	0.311011	65.671717
final adjustable model			0.104503	0.311011	65.671717

Table 6.9: Fitting sequence for Fujinon/Photometrics parameter models for aperture

relatively low. If this variation is uniformly low across the desired focus and zoom motor ranges, then the sampling strategy for the camera’s operating space would require very few aperture settings. In addition, the extension of the polynomial parameter models to include aperture could be accomplished using low orders for the aperture variable.

## 6.7 Discussion

In applying our methodology to produce an adjustable perspective-projection camera model we touched upon several issues that deserve further discussion.

### Sampling strategy

Our objective in building an adjustable camera model is to provide sets of “useful” parameter values for the fixed camera model for any given lens setting. “Useful” parameter values are those that yield a fixed camera model that accurately captures the lens’s image-formation process at a given lens setting.

To calibrate the adjustable camera model we need to take measurements of the camera’s imaging properties at different lens settings. The limit for the closeness of these measurements is the control resolution of the lens actuators. However, the number of lens actuators together with their respective control resolutions typically make an exhaustive sampling of

the lens's physical operating space impossible. For practicality we subsample the camera's control space.

To collect the calibration data used in this chapter we used evenly spaced samples in each of the lens control parameters. We assumed that the chosen sampling rate was sufficiently high to produce parameter models that accurately interpolated the lens's imaging behavior between all of the sampled lens settings. To reduce the number of samples required to calibrate our lens model we could use variable or adaptive sample spacings.

One approach to determine where samples need or do not need to be taken is to use a dense survey of the operating space (as demonstrated in Section 4.2.2). To further reduce the number of samples taken the dense surveying could be performed along single control variables (i.e. taking cuts across the lens's physical operating space). Another approach would be to look for regions of high curvature in the parameter models produced from a given set of calibration data. More calibration data could be then taken around these regions and a new set of parameter models built. The process could be repeated until the differences between successive parameter models fell below some threshold.

## Formulation and fitting of parameter models

For the parameter models used in this chapter we used polynomials that were symmetric in each of the model's independent variables (i.e. had the same polynomial order for each control motor). From the graphs of the final parameter models for  $f$  (Figs. 6.15 and 6.41) and  $T_z$  (Figs. 6.16 and 6.42) we can see that the dependence between the model parameters and the focus and zoom motors is not the same. Permitting different polynomial orders for each independent variable would allow the overall complexity of the adjustable camera model to be reduced without significantly degrading the model's performance.

To fit the polynomial models in this chapter we used ascending polynomial order, greedy within order sequencing followed by an iterative refinement step. Throughout this process the fitting and refitting was performed one parameter model at a time. One possible way to speed convergence would be to fit or refit two or more parameter models at the same time. We tried the extreme of this approach by initially fitting all of the parameter models at the same time. In this case, however, the resulting adjustable model was poor: it required numerous iterative refinements and was trapped in a local minima.

For computational efficiency we used a least-squared-error technique to fit the parameter models in this chapter. Unfortunately the data for the models was generated using iterative non-linear optimization on error surfaces that, in practice, have many local minima. If we consider different sets of calibration data taken at the same lens settings, the variation in the generated model parameters is unlikely to have a Gaussian, zero-mean distribution. Given this, the least-squared-error technique is not necessarily the best approach to fitting the data. To deal with the non-Gaussian statistics of our data we have experimented with using robust fitting techniques. For our calibration data, however, this approach did not significantly change the performance of the final adjustable model.

To build the adjustable models in this chapter we chose the order of the parameter models first and then fit the parameter models to the data. Another approach would be to vary the model order during the fitting process. One way to look at this is to consider the fitting the parameter models as imposing a regular structure on the calibration data. Re-estimation of the fixed camera model’s parameter values essentially allows the calibration data to re-conform to the structure being imposed on it. In this paradigm low-order parameter models constitute a “stiff” adjustable camera model to which the calibration data may not be able to conform well (making the adjustable model’s error large). Higher-order parameter models constitute a more “flexible,” adjustable camera model to which the calibration data can conform better (making the adjustable model’s error smaller). One way to integrate the formulation and fitting process then would be to fit a very flexible adjustable camera model to the calibration data and then reduce the polynomial orders of the independent variables in the parameter model one at a time to stiffen the adjustable camera model until a balance is achieved between the complexity of the adjustable model and the model’s performance. Alternately we could start by fitting a stiff adjustable model and then progressively make it more flexible by increasing the polynomial orders of the independent variables in the parameter models. In both of these approaches the MDL framework could be used to guide the decision making process.

### Calibration across multiple poses

The calibration data we used in this chapter came from images of a single target with the camera held at a single pose (i.e. the exterior orientation remained constant). For broader ranges of focus or zoom it may be necessary to use different calibration targets for different *subranges* of the camera’s operating space. In most cases changing the calibration target will change the relationship between the world coordinate frame (which the target’s reference points are located in) and the camera’s coordinate frame. To accommodate this the extrinsic parameter models,  $g_{R_x}$ ,  $g_{R_y}$ ,  $g_{R_z}$ ,  $g_{T_x}$ ,  $g_{T_y}$ , and  $g_{T_z}$  will have to be calibrated separately for each subrange of the operating space. At the same time the intrinsic parameter models,  $g_f$ ,  $g_{T_z}$ ,  $g_{\kappa_1}$ ,  $g_{C_x}$ ,  $g_{C_y}$ , and  $g_{s_x}$  must still be calibrated across the full operating space. This has significant implications for the fitting of the  $g_{T_z}$  parameter which has both intrinsic and extrinsic characteristics. While the value of  $g_{T_z}$  may change when a different subrange is calibrated, generally we would like the derivatives of  $g_{T_z}$  to be continuous as we move from one subrange to the next.

## 6.8 Summary

In this chapter we have described the adjustable perspective-projection camera model we developed based on Tsai’s fixed camera model. This model was designed to hold calibration across continuous ranges of both focus and zoom. We have also described a version of this model that holds calibration across a continuous range of aperture. Our model was calibrated and tested on two distinctly different camera systems. Both systems performed across their

calibrated operating ranges of focus and zoom with an average error of less than 0.14 pixels between the predicted and measured positions of features in the image plane. Our model was also calibrated and tested on one system for a continuous range of aperture, where it achieved similar results.



# Chapter 7

## Conclusions

This thesis presents new results for the modeling, calibrating, and control of cameras with automated zoom lenses. In this thesis we show that traditional approaches cannot capture the complex relationships between control parameters and imaging processes. Furthermore, we show that the idealized behavior tradition models assume can lead to significant performance problems in color imaging and focus ranging. We demonstrated that we could reduce or eliminate these performance problems by using more complex models and control strategies. As an outgrowth of our examination of the behavior of real lenses in both fixed- and variable-parameter systems we have also developed a comprehensive taxonomy for the property of “image center.”

In this thesis we describe a methodology for producing accurate camera models for systems with variable-parameter lenses. We used this approach to produce an adjustable, perspective-projection camera model based on Tsai’s fixed camera model and calibrated and tested our model on two different, automated-camera systems. In both cases the calibrated models operated, across continuous ranges of focus and zoom, with an average error of less than 0.14 pixels between the predicted and the measured positions of features in the image plane. We have also calibrated and tested our model on one system across a continuous range of aperture and achieved similar results.

### 7.1 Contributions

This research makes contributions to computer vision in the theory, modeling, calibration, and control of variable-parameter camera systems, specifically:

- *A new methodology for producing adjustable camera models.* Our approach involves first calibrating a conventional fixed camera model at a number of lens settings spanning the adjustable model’s physical operating space. We then characterize how the parameters of the fixed model vary with lens setting by alternately fitting polynomials to individual model parameters and reestimating the as yet unfitted parameters using the calibration

data. The process is repeated until all of the fixed camera model’s terms have been replaced with polynomial functions of the lens control parameters. The result is a predictive camera model that can interpolate between the original sampled lens settings to produce — for any lens setting — a set of values for the parameters in the fixed camera model. As part of our methodology we use a preliminary measurement survey of the physical operating space to identify regions where rapid changes in imaging properties may require denser sampling for calibration data.

Unlike previous modeling approaches, this approach makes no a priori assumptions about the dependencies between the parameters of the fixed camera model and the lens settings. Our approach handles imaging behavior that cannot be explained using the abstract models of the image-formation process.

This modeling approach is general and can be applied to produce — for any image property — an adjustable camera model from the fixed one. Any number of independent lens control variables can be incorporated into the adjustable camera model. The degree of accuracy and complexity, and the required calibration effort can be chosen arbitrarily.

- *A new adjustable perspective-projection camera model.* Our adjustable camera model is based on Tsai’s 11-parameter fixed camera model for perspective projection. Five of the exterior orientation parameters,  $R_x$ ,  $R_y$ ,  $R_z$ ,  $T_x$ , and  $T_y$ , are modeled with fixed constants, simplifying the adjustable model’s calibration and use. Four of the interior orientation parameters,  $f$ ,  $C_x$ ,  $C_y$ , and  $\kappa_1$ , and the  $T_z$  exterior orientation parameter are modeled with multivariate polynomials whose order can be increased or decreased to accommodate different rates of dependency between the imaging behavior and lens settings in different camera systems. The last interior orientation parameter,  $s_x$ , is modeled with a constant.

This model can be calibrated for continuous ranges of focus, zoom, and aperture.

- *A new taxonomy for the property of “image center.”* First-order models of lens behavior, such as the pinhole-camera model or the thin-lens model, suggest that the image center is a single, fixed, intrinsic parameter of the lens. In reality we find that there are many possible definitions of image center, and most do not have the same coordinates. Moreover, image centers move as lens parameters are changed. In this research we have developed a new and comprehensive taxonomy that includes 15 different definitions of image center and procedures for measuring them.
- *A new algorithm for color imaging.* In color imaging, the first-order model of lens behavior assumes that magnification, focus, and center-of-field-of-view are independent of the color of light being imaged. We show that the second-order effect of chromatic aberration can cause significant defocusing and misregistration between the bands of a color image. We have developed a new approach, called Active Color Imaging, that refocuses, re-zooms, and re-centers the lens as the images are taken, thereby reducing the levels of misregistration by an order of magnitude.
- *New theory for focus ranging.* In focus ranging the first-order model of lens behavior assumes that the position of the peak in the sharpness criterion function is only a

function of the lens's focus control. We show theoretically and experimentally how the second-order effect of focus magnification can add a target-dependent bias to the position of the criterion function peak, causing error in the range estimate. By maintaining a constant image magnification during focus ranging by varying both the focus and zoom motors we can eliminate this bias.

## 7.2 Directions for future research

This research has, for the first time, demonstrated that we can construct accurate camera models of complex imaging properties for systems with variable-parameter lenses. In developing and applying our methodology many issues regarding the building of adjustable camera models were only briefly addressed. In addition, the new found availability of adjustable camera models, such as our perspective-projection model, opens the question of where they might be applied.

In section 6.7 we outlined several issues on the building of adjustable models that require further study, including sampling strategies, formulation and fitting of parameter models, and calibration across multiple poses. While we have produced an adjustable camera model that is quite usable as is, many of the approaches we used in this thesis could be refined to generate more compact models and reduce the amount of calibration required.

Automated lenses are mechanical devices containing moving parts. Whenever we have moving parts we must deal with mechanical play and wear. The experiments described in this thesis were performed in a controlled laboratory environment. If variable-parameter lenses are to be used outside of the laboratory, attention will have to be given to such factors as vibration, acceleration, temperature, and wear, that may affect the accuracy of our models.

The technology used in variable-parameter lenses is continually evolving. For example, broadcast quality zoom lenses are now available with focal length ranges of 13.5-740 mm (55 $\times$ ). While this expanded range in optical parameters increases their potential usefulness, it also makes calibration of the lens more difficult. One way to approach this problem would be to use active or adaptive calibration targets.

More advanced zoom lenses (e.g.[13][32]) use computer-controlled actuators rather than mechanical slides and cams to position the optical elements within the lens body. These actuators allow the possibility of modeling directly from the parameters of an abstract camera model to the positioning of the individual lens elements in the lens body. The advantage of direct modeling is it alleviates many of the potential mechanical problems, such as hysteresis, discontinuity, and wear, that must be dealt with in more conventional lenses. Other technological advances, such as liquid, graded index, holographic optics, and electro-optics (e.g.[47]), will likely affect the way variable-parameter lenses are modeled and calibrated.

In this thesis we have shown that accurate, predictive camera models can be built for variable-parameter lenses. These models have great potential as tools in machine vision but we have only started to consider how they might be applied. One interesting possibility would be to

use our adjustable, perspective-projection camera model in a stereo system having variable-parameter lenses. By changing the focus and aperture of the lenses in such a system we can blur out the scene's foreground and/or background. This capability can be used to simplify feature matching both by reducing the number of features detected in the scene and by constraining the range of their allowable disparities. Adjusting the lens's zoom can also maximize the disparity of matched features, thereby increasing the accuracy of range measurements.

# Appendix A

## Calibrated Imaging Lab

The Calibrated Imaging Lab at CMU provides a highly controlled environment for experiments in physics-based machine vision, including color inter-reflection, surface roughness, texture analysis, and camera modeling and calibration. The general lab facilities include a light-tight imaging area and a variety of lighting, positioning and measurement equipment. This lab also has two high-precision, automated camera systems developed by us specifically for physics-based, machine-vision research. In this appendix we briefly describe the general facilities in the lab and the development of the two camera systems with their automated lenses.

### A.1 General facilities

The key feature of the Calibrated Imaging Lab is a  $21 \times 10 \times 12$ -foot “imaging studio” that can be sealed off from the computer equipment and other work areas in the room. This studio presents a controllable environment for acquiring image data to support research in machine vision.

The studio area has a black ceiling and black carpet; black curtains are also normally used, to eliminate all reflections from the walls. This allows direct illumination to be configured and modeled relatively easily. There are also white curtains that can be pulled around the entire area for studies of diffuse illumination (shadows, etc.).

Lighting facilities include a battery of six Macbeth multi-spectral fixtures for diffuse lighting by three different sources (incandescent or cool or warm fluorescent). For directional lighting, the lab has several 1000 W theater spotlights hung from a grid of pipe work on the ceiling. For many purposes, free-standing 600 W photographers’ lights are used with manually operated dimmers.

The lab features an optical table used for placing objects to be viewed. On this table is a large (60-inch) travel rail table fitted with an 8-inch diameter, 360-degree Daedal rotation platform. Test targets and objects of various types are placed on the platform for stationery

imaging or controlled motion.

For moving the camera, the lab has a full 6-degree-of-freedom (6-DOF) rotation/translation jig that moves a 12×18-inch platform with high precision (0.001 inch and 0.01 degree resolution). Combined with the rail table, the camera-to-object distance can be varied from about 3 feet to 10 feet (1 m to 3 m). The rail table, rotation platform, and camera jig are all controlled by the lab's Sun4 workstation.

The CIL also has calibration equipment and targets. For geometric calibration, the lab has a pair of Lietz DT20E 20-arc-second surveyor's theodolites; with these, we can determine the position of any point in the lab with a precision uncertainty of about 0.3 mm in each dimension. For radiometric calibration, the lab has a Gamma Scientific spectroradiometer. This is used as a unit for measuring reflectance, transmittance, or illumination; and it can be disassembled for measuring the spectral responsivity properties of a camera. The CIL has numerous commercial test targets such as gray-scale charts; and has designed many additional targets such as the CMU Color Chart, which is particularly well-suited for color cameras used in machine vision.

Taken together, in conjunction with the cameras and lenses describe below, the facilities of the CIL provide an environment for controlled experimentation in a wide variety of topics of machine-vision research.

## A.2 Cosmicar/Panasonic camera system

The Cosmicar/Panasonic system was our first attempt at building a precision camera system. Our objective was to take a conventional CCTV camera system and see what type of performance we could get out of it. The system consists of a Cosmicar CCTV grade zoom lens mounted on a Panasonic GP-MF702 MOS solid state camera. The video signal from the camera is digitized to 8-bits/pixel by a Matrox VIP frame grabber in the lab's Sun4 workstation. To select the color band the system uses a seven slot motorized filter wheel housing six filters, including red, green, and blue (Wratten #25, 58, and 47B), 0.3 neutral density, 0.6 neutral density, and 0.9 neutral density. As CCD cameras are generally sensitive down well into the infra red, and since all of the band pass filters that we use are transparent in the IR regions, we use a total IR block (Schott KG5) mounted directly on the camera lens.

While the Panasonic video camera is fairly high quality, the camera's images suffer from the timing jitter and noise problems common to all video cameras [4]. The limited dynamic range of the camera combined with a fixed sensor gain and exposure time make experimentation with the full range of aperture positions very difficult to do. Another shortcoming of the system results from the low sensitivity of the camera in the blue region as well as the low light transmission of the blue filter. This necessitates the use of wide apertures to improve the SNR of the blue images during color imaging.

## A.3 Fujinon/Photometrics camera system

The Fujinon/Photometrics camera system results from our efforts to improve on the shortcomings of the Cosmicar/Panasonic camera system. The system consists of a Fujinon ENG color camera grade zoom lens mounted on a Photometrics Star I scientific camera. Images are digitized to 12-bits/pixel in the camera head by the camera's controller and then transferred to the lab's host computer over a GPIB interface. To select the color band, the system has a built in six slot motorized filter wheel housing five filters, including red, green, blue, 0.3 neutral density, and 0.9 neutral density. The system also has an IR block mounted directly on the camera lens.

In the Fujinon/Photometrics system the 2/3-inch image format produced by the lens does not completely cover the camera's sensor array and thus the usable image size for the camera ( $338 \times 338$  pixels) is smaller than the full sensor size. There were several reasons for this tradeoff. First, the Photometrics camera has a relatively large back flange focal length (on the order of 45 mm, even with a custom machined camera head). This precludes using commercially available CCTV camera lenses which typically have back flange focal lengths in the range of 10-20 mm. Second, in our experience, to be able to precisely motorize the lens body we needed a lens where the focal length DOF was rotational. This eliminated almost all commercially available 35 mm format zoom lenses. Of the remaining 35 mm format lenses that we examined none had a sufficiently large enough range of focal lengths or a solid enough mechanical movement for our purposes. This left us with commercial ENG/EFP lenses and broadcast quality TV lenses to choose from. For both of these types of lenses the cost of a 1-inch format lens was prohibitive. In the end, the Fujinon lens that we chose had the necessary range of focal lengths, three rotational DOFs, robust mechanical movement, and a 48 mm back flange focal length.

The Fujinon/Photometrics camera has several major advantages over the Cosmicar/ Panasonic system. The Photometrics camera features direct pixel-by-pixel digitization, virtually eliminating timing jitter as a source of image noise. In addition, the camera is cooled to -40°C which further improves the SNR of the image. The 12-bit dynamic range of the Photometrics camera is also a significant improvement over the Panasonic camera's 8-bit dynamic range. The variable exposure time and sensor gain of the Photometrics camera gives us an ability to experiment over the full range of aperture values that was unavailable with the Panasonic camera's fixed exposure time and gain. Finally the Fujinon/Photometrics camera was also designed with a larger focal length range and with finer control of the focus and focal length DOFs compared to the Cosmicar/Panasonic camera.

The Fujinon/Photometrics camera system is not without its limitations, however. To minimize the effects of variations in the shutter rise and fall times and to improve the SNR of the images, exposure times typically must be greater than 0.5 seconds for the Photometrics camera. The longer exposure time combined with the 3-4 seconds required to transfer a full image from the camera controller to the lab's workstation results in a relatively slow imaging speed for the camera. In comparison the Cosmicar/Panasonic camera system can image at a rate of 30 frames/second. At  $23 \times 23 \mu\text{m}$ , the pixel size for the Photometrics

camera is very large and results in aliasing problems in images with high spatial frequency content. The  $338 \times 338$  pixel effective image size is also a major drawback. Finally, saturation and blooming in the Photometrics camera has proved to be very difficult to control. In conclusion, rather than replace the initial Cosmicar/Panasonic camera system, the Fujinon/Photometrics camera system is used to complement it for high radiometric precision and geometric stability at a cost of speed and spatial resolution.

## A.4 Lens automation

The strength of the CIL's camera systems lies in their high-precision lens automation. Automation for both the Cosmicar and Fujinon lenses is provided by digital microstepping motors on each of the focal length, focus, and aperture DOFs. The nature of the microstepping motors ensures precise, drift-free repeatability for all of the lens control parameters, even across power downs.

For the Cosmicar lens all three motor shafts are connected to the rotational rings on the lens body by pushrod assemblies similar to the one illustrated in Fig. A.1. This design provides a 1:1 transfer of the motor's angular position to the lens body with virtually no mechanical play.

For the Fujinon lens the larger range of motion for the lens's focus distance ring requires the double pushrod assembly illustrated in Fig. A.2 to minimize the amount of lateral force exerted on the lens body. The double pushrod design requires an additional motor pivot to compensate for the translation of the lens's focus element along the optical axis, but has the advantage over the single pushrod approach of transferring a moment between the motor and the lens body rather than a force. The small range of motion for the Fujinon lens's zoom ring combined with the lens's large range of focal lengths necessitates the step-down wire pulley arrangement illustrated in Fig. A.3 to achieve sufficient resolution of the lens magnification. The tensioned-pulley design provides a 3.5:1 transfer of the motor's angular position to the angular position of the lens's zoom ring, again with no mechanical play. Finally the aperture of the Fujinon lens is controlled using the single pushrod assembly illustrated in Fig. A.1.

In most zoom lenses the focal length of the lens is changed by moving groups of lens elements relative to one another using internal cams and slides. In our system we minimize the effects of this internal source of mechanical play and any external stepper motor hysteresis by always approaching a desired motor position from one direction.

## A.5 Summary of computer-controlled parameters

These two systems - a CCTV system and a scientific camera system - give the CIL precise computer control over the most common camera DOFs. The Cosmicar/Panasonic camera system and camera jig provide a total of 10 computer-controlled DOFs, listed in Table A.1.

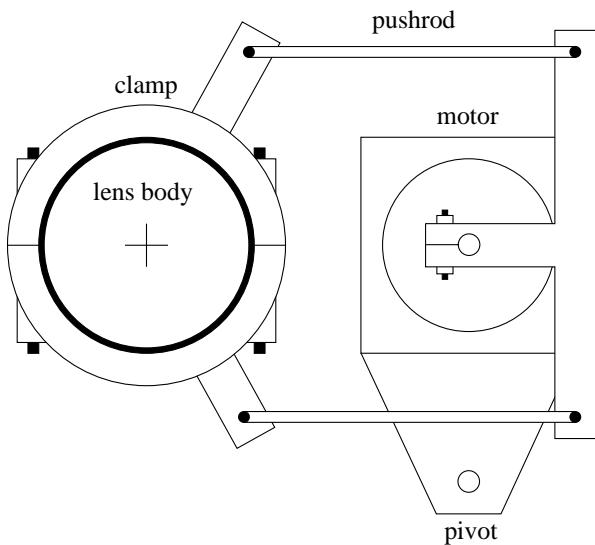
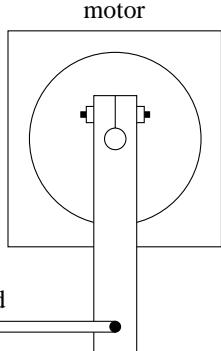
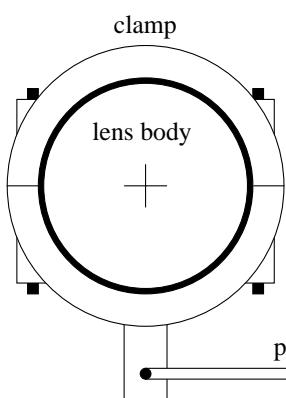


Figure A.1: Motor/lens pushrod assembly

Figure A.2: Motor/lens double pushrod assembly

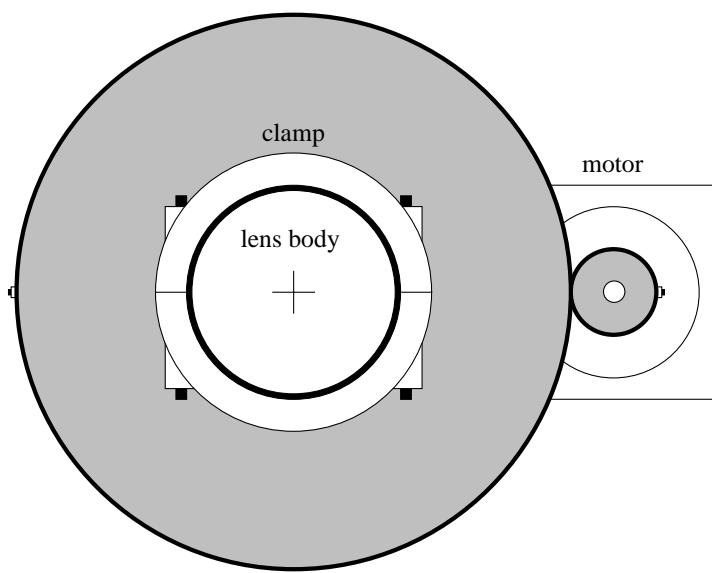


Figure A.3: Motor/lens pulley assembly

Table A.1: Controllable parameters for the Cosmicar/Panasonic camera system and jig

parameter	operating range	control resolution
focus distance	1.0 m - $\infty$	3900 steps (110° of travel)
focal length	12.5 - 75 mm (6×)	4000 steps (113° of travel)
aperture	$f/1.8$ - closed	2700 steps (76° of travel)
color band	none, red, green, blue	7 filters
$x, y, z$	24 in, 12 in, 80 in	0.001 inches
$\theta_x, \theta_y, \theta_z$	$\pm 20^\circ, \pm 20^\circ, \pm 10^\circ$	0.01 degrees

Table A.2: Controllable parameters for the Fujinon/Photometrics camera system and jig

parameter	operating range	control resolution
focus distance	1.0 m - $\infty$	5100 steps (143° of travel)
focal length	10 - 130 mm (13×)	11100 steps (89° of travel)
aperture	$f/1.7$ - closed	2700 steps (76° of travel)
color band	none, red, green, blue	6 filters
$x, y, z$	24 in, 12 in, 80 in	0.001 inches
$\theta_x, \theta_y, \theta_z$	$\pm 20^\circ, \pm 20^\circ, \pm 10^\circ$	0.01 degrees
exposure time	0.0 - 6553.6 seconds	0.1 seconds
sensor gain	20 or 5 electrons / LSB	high/low

The Fujinon/Photometrics camera system and camera jig provide the same 10 DOFs as the Cosmicar/Panasonic system, plus the ability to set the sensor's gain and exposure time. The automation for the Fujinon lens also has 30% more control resolution in the focus and focal length DOFs. The 12 DOFs are summarized in Table A.2.

# Appendix B

## Chromatic Aberration Measurement

Chromatic aberration in camera lenses causes differences in the magnification, focus and centering of images taken in different color bands. In this appendix we describe the procedure we developed for measuring the level of lateral chromatic aberration in a camera lens using a single color image of a simple test target. As described the results for tests on 12 different camera lenses.

### B.1 Measurement approach

To measure the level of lateral chromatic aberration in a camera lens we use the lens to take a color image of a black-on-white checkerboard test target such as the one shown in Fig. B.1. If there were no lateral chromatic aberration in the lens the positions of the black-to-white edges would be the same in each image band. Any difference in the position of the edge between image bands provides a measure of the lens's lateral chromatic aberration perpendicular to the direction of the edge. By using a relatively dense set of vertical and horizontal black-to-white edges we can approximate the magnitude of the misregistration caused by lateral chromatic aberration at any position in the image by using the nearest pairs of horizontal and vertical edges to obtain the orthogonal components of the misregistration.

To measure lateral chromatic aberration we use one band of the checkerboard's color image to identify a set of *reference edges* whose horizontal or vertical positions can accurately and reliably be measured. We start by using  $3 \times 3$  Sobel operator and a threshold to produce a binary edge map from the checkerboard image (Fig. B.2). We then look for suitable reference edges in the image. Reference edges must meet four criteria: they must be fairly close to horizontal or vertical; they should have sufficient length so that any image shifting in the direction parallel to the edge will still leave enough of the edge to measure the cross edge shifting; they should be sharp enough for the edge position to be accurately calculated; and all horizontal and vertical reference edges should be chosen far enough apart so that there is no ambiguity in tracking them between image bands.

To pick out reference edges we test the edge image using a  $21 \times 21$ -pixel test mask. The test

mask for finding the horizontal edges used for detecting vertical shifting, shown in Fig. B.3, contains one *edge area* and two *non edge areas*. If the region of the image under the test mask contains an edge along the mask's edge area, and no edge in the mask's non edge area, and if the mask does not overlap any previously successful masks, then the section of the edge under the mask is labeled as a reference edge. Given a reference edge, the image column (or row) running across the center of the edge is used to measure the position of the edge to sub-pixel accuracy in the three image bands. Figure B.4 shows the edge image with the vertical and horizontal test masks superimposed. The tick marks across the edges identify the columns and rows used to measure the positions of the edges. In this image there are 415 vertical reference edges and 275 horizontal reference edges.

Once we have identified a set of reference edges we next determine their vertical or horizontal positions to sub-pixel accuracy by using a standard Laplacian-of-Gaussian (LoG) edge localization technique. The LoG technique is nearly optimal for edge localization [15], is relatively easy to implement and is computationally very efficient. To locate the position of a reference edge a 21-pixel segment of the row or column crossing the reference edge is first convolved with the second derivative of a Gaussian convolution kernel<sup>1</sup>. The location of the convolution's zero crossing is then interpolated to give the exact position of the edge.

To display the misregistration between image bands we plot the difference between the positions of the edges in the blue and red images and the blue and green images. This can be done for both the vertical edges and the horizontal edges. If the vertical and horizontal reference edges are close enough we can approximate the total magnitude of the misregistration at any point in the image by using the misregistration at nearest pairs of vertical and horizontal reference edges as the orthogonal components.

Figures B.5 and B.6 show the vertical misregistration between the blue and red and the blue and green images respectively. Figures B.7 and B.8 show the horizontal misregistration between the blue and red and the blue and green images respectively. Finally, Figures B.9 and B.10 show the total misregistration between the blue and red and the blue and green images respectively.

---

<sup>1</sup>We use a standard deviation of 4.0 pixels for the Gaussian convolution kernel.

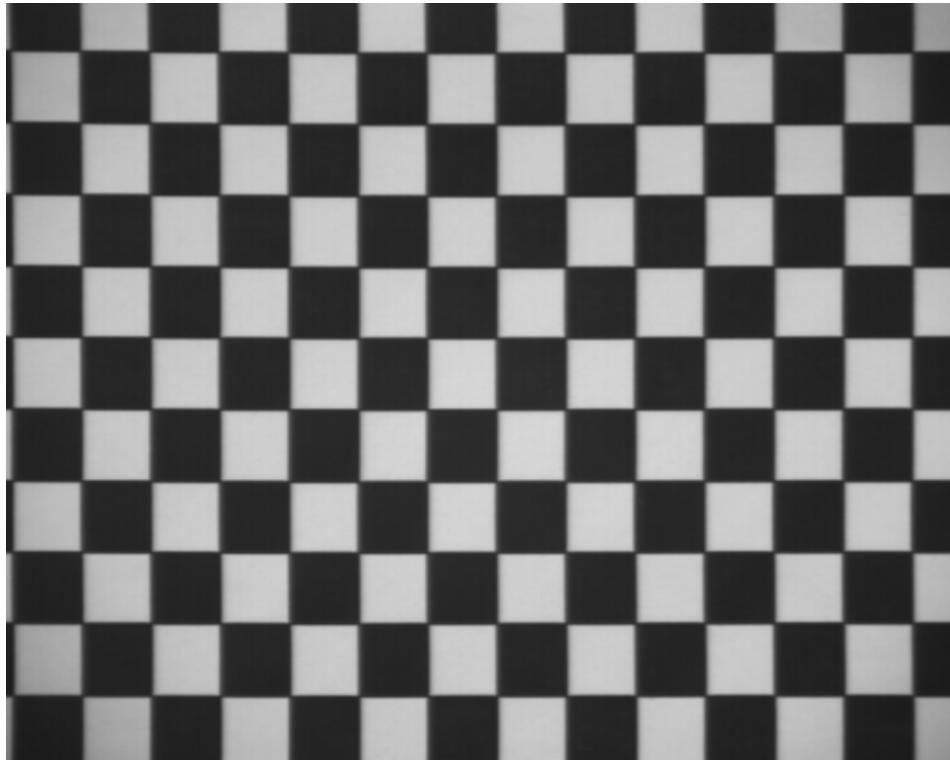


Figure B.1: Checkerboard target used for measuring lateral chromatic aberration

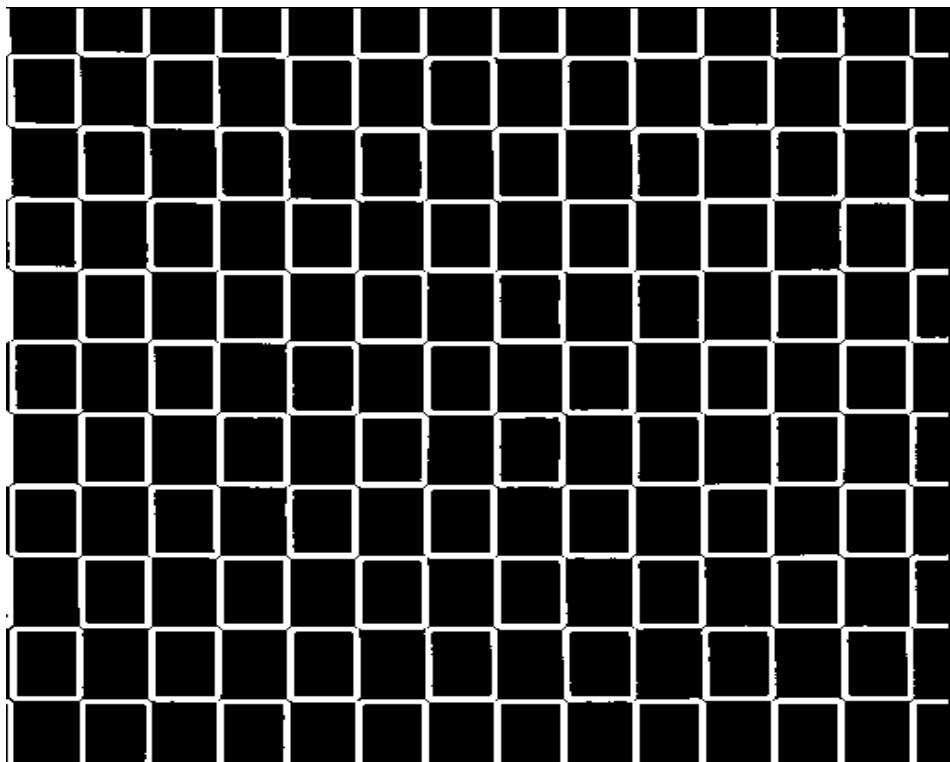


Figure B.2: Thresholded edge image of checkerboard

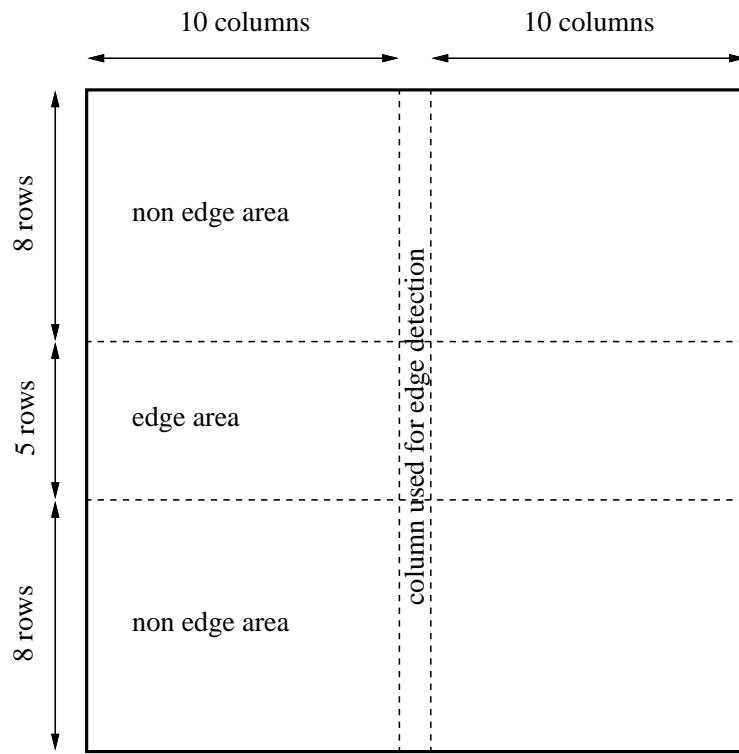


Figure B.3: Test mask for finding horizontal reference edges on checkerboard

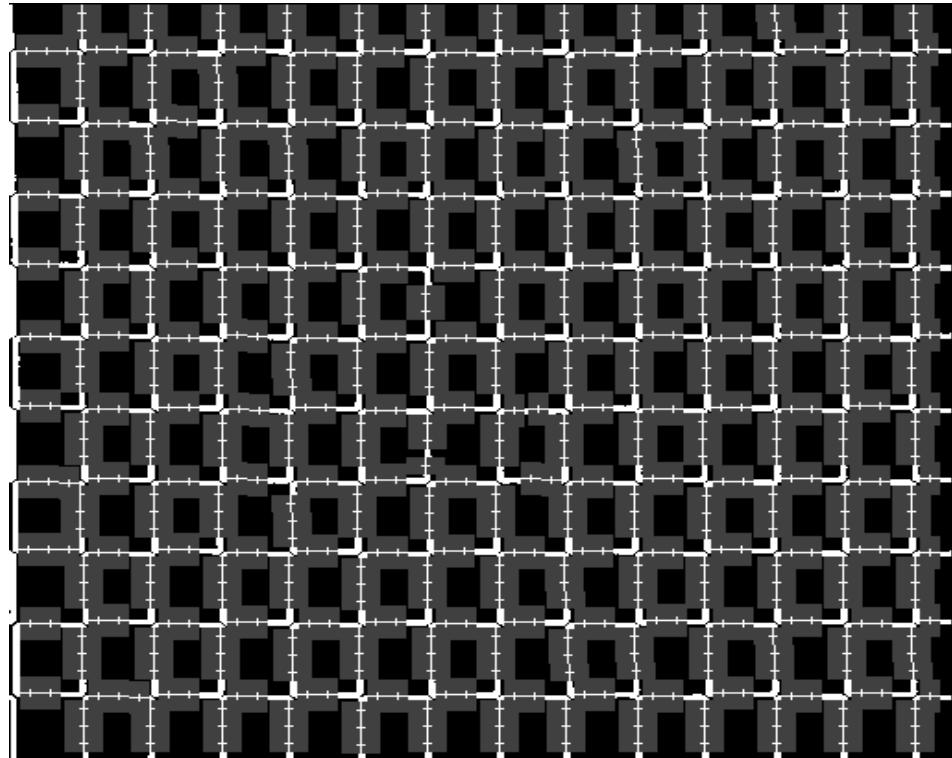


Figure B.4: Edge image with highlighted reference edges

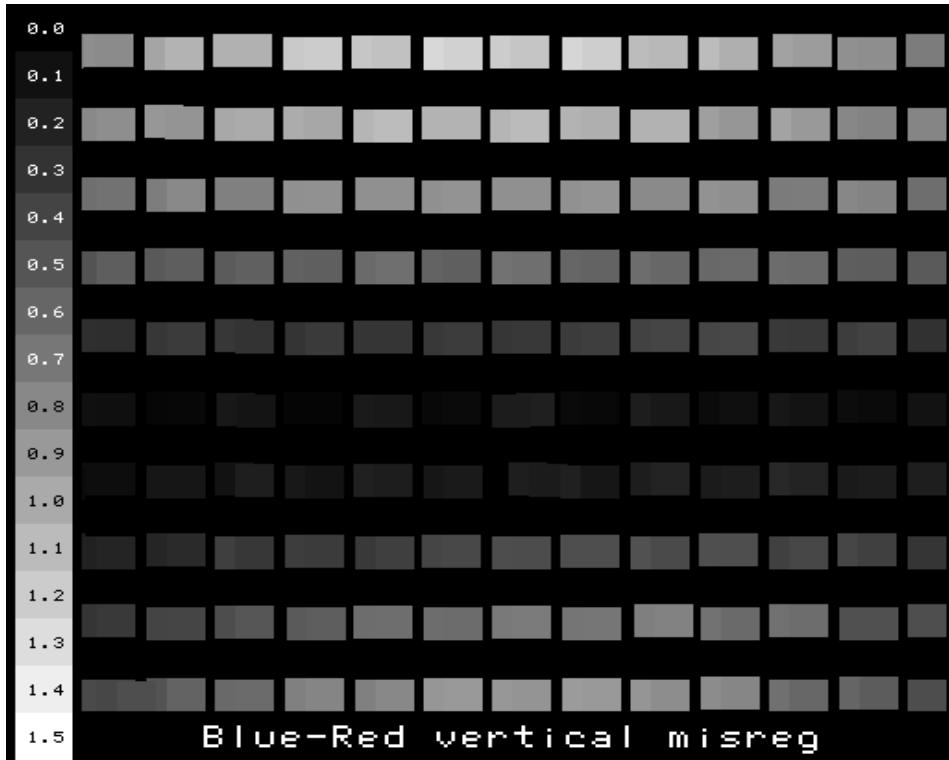


Figure B.5: Blue/Red vertical misregistration

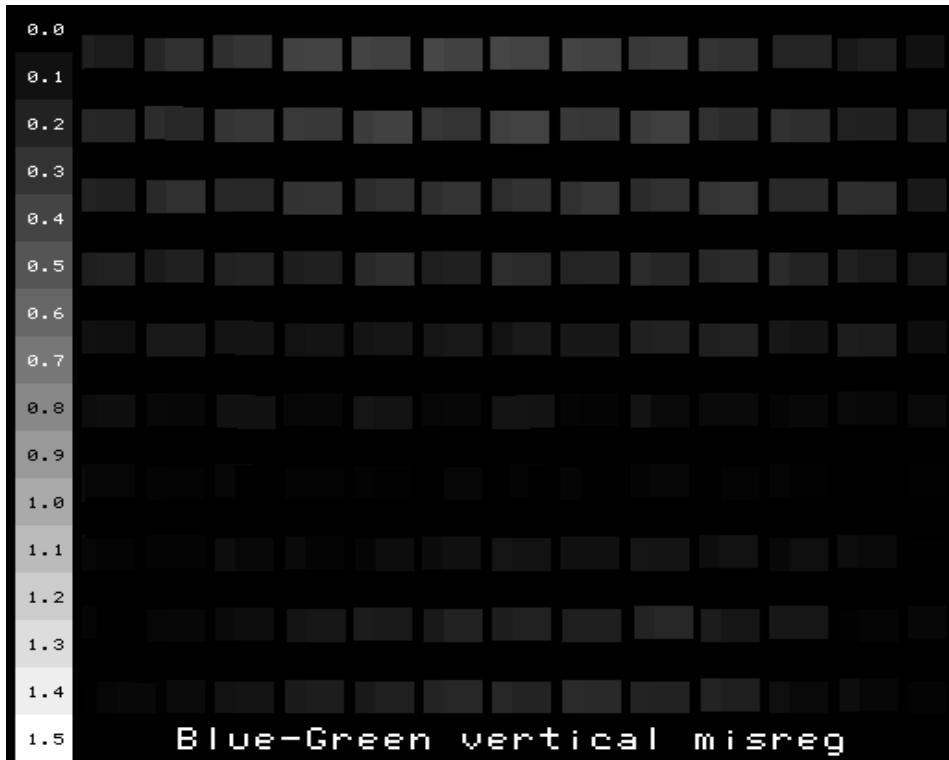


Figure B.6: Blue/Green vertical misregistration

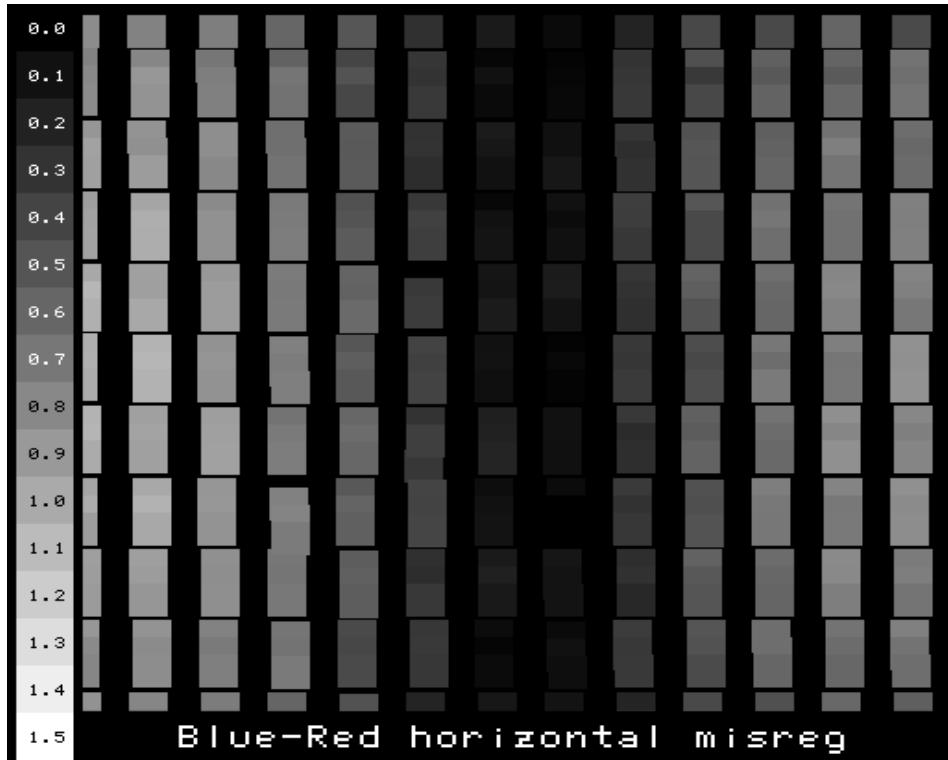


Figure B.7: Blue/Red horizontal misregistration



Figure B.8: Blue/Green horizontal misregistration

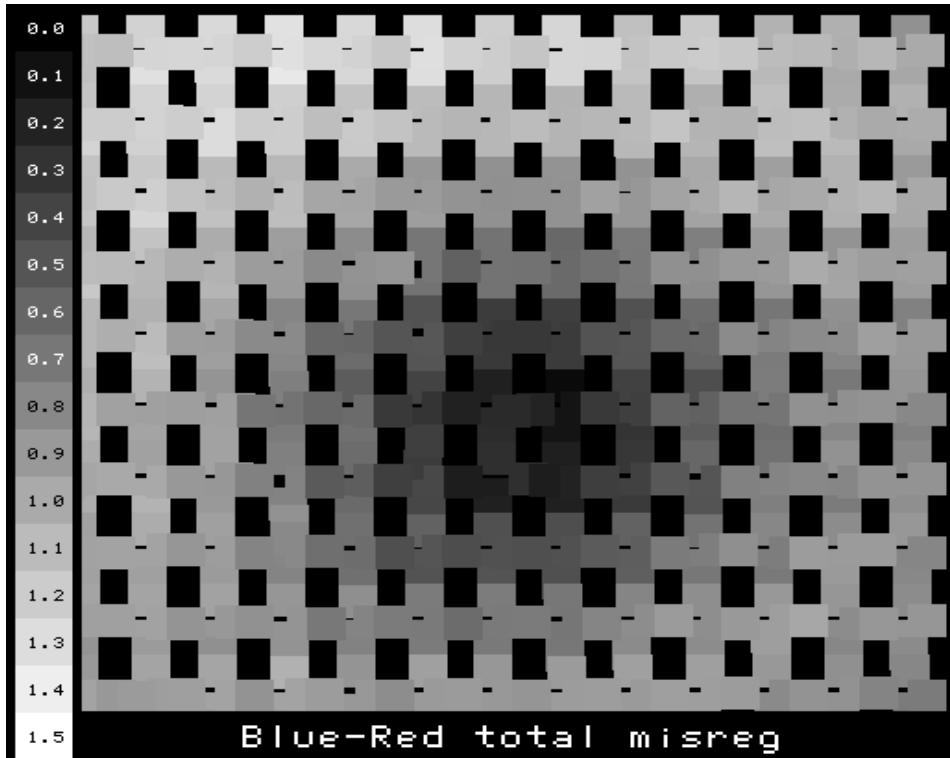


Figure B.9: Blue/Red total misregistration

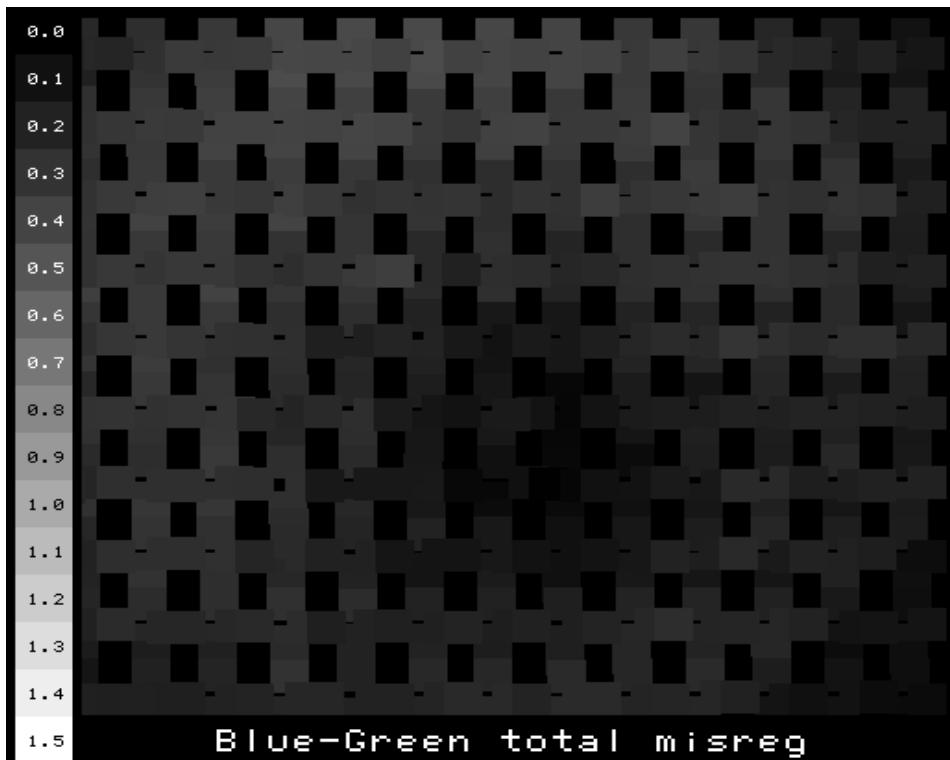


Figure B.10: Blue/Green total misregistration

## B.2 Examples

To determine the levels of lateral chromatic aberration that could be found in off-the-shelf lenses we measured a 12 different fixed and variable focal length lenses, including multiple copies of the same lens model. The results are summarized in Table B.1 and in Figs. B.11 to B.46.

Lens 1 is a CCTV grade lens typically used with black and white cameras for surveillance applications. Lenses 2 through 12 are 35 mm SLR camera lenses typically used in color photography and thus have better color correction.

Lenses 1, 2 and 3 are zoom lenses and thus the level of lateral chromatic aberration depends on the focal length setting. The lenses were measured at representative focal lengths.

lens	make and focal length	type	camera system
1	Kowa 12.5-75 mm at 55 mm	CCTV zoom	General Imaging + Matrox
2	Nikon 35-105 mm at 85 mm	35 mm SLR zoom	General Imaging + Matrox
3	Tokina 35-200 mm at 150 mm	35 mm SLR zoom	General Imaging + Matrox
4	Nikon 200 mm	35 mm SLR telephoto	General Imaging + Matrox
5	Nikon 50 mm AF Nikkor	35 mm SLR	General Imaging + Matrox
6	Nikon 50 mm Nikkor	35 mm SLR	General Imaging + Matrox
7	Nikon 50 mm Nikkor	35 mm SLR	General Imaging + Matrox
8	Nikon 24 mm Nikkor	35 mm SLR	General Imaging + Matrox
9	Nikon 20 mm AF Nikkor	35 mm SLR	Kodak Megaplus XRC
10	Nikon 28 mm Nikkor	35 mm SLR	Kodak Megaplus XRC
11	Nikon 55 mm micro-Nikkor	35 mm SLR	Kodak Megaplus XRC
12	Elicar 90 mm V-HQ	35 mm SLR	Kodak Megaplus XRC

Table B.1: Description of lenses measured for lateral chromatic aberration

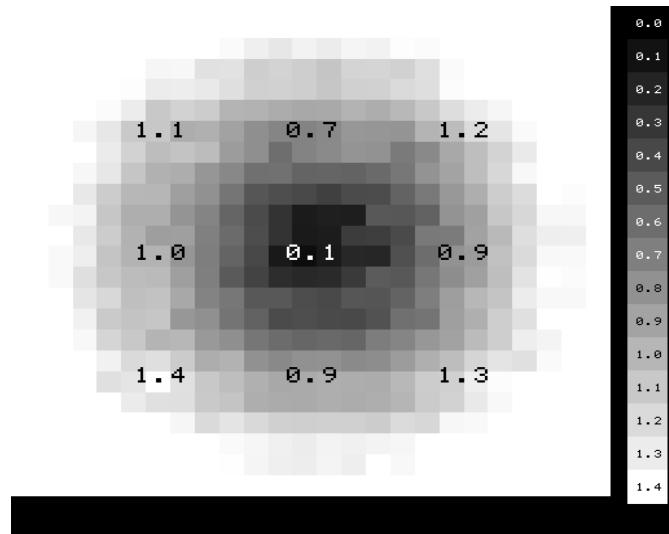


Figure B.11: Kowa 12.5-75 mm at 55 mm - blue/red misregistration

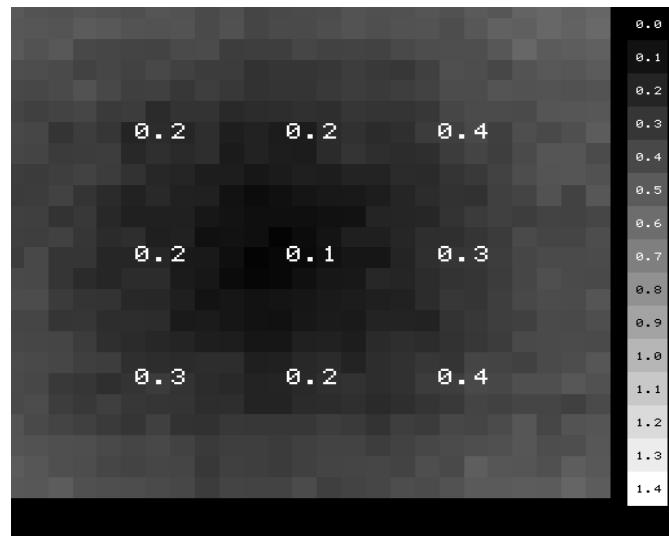


Figure B.12: Kowa 12.5-75 mm at 55 mm - blue/green misregistration

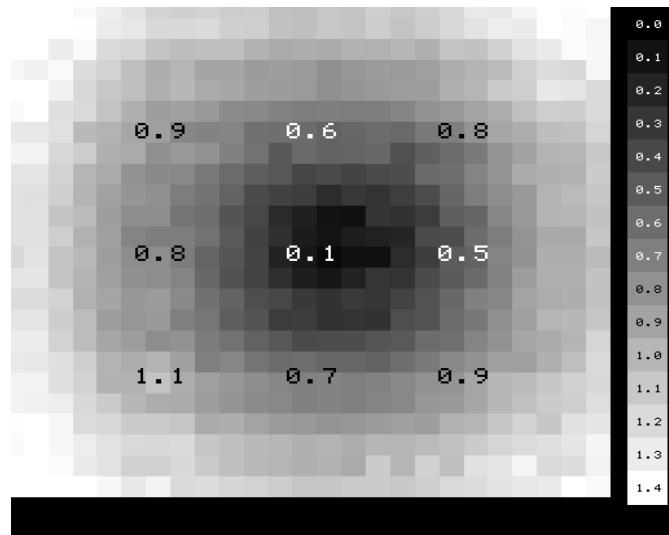


Figure B.13: Kowa 12.5-75 mm at 55 mm - red/green misregistration

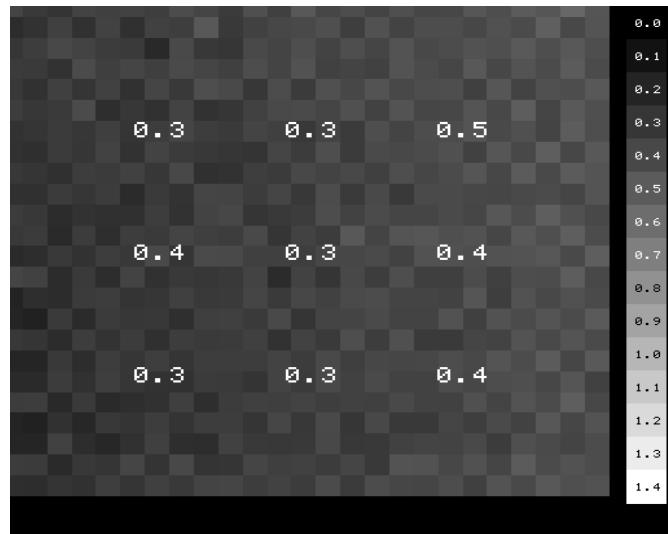


Figure B.14: Nikon 35-105 mm at 85 mm - blue/red misregistration

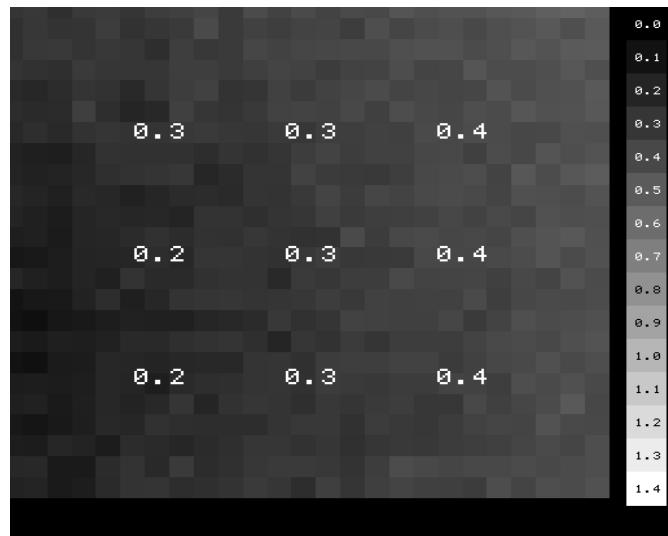


Figure B.15: Nikon 35-105 mm at 85 mm - blue/green misregistration

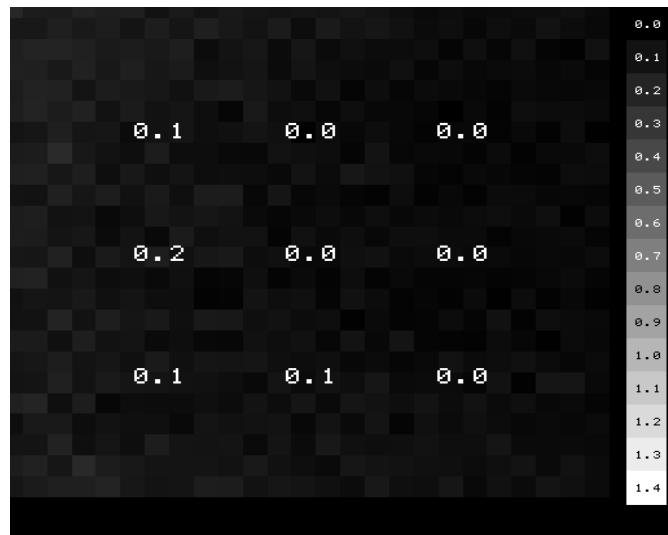


Figure B.16: Nikon 35-105 mm at 85 mm - red/green misregistration

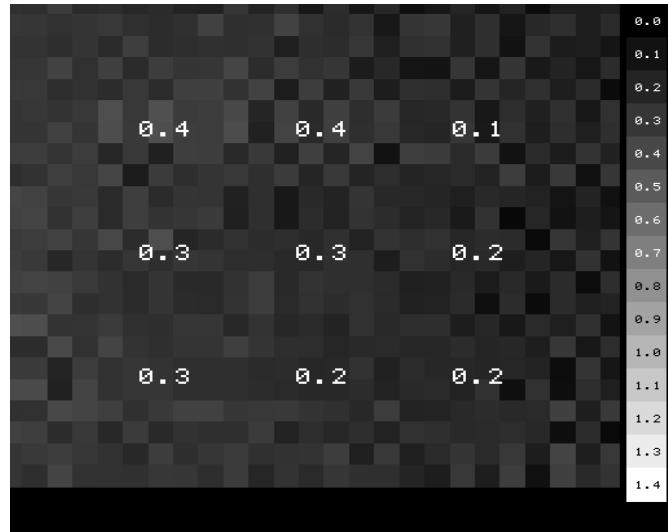


Figure B.17: Tokina 35-200 mm at 150 mm - blue/red misregistration

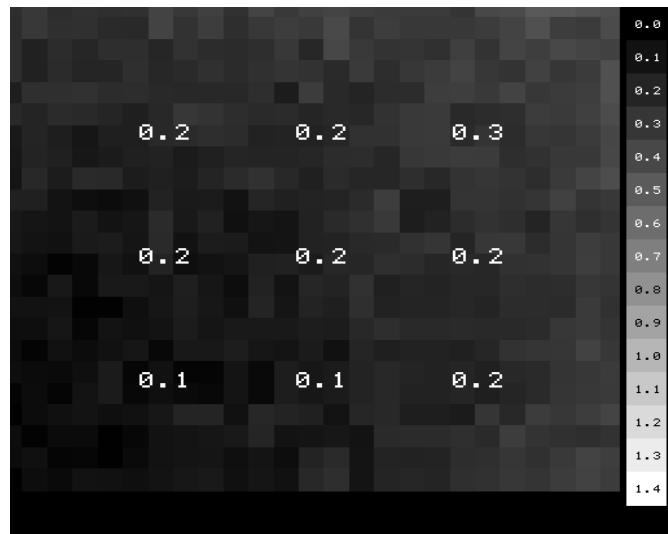


Figure B.18: Tokina 35-200 mm at 150 mm - blue/green misregistration

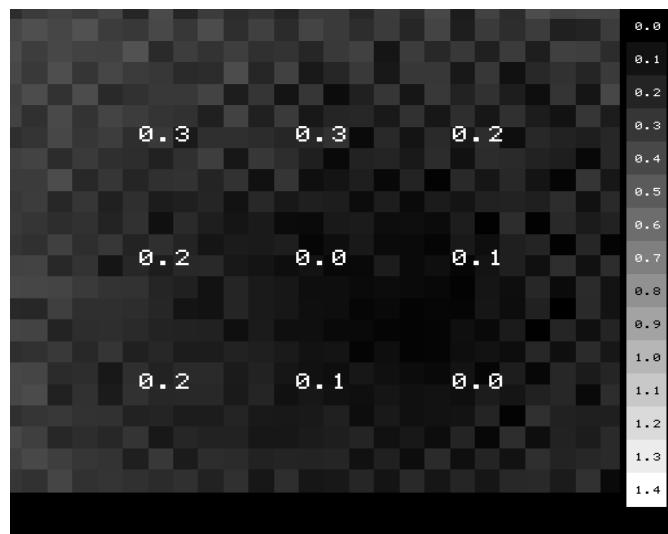


Figure B.19: Tokina 35-200 mm at 150 mm - red/green misregistration

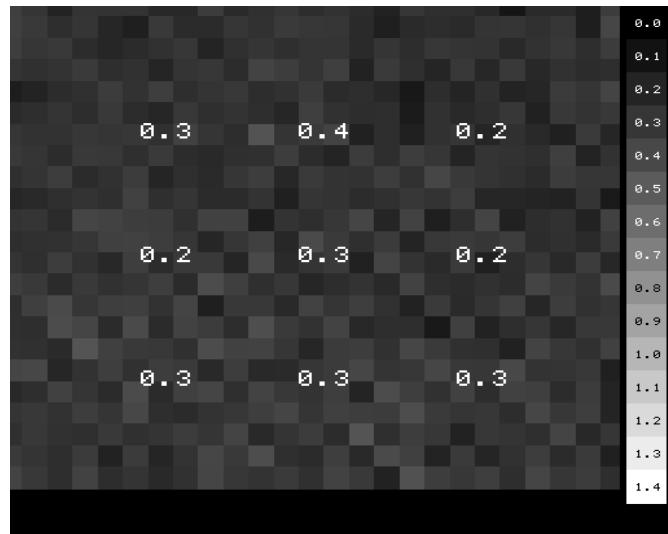


Figure B.20: Nikon 200 mm - blue/red misregistration

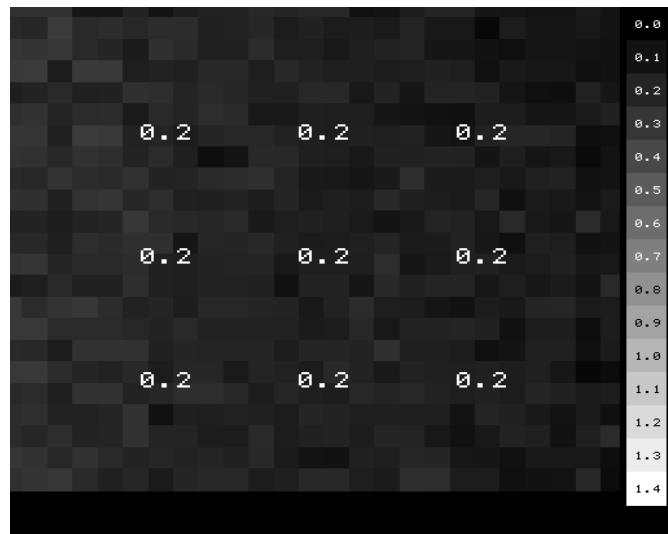


Figure B.21: Nikon 200 mm - blue/green misregistration

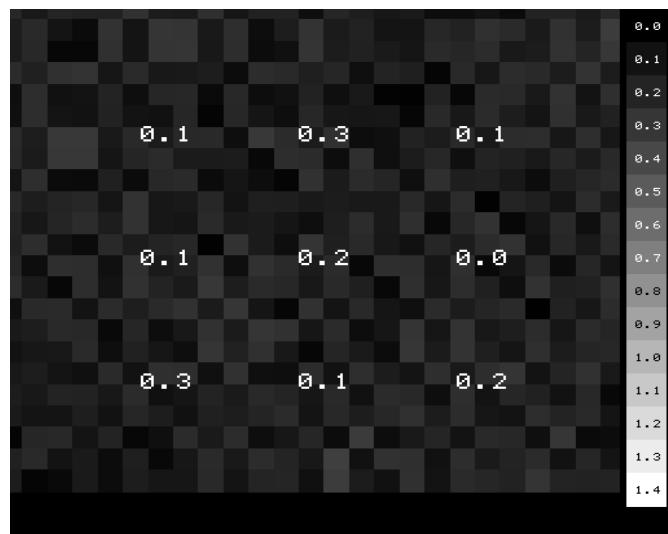


Figure B.22: Nikon 200 mm - red/green misregistration

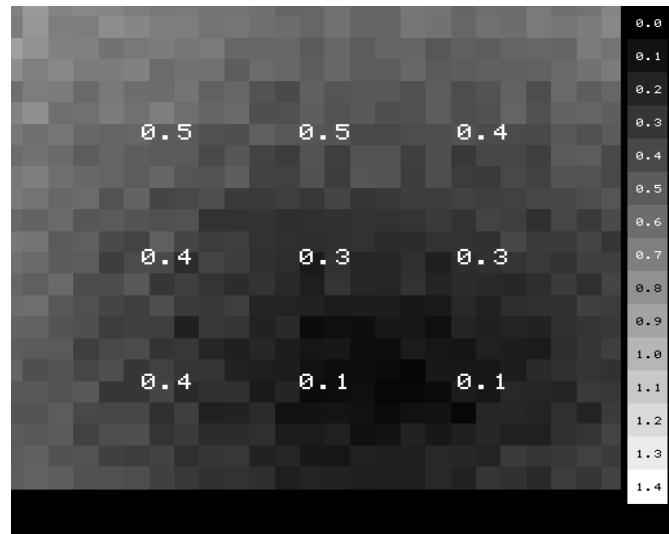


Figure B.23: Nikon 50 mm AF Nikkor - blue/red misregistration

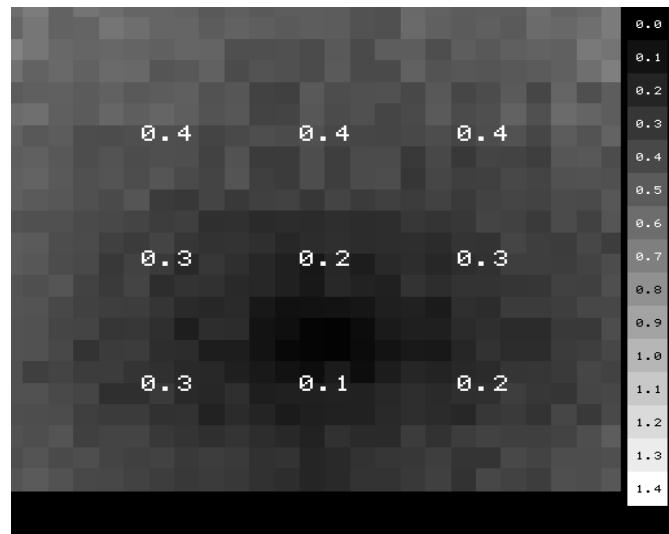


Figure B.24: Nikon 50 mm AF Nikkor - blue/green misregistration



Figure B.25: Nikon 50 mm AF Nikkor - red/green misregistration

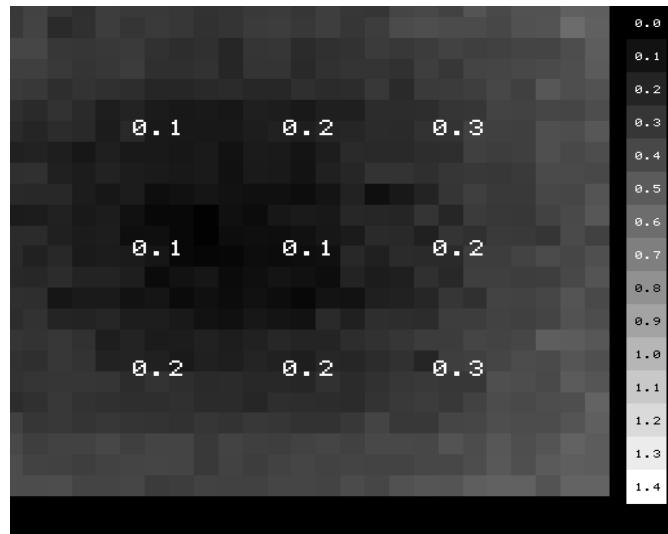


Figure B.26: Nikon 50 mm Nikkor - blue/red misregistration

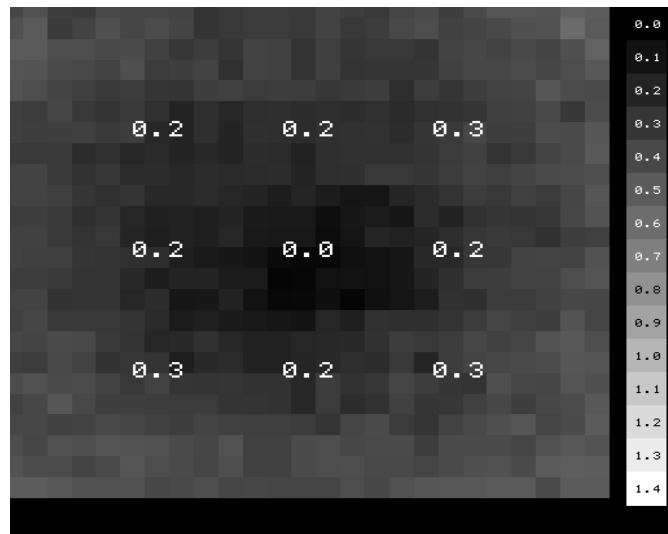


Figure B.27: Nikon 50 mm Nikkor - blue/green misregistration

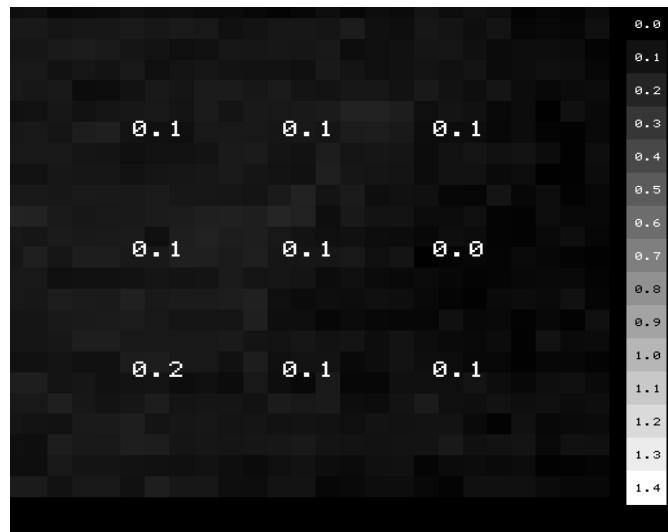


Figure B.28: Nikon 50 mm Nikkor - red/green misregistration

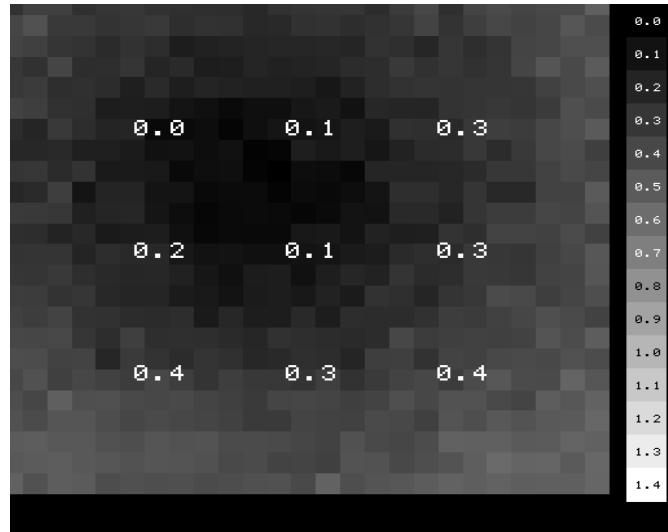


Figure B.29: Nikon 50 mm Nikkor - blue/red misregistration

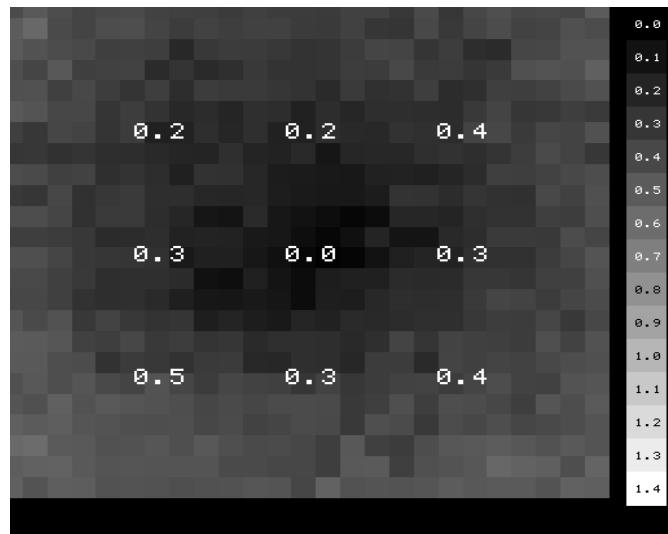


Figure B.30: Nikon 50 mm Nikkor - blue/green misregistration

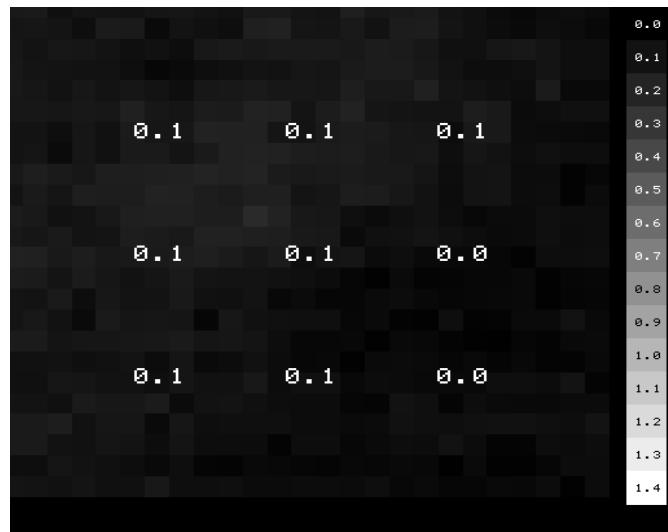


Figure B.31: Nikon 50 mm Nikkor - red/green misregistration

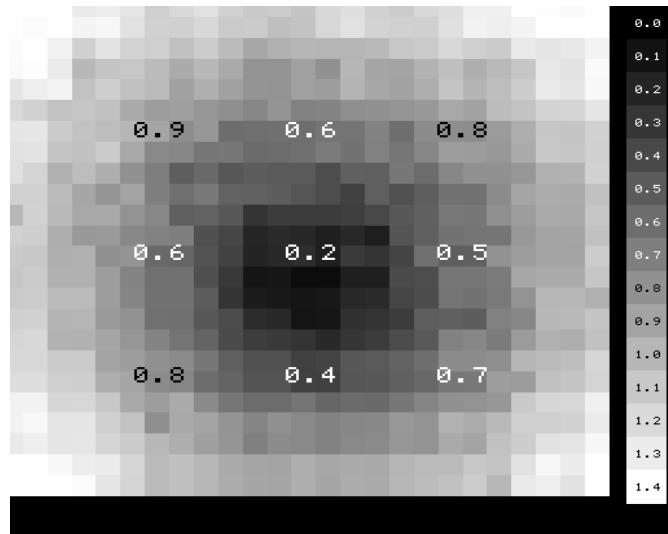


Figure B.32: Nikon 24 mm Nikkor - blue/red misregistration

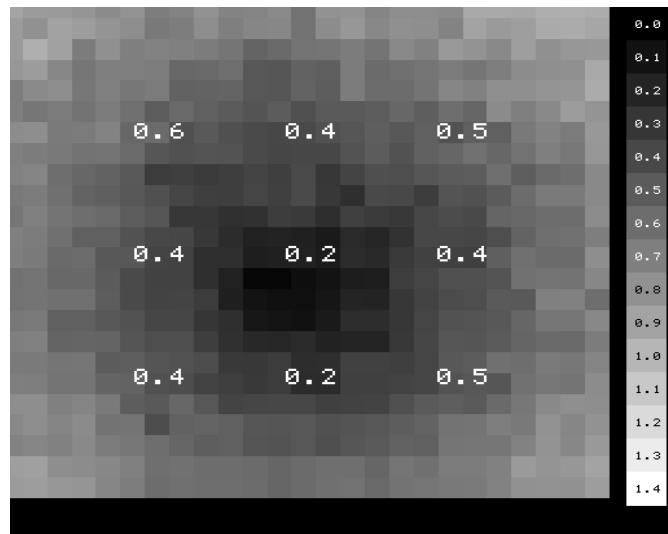


Figure B.33: Nikon 24 mm Nikkor - blue/green misregistration

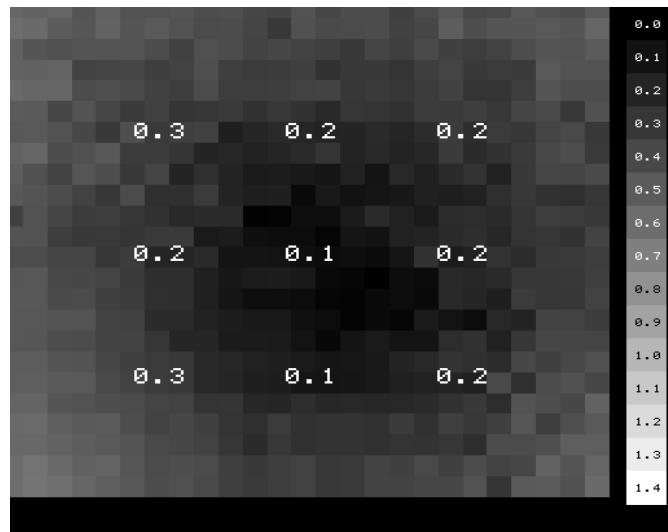


Figure B.34: Nikon 24 mm Nikkor - red/green misregistration

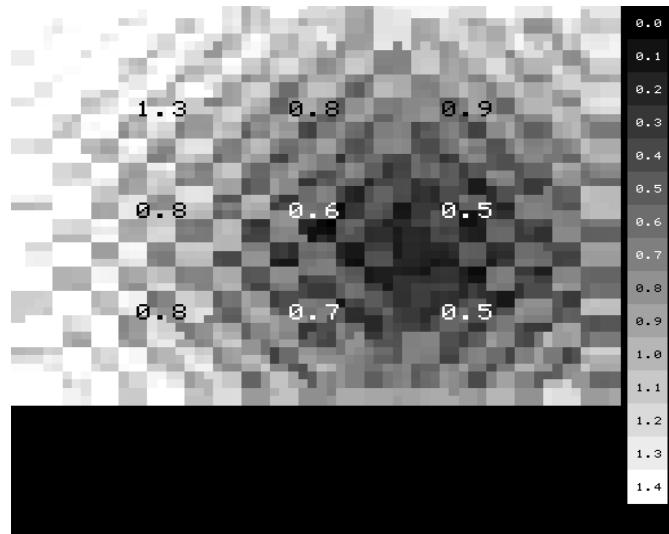


Figure B.35: Nikon 20 mm AF Nikkor - blue/red misregistration

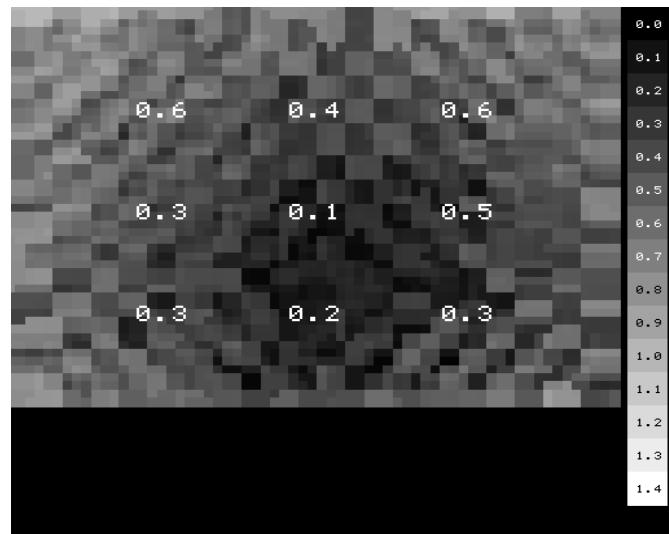


Figure B.36: Nikon 20 mm AF Nikkor - blue/green misregistration

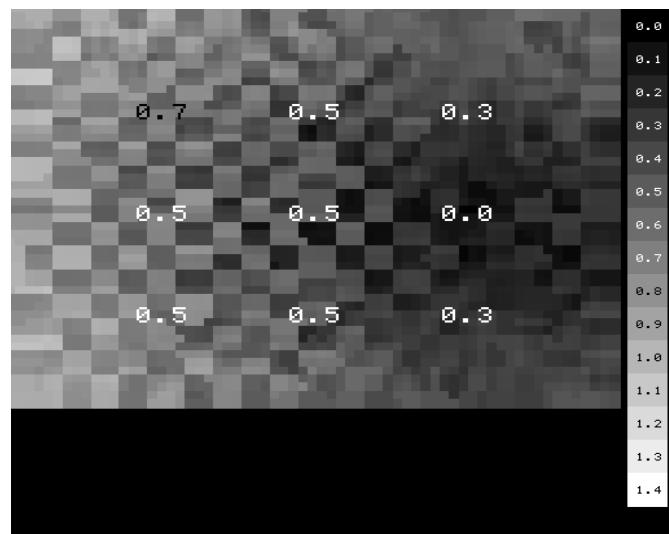


Figure B.37: Nikon 20 mm AF Nikkor - red/green misregistration

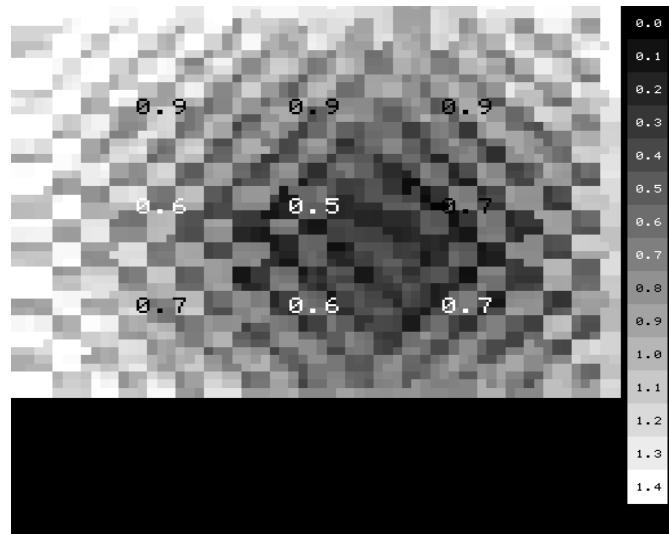


Figure B.38: Nikon 28 mm Nikkor - blue/red misregistration

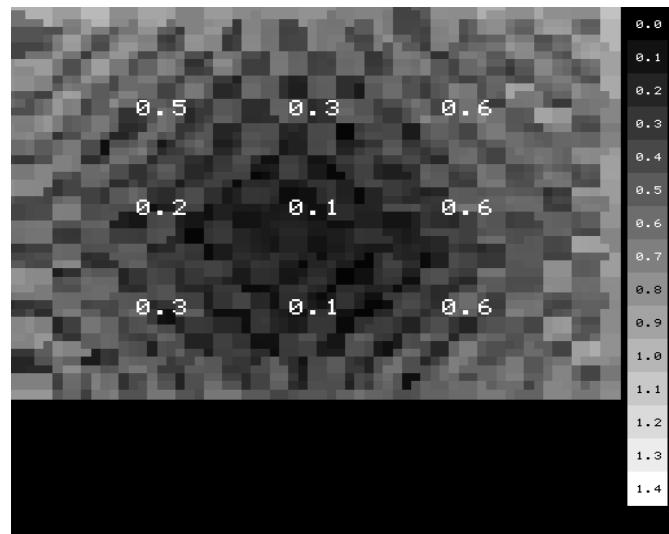


Figure B.39: Nikon 28 mm Nikkor - blue/green misregistration

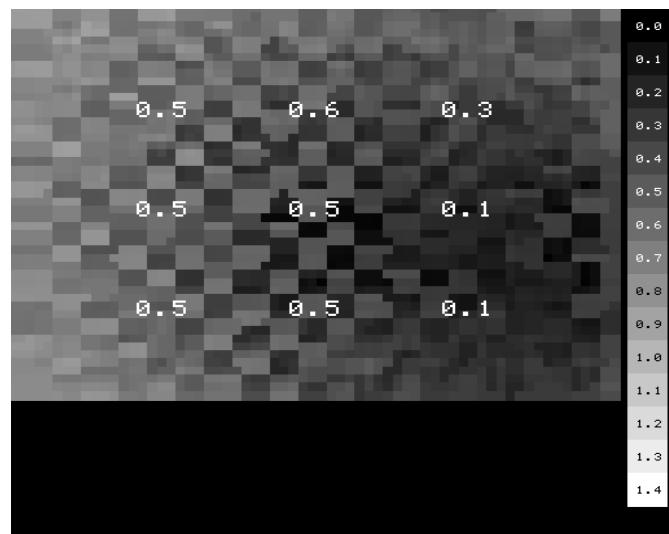


Figure B.40: Nikon 28 mm Nikkor - red/green misregistration

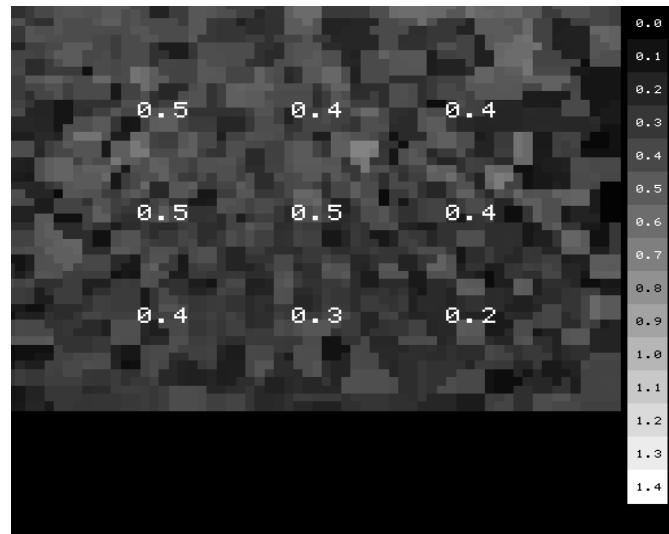


Figure B.41: Nikon 55 mm micro-Nikkor - blue/red misregistration

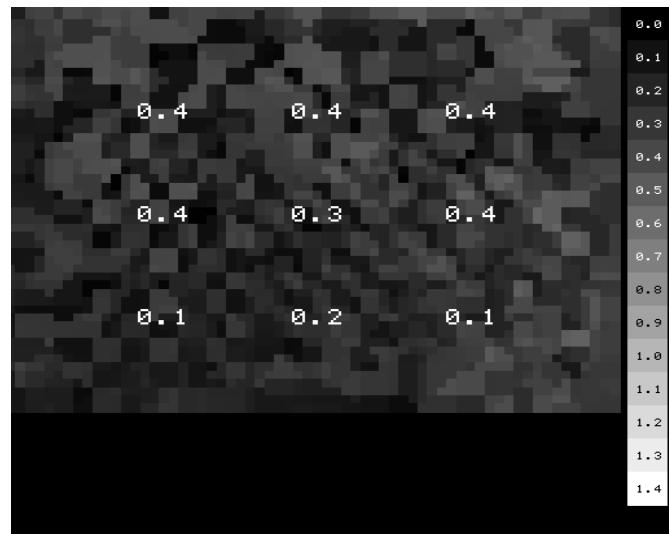


Figure B.42: Nikon 55 mm micro-Nikkor - blue/green misregistration

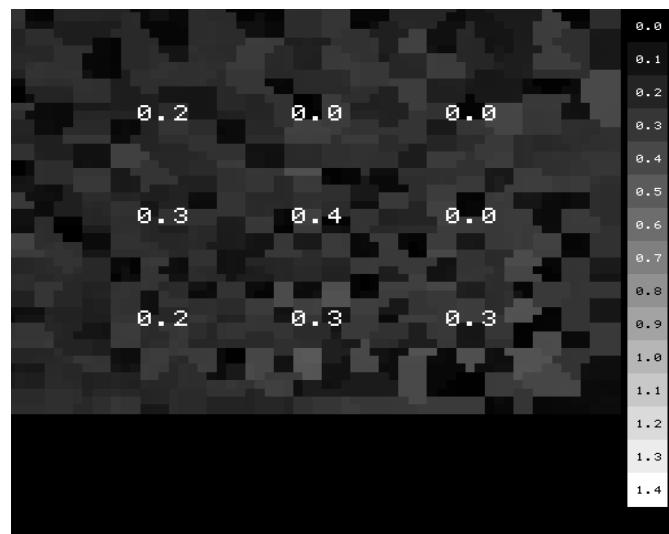


Figure B.43: Nikon 55 mm micro-Nikkor - red/green misregistration

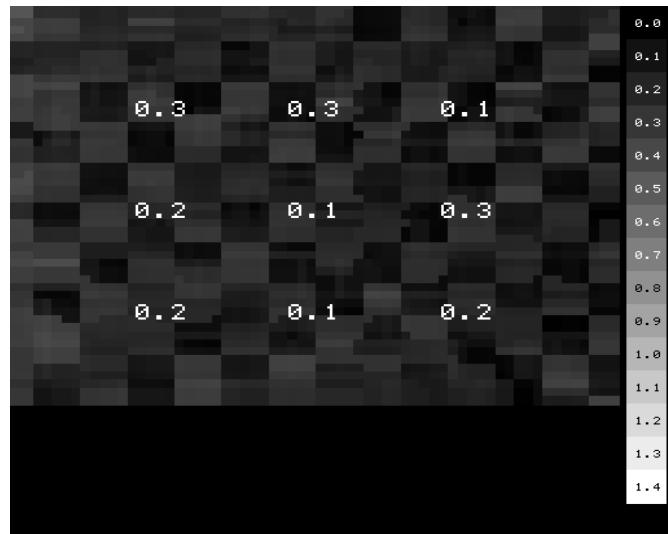


Figure B.44: Elicar 90 mm V-HQ - blue/red misregistration

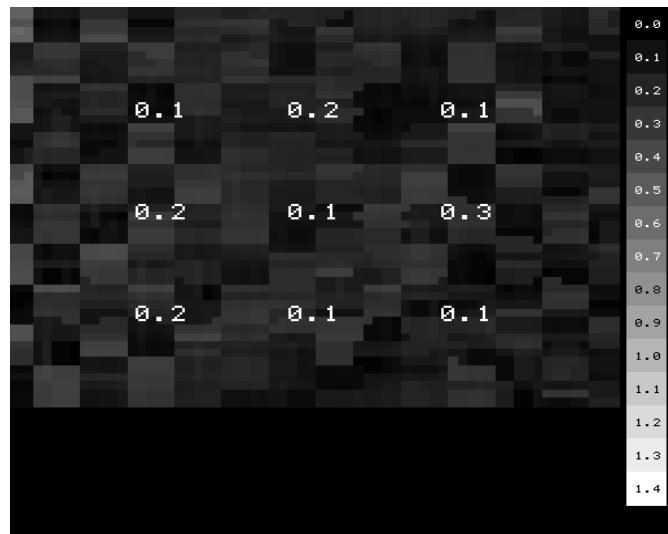


Figure B.45: Elicar 90 mm V-HQ - blue/green misregistration

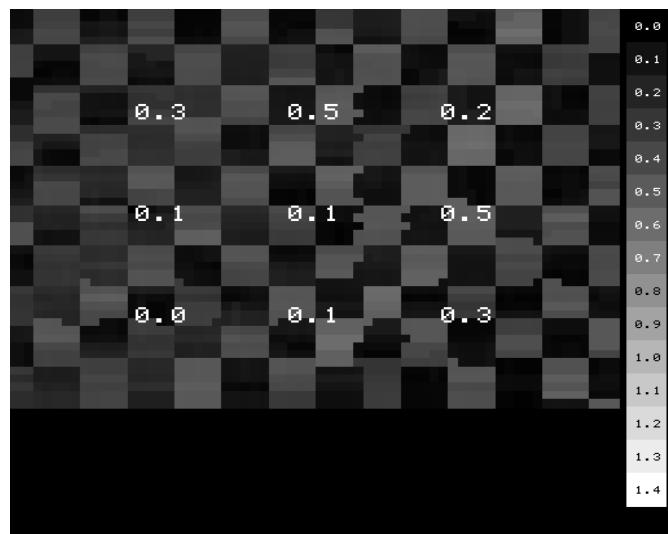


Figure B.46: Elicar 90 mm V-HQ - red/green misregistration

# Appendix C

## Sub-pixel Centroid Measurement

To obtain calibration data for our perspective-projection camera models we must measure the positions reference marks, such as the one shown in Fig. C.1, across wide ranges of both focus and magnification<sup>1</sup>. In this appendix we describe the method that we have developed to measure the centroids of circular dots to sub-pixel accuracy. This method is also used to measure the center of an autocollimated laser's image in Section 3.3.2.

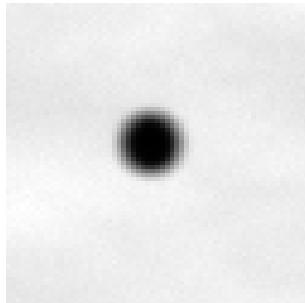


Figure C.1: Typical image of a black reference point on a white background

Many methods have been developed for measuring the centroids of dots or point sources [16][6][31]. To deal with the wide ranges of focus and magnification that are somewhat unique to our application we have developed a new approach.

To find the centroid of the dot we first find the centroids of a series of 1D cross sections of the dot in the row and column directions (i.e. we use the rows and columns crossing the dot's image). We then fit lines through the centers found for both directions. We consider the centroid of the dot to be the intersection of these two lines.

---

<sup>1</sup>Even though we may never intend to use our camera when the image is defocused, we may still need to collect calibration data under these conditions.

## The center of a 1D cross section

To find the center of each 1D cross section of the dot we use the zero crossings of the cross section's first and second derivatives. Figures C.2 and C.3 show the two cross sections that are possible for the image of a dot, along with the first and second derivatives of the cross sections.

In Fig. C.2 the width of the cross section of the dot is comparable to the width of the blur in the image. In this situation either the zero crossing point of the first derivative (the position of the peak in the data) or the midpoint between the two zero crossings of the second derivatives (the midpoint between the left and right inflection points) can be used as the center of the cross section.

In Fig. C.3 the width of the cross section of the dot is somewhat greater than the width of the blur in the image. In this situation the zero crossing of the first derivative is of no use (there is no well defined peak in the data) and we must rely on the midpoint between the two zero crossings of the second derivative for the center of the cross section.

To obtain the first and second derivatives ( $F'(n)$  and  $F''(n)$ ) of the cross section  $F(n)$  we use the convolution properties

$$\begin{aligned}(F * G)' &= F * (G') \quad \text{and} \\ (F * G)'' &= F * (G'')\end{aligned}$$

where  $G$  is a Gaussian function and  $G'$  and  $G''$  are the first and second derivatives of the Gaussian.

The first derivative Gaussian convolution kernel is calculated using

$$G'(n) = \frac{-n}{\sqrt{2\pi}\sigma^3} \exp^{\frac{-n^2}{2\sigma^2}}$$

The second derivative Gaussian convolution kernel is calculated using

$$G''(n) = \frac{n^2 - \sigma^2}{\sqrt{2\pi}\sigma^5} \exp^{\frac{-n^2}{2\sigma^2}}$$

For both kernels we use a  $\sigma$  of 3 pixels and a kernel length of 31 pixels.

For the data shown in Figs. C.2 and C.3 (i.e. a white dot on a black background) we obtain the first derivative center  $C_1$  using

$$C_1 = C_{1_{high-to-low}}$$

where  $C_{1_{high-to-low}}$  is the simple linear interpolation of the strongest high-to-low zero crossing of  $F'$ . As a measure of the strength of  $C_1$  we use

$$C_{1_{strength}} = \frac{|F''(C_1)|}{\text{MAX}(|F''|)}$$

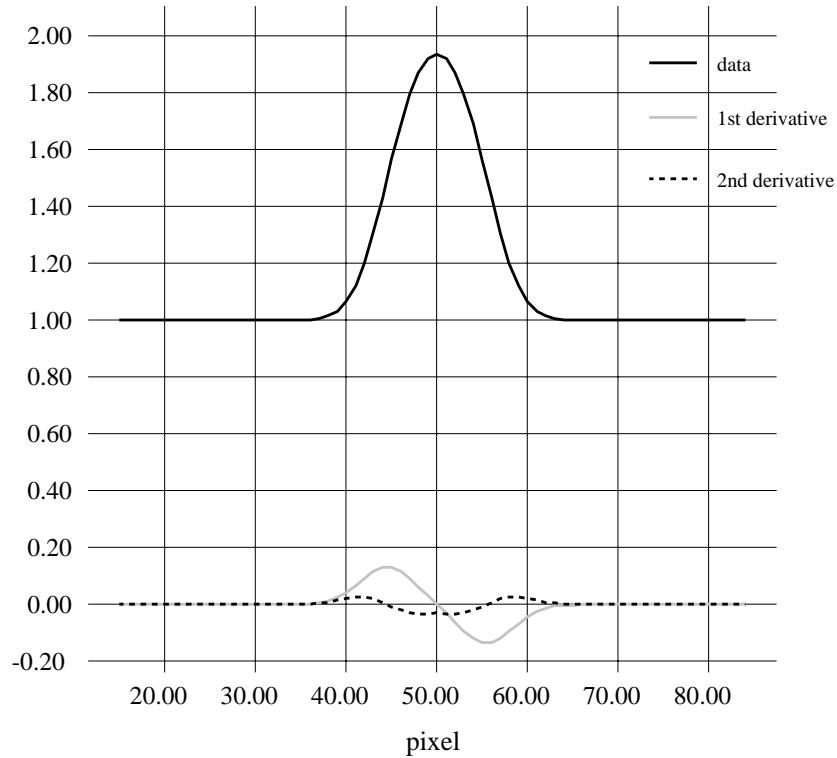


Figure C.2: Intensity profiles when blur width is greater than width of dot's cross section

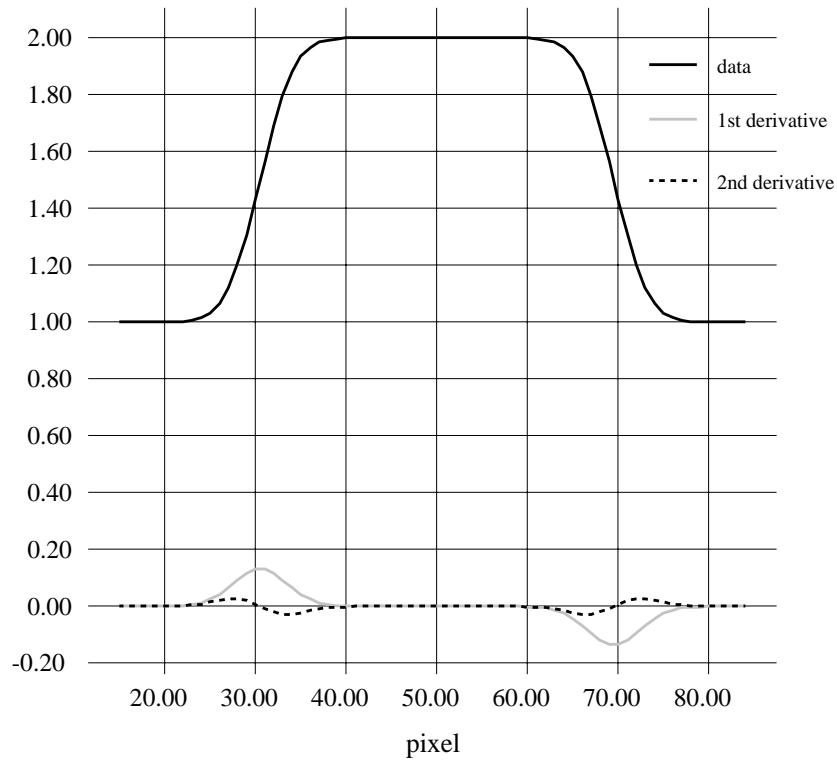


Figure C.3: Intensity profiles when blur width is less than width of dot's cross section

To calculate the second derivative center  $C_2$  we use

$$C_2 = \frac{C_{2_{high-to-low}} + C_{2_{low-to-high}}}{2}$$

where  $C_{2_{high-to-low}}$  and  $C_{2_{low-to-high}}$  are the positions of the two second derivative zero crossings. The two second derivative zero crossings are calculated using simple linear interpolations of the strongest high-to-low and low-to-high zero crossings of  $F''$ . As a measure of the strength of  $C_2$  we use

$$C_{2_{strength}} = \frac{\text{MIN}(|F'(C_{2_{high-to-low}})|, |F'(C_{2_{low-to-high}})|)}{\text{MAX}(|F'|)}$$

To determine the center of the cross section we use the weighted sum

$$C = \frac{C_1 C_{1_{strength}} + C_2 C_{2_{strength}}}{C_{1_{strength}} + C_{2_{strength}}}$$

As a measure of the strength of  $C$  we use

$$C_{strength} = \text{MAX}(|F'|)$$

For a black dot on a white background the sense of the high-to-low and low-to-high are reversed.

## The centroid of the 2D dot

To find the center of the dot we first smooth the dot's image with a 2D Gaussian convolution kernel to suppress any per-pixel noise. Next we find the centroids (and their strengths) for all cross sections along the rows and columns of the image. We then find the maximum strength for all centroids and discard any centroids whose strength is less than 5% of this. This effectively removes measurements from cross sections that did not contain any part of the dot. Following this we fit lines to the centroids in each of the row and column directions using standard weighted least squares error. For each of the data points we use a weight of  $(1/C_{strength})^2$ . Finally, we calculate the centroid of the dot by finding the intersection of the two fitted lines.

As a quick test of the algorithm we took a series of 9 images of a black dot on a white background using the Fujinon/Photometrics camera system. The test was repeated for a total of three different levels of defocus. In all of the tests the standard deviation of the  $x$  and  $y$  coordinates of the dot's centroid were well under 0.01 pixels.

The advantages of this algorithm over other approaches are 1) it works for a wide range of dot sizes without having to be retuned and 2) it works across a wide range of defocus. The primary disadvantage of this algorithm is its computational cost.

# Appendix D

## Aperture Calibration

The aperture in a camera system affects both the radiometric and the geometric imaging properties of the system. In this appendix we describe the relative and photometric aperture models that we have developed for our cameras. We also discuss the development of a depth-of-field model. Finally we discuss why aperture setting may be a necessary parameter in any geometric camera model.

### D.1 Aperture stops

In an ideal pinhole camera exactly one ray of light from a point in object space will pass through the camera's pinhole to strike the image plane (see Fig. D.1). With a lens however, light from a point in object space is collected from a solid angle of rays and projected through the lens onto the image plane, as illustrated in Fig. D.2. The extent of this solid angle of rays is limited by the lens elements and by the diameter of any diaphragms along the optical path. The limiting diaphragm is called the *aperture stop* of the lens.

Typically the aperture stop of a lens (i.e. the iris) is constructed from a set of mechanical blades whose positions can be varied to change the stop's diameter[21]. The Fujinon and Cosmicar lenses used in our work both use six blade apertures. Both lenses have 2700 steps of aperture position.

### D.2 Relative and photometric aperture

The relative aperture or *f*-number of a lens is defined as

$$f\text{-number} = \frac{f}{a}$$

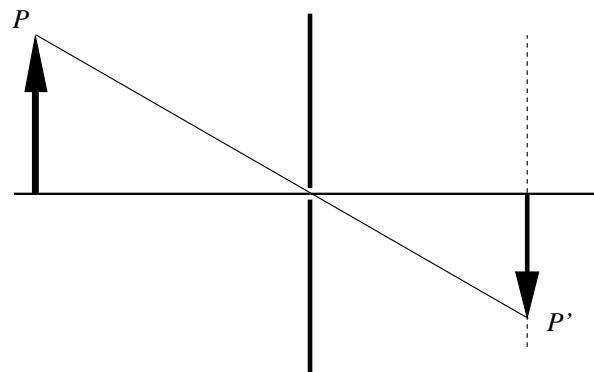


Figure D.1: Basic pinhole model

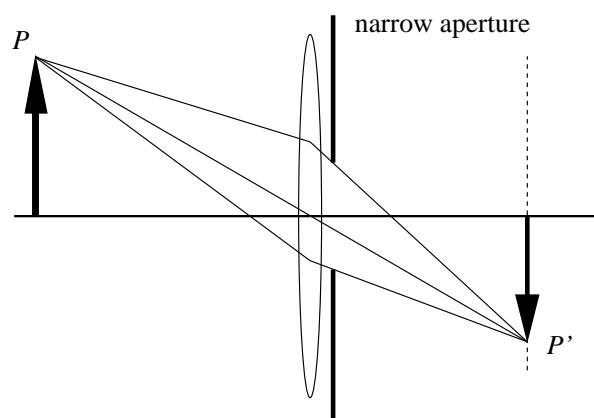
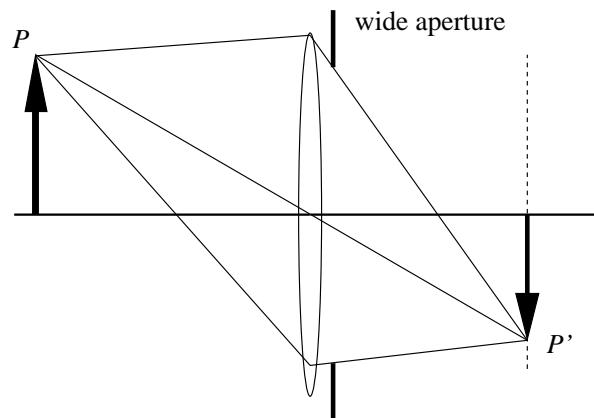


Figure D.2: Basic thin-lens model with wide and narrow apertures

where  $f$  is the effective focal length<sup>1</sup> and  $a$  is the diameter of the clear aperture or entrance pupil of the lens. The clear aperture of the lens is the *image* of the aperture stop as it would be seen from an axial position in front of the lens.

In the ideal case the illumination of the image plane is proportional the area of the clear aperture of the lens and inversely proportional to the square of the relative aperture. In practice, inter-reflections between lens elements reduce the amount of light reaching the image plane. For more precise radiometric work the photometric aperture, or T-number, is used. The T-number of a lens is defined as

$$\text{T-number} = \frac{f\text{-number}}{\sqrt{t}}$$

where  $t$  is the lens transmittance. The transmittance is the ratio of the light flux leaving the lens to the light flux entering the lens.  $t$  can be determined using the expression

$$t = \frac{1 - r}{1 + (K - 1)r}$$

where  $r$  is the reflectance of each lens surface and  $K$  is the number of glass/air interfaces in the lens[24].

If we could measure or estimate values for  $f$  and  $a$  at different aperture settings then we could directly formulate and calibrate a relative aperture model. Unfortunately there is no easy way to measure these quantities short of removing the lens from the camera and mounting it on an optical test bench as Tarabanis did to measure the optical parameters of his lens[49]. However, we can measure the relative intensity of light passing through the lens. If we know the T-number for the light intensity at one aperture setting we can use the relative light intensity at other aperture settings to calculate their respective T-numbers.

To collect data for our model we took images of an extended light source at 55 aperture settings from 2700 motor units (fully open) to 0 motor units (fully closed). For each aperture setting the exposure time was adjusted so that the maximum pixel value in the image was just below full scale (4095 greyscales for our 12-bit camera). The ratio of the the pixel values to the exposure time provides a measure of the light flux striking the image sensor in greyscales/second.

Figure D.3 shows a plot of the light flux versus aperture setting for the Fujinon/ Photometrics camera. At settings below 200 a pair of cutoff blades engage to completely close off the aperture. At settings above 2500 the blades comprising the aperture have completely cleared away from the bundle of light passing through the lens. At this point a fixed diaphragm located elsewhere along the lens acts as the lens' aperture stop. To bypass any transient effects we chose the operating region for our aperture models to be from 300 to 2450 motor units.

The Fujinon lens specifications list the maximum relative aperture ( $f$ -number) as 1.7 for focal lengths from 10-100 mm. Between 100 and 130 mm the maximum  $f$ -number rises to

---

<sup>1</sup>The effective focal length here is not equivalent to the focal length of the pinhole camera model used in Chapters 4, 5 and 6.

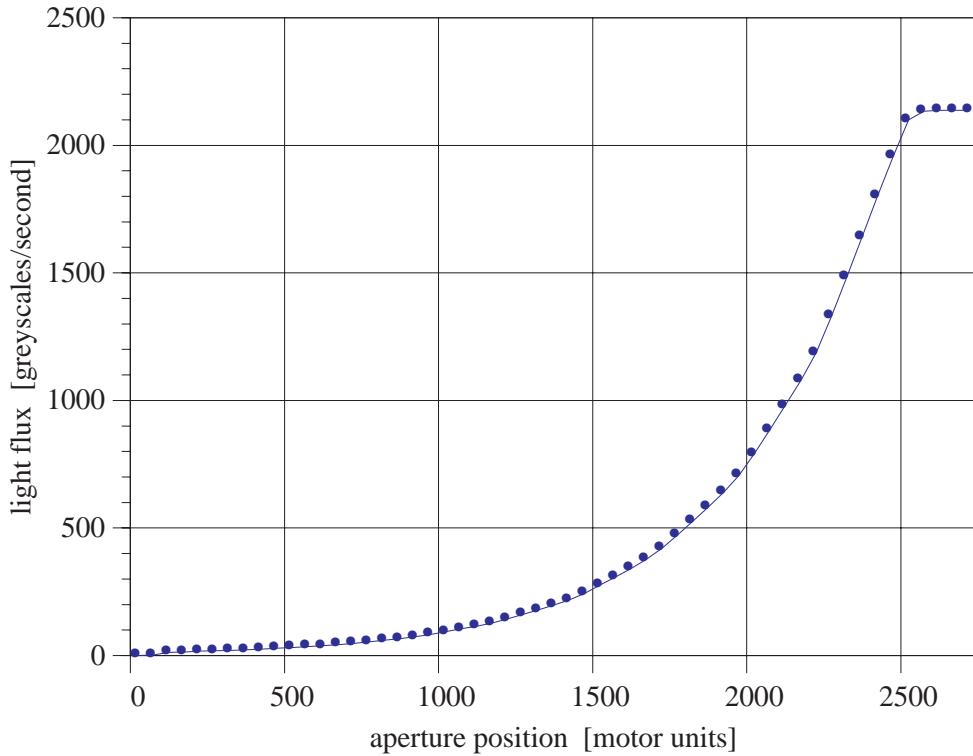


Figure D.3: Light flux versus aperture motor

2.1 indicating that within this range the limiting diaphragm for the lens is likely the outer diameter of one of the lens groups rather than the main fixed diaphragm. For our calibration data the zoom was set to 2000 motor units, which corresponds to a focal length of slightly under 100 mm. The maximum T-number for the Fujinon lens is specified as 1.9. The transmittance for the lens is thus

$$t = \left( \frac{1.7}{1.9} \right)^2 = 0.80 = 80\%$$

To model the photometric aperture we use the fact that the intensity of the light flux  $I$  at any aperture setting  $m_a$  is inversely proportional to the square of the T-number at that setting, i.e.

$$I(m_a) = c \left( \frac{1}{\text{T-number}(m_a)} \right)^2$$

where  $c$  is a constant of proportionality. In addition, from the calibration data (Fig. D.3) we know the level of the light flux,  $I_{MAX}$ , that corresponds to the maximum T-number of 1.9. Together this gives us the relationship

$$\text{T-number}(m_a) = 1.9 \sqrt{\frac{I_{MAX}}{I(m_a)}}.$$

Since the transmittance  $t$  of the lens is constant with aperture, the  $f$ -number relationship is

$$f\text{-number}(m_a) = 1.7 \sqrt{\frac{I_{MAX}}{I(m_a)}}.$$

To determine a formulation for  $I(m_a)$  we replot the calibration data on a logarithmic scale. The plot, shown in Fig. D.4, is roughly linear between 300 and 2450 motor units indicating that  $I(m_a)$  can be represented by an exponential model. The model that we use is

$$I(m_a) = 2^{(c_1 + c_2 m_a + c_3 m_a^2)}.$$

To determine the calibration constants  $c_1$ ,  $c_2$ , and  $c_3$  we fit a least squares error quadratic curve to the base 2 logarithm of the flux data between settings 300 and 2450. Figures D.6 and D.7 show plots of the final  $f$ -number and T-number models.

Figure D.5 shows the percentage error between the original flux data and the calibrated flux model. The maximum error is 3.4%. The deviations between the model and the data are repeatable across different experiments indicating that they are a property of the aperture mechanism. For a more accurate flux model we could store the calibration data in a lookup table and use interpolation to determine values for intermediate settings.

The models we built for relative and photometric aperture were calibrated for one focus and zoom setting. Typically zoom lenses are designed to try and hold the  $f$ -number constant as the focal length of the lens is varied. However, this may not be practical in all cases. In this event the  $f$ -number and T-number models may also have to be calibrated across focus and zoom.

### D.3 Depth of field

In camera systems the finite resolution of the image sensor leads to a zone around the plane of best focus of the camera called the depth of field. Light rays originating from any point within this zone project to the same pixel in the camera and thus are indistinguishable from one another. Consider the thin lens shown in Fig. D.8. In this diagram the cone of rays from point  $P$  converge at point  $P'$  behind the lens. The point  $P$  is related to the point  $P'$  by the thin-lens formula

$$\frac{1}{P} + \frac{1}{P'} = \frac{1}{f}$$

where  $f$  is the focal length of the thin lens. The extent of the cone of rays is limited by the diameter of the thin lens,  $a$ . If the minimum resolvable dimension on the sensor is  $b$  we can determine a point  $P_N$  whose cone of rays will form a circle of diameter  $b$  on a sensor placed at position  $P'$ . Similarly we can determine a point  $P_F$  whose cone of rays will also form a circle of diameter  $b$  on the sensor at  $P'$ . Given the geometry of the situation the cone of rays from any point lying between  $P_N$  and  $P_F$  will fall completely within the circle of diameter  $b$  and thus be unresolvable from any other point in the interval. The unresolvable region between  $P_N$  and  $P_F$  is the depth of field for the camera for the focused distance  $P$ .

To determine the depth of field we use the similar triangles in Fig. D.8 to give us

$$\frac{a}{b} = \frac{P'_N}{P'_N - P'} = \frac{P'_F}{P' - P'_F}.$$

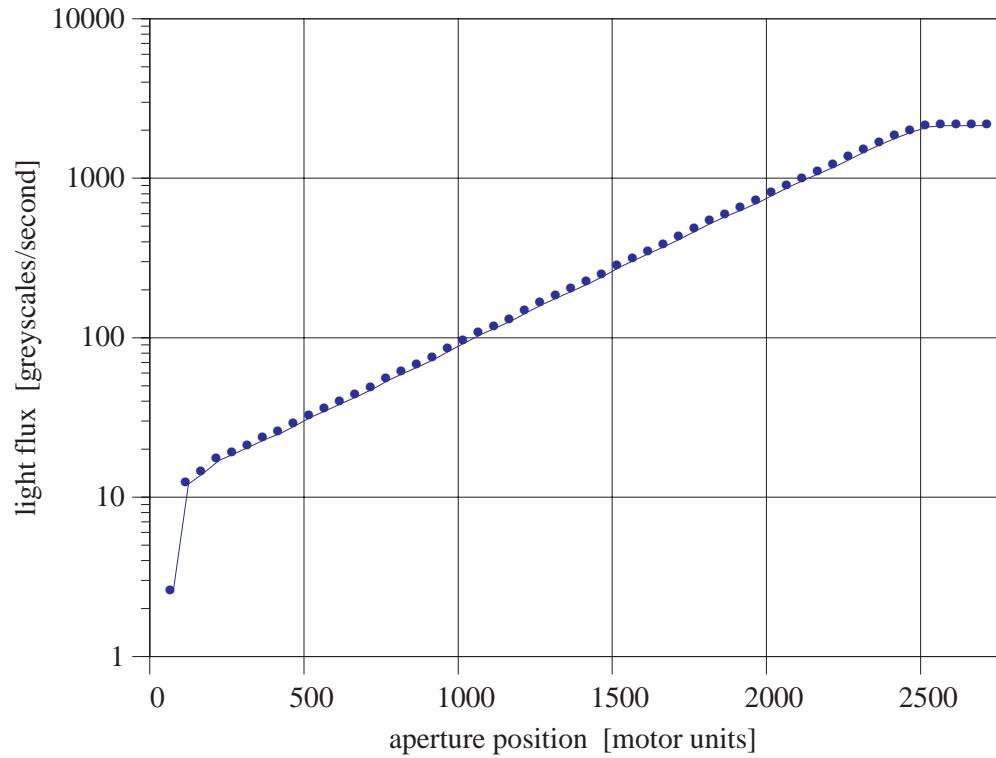


Figure D.4: Light flux (log scale) versus aperture motor

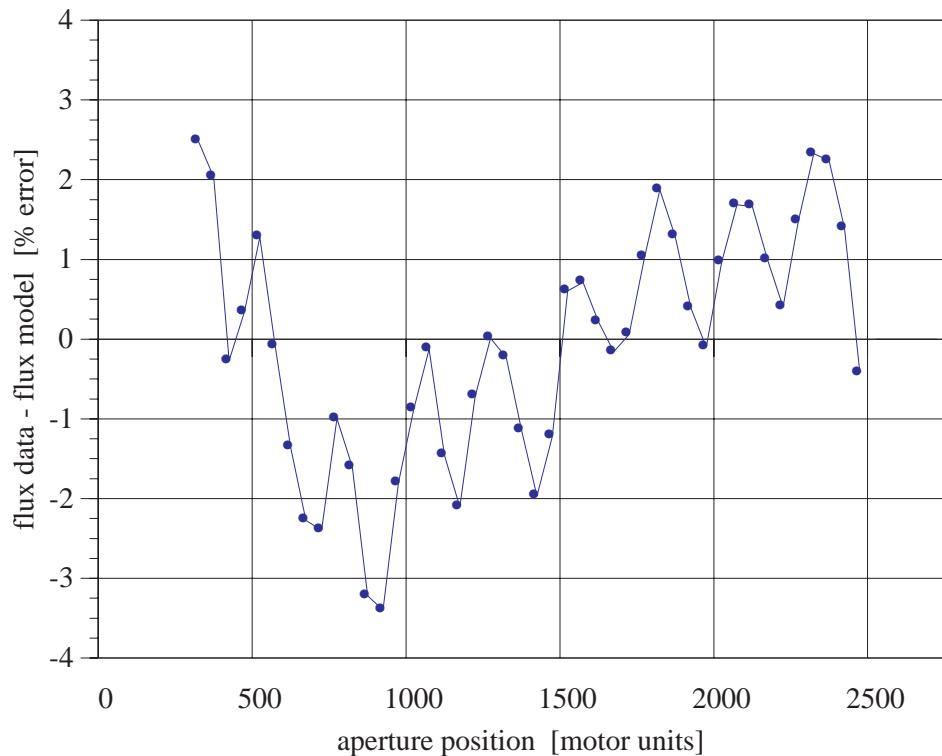


Figure D.5: Error between flux data and flux model versus aperture motor

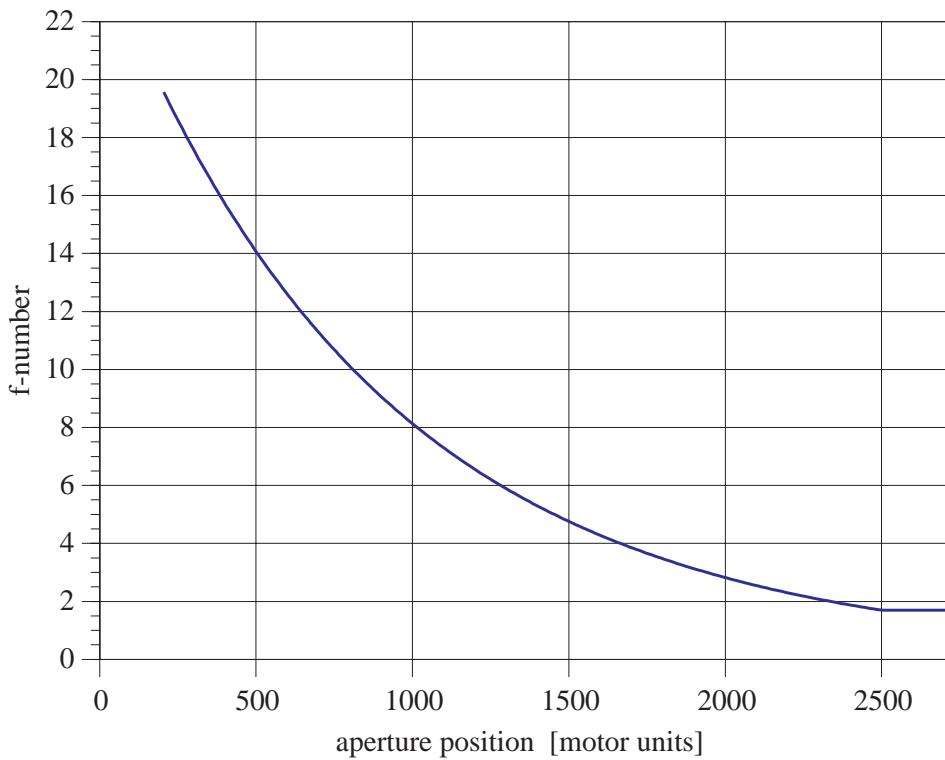
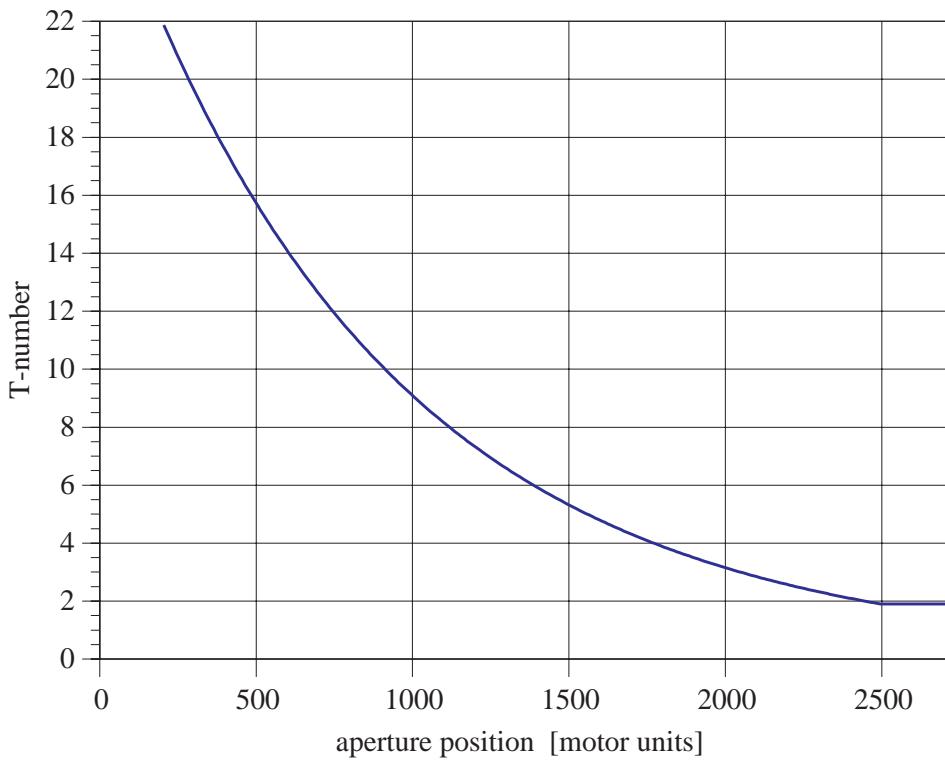
Figure D.6:  $f$ -number versus aperture motor

Figure D.7: T-number versus aperture motor

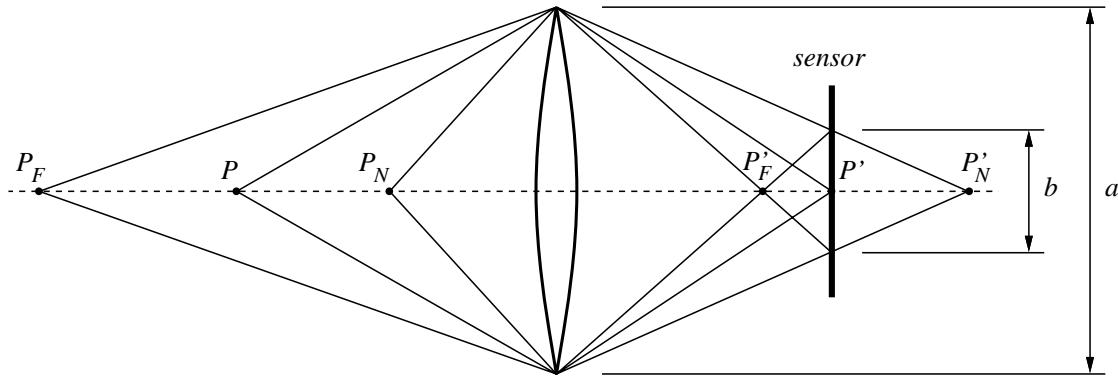


Figure D.8: Depth of field for a thin lens

Applying the thin-lens law to  $P'$ ,  $P'_N$ , and  $P'_F$  to obtain  $P$ ,  $P_N$ , and  $P_F$  and solving for  $P_N$  and  $P_F$  we have

$$P_N = \frac{afP}{af - bf + bP}$$

and

$$P_F = \frac{afP}{af + bf - bP}.$$

The extent of the depth of field is therefore

$$D = P_F - P_N = \frac{2abPf(P - f)}{a^2f^2 - b^2f^2 - b^2P^2 + 2b^2fP}.$$

To calculate the depth of field for the camera we need to know the values of  $f$ ,  $a$ ,  $b$ , and  $P$ . Values for  $b$  can be determined from the dimensions of the pixels in the camera's sensor. Values for  $P$  can be determined from a focused distance model and the range-from-focus techniques discussed in Chapter 2. The ratio  $f/a$  can be determined using the relative aperture model developed above. However, to obtain the value for either  $a$  or  $f$  that would be necessary to determine the actual depth of field it is likely that the lens would have to be removed from the camera and measured on an optical bench.

## D.4 Lens aberrations and aperture

The aperture of the lens has a direct influence on the radiometric imaging properties of the camera system as well as on the geometric imaging properties related to the camera's depth of field. However, the aperture also has more subtle but potentially significant effects on other imaging properties.

The quality of the image formed by a lens is determined by the contributions of the lens aberrations to the image. These contributions depend on the position of the aperture stop along the optical axis and on the aperture's radius. The relationships between the radius  $r$  of the aperture and the primary lens aberrations are shown in Table D.1. Changing the

Table D.1: Effect of aperture radius on primary lens aberrations (From [43])

primary aberration	relation to aperture radius
spherical (longitudinal)	$r^2$
spherical (transverse)	$r^3$
coma	$r^2$
length of astigmatic lines	$r$

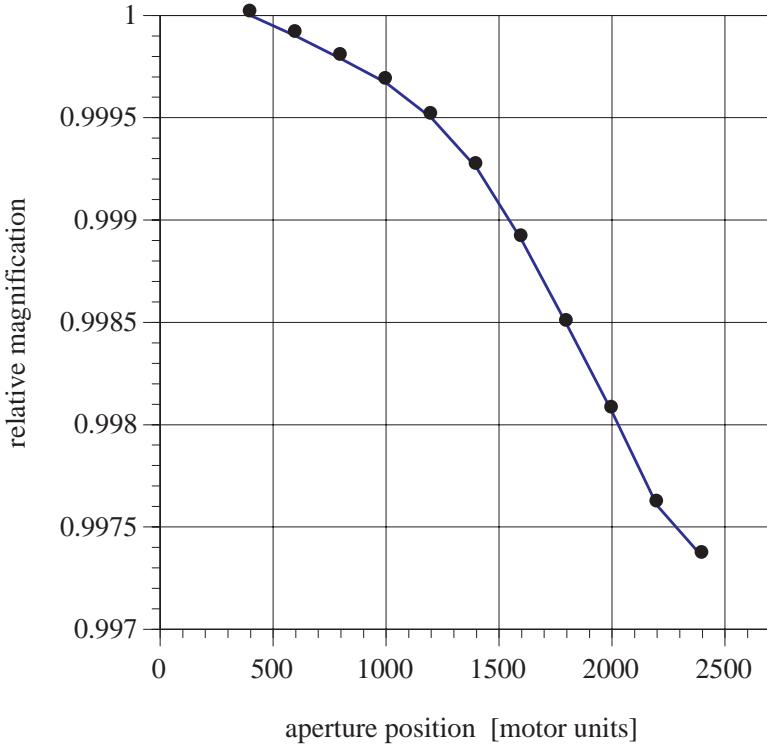


Figure D.9: Image magnification versus aperture motor

aperture changes the contributions of the lens aberrations, which in turn can alter image properties such as image magnification. Figure D.9 shows the change in image plane magnification for the Fujinon/Photometrics camera as a function of the aperture motor<sup>2</sup>. The subtle dependency between the lens aberrations and the aperture means that precise geometric camera models need to be calibrated across aperture setting, as we demonstrated in Section 6.6.

---

<sup>2</sup>For this experiment the relative magnification was measured using the target in Fig. 5.5 and the algorithm in Section 3.3.3.



# Bibliography

- [1] A. L. Abbott. *Dynamic Integration of Depth Cues for Surface Reconstruction From Stereo Images*. PhD thesis, University of Illinois at Urbana-Champaign, January 1991.
- [2] Y. I. Abdel-Aziz and H. M. Karara. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In *Proceedings of Symposium on Close-Range Photogrammetry*, pages 1–18, Urbana, IL, January 1971.
- [3] A. Bani-Hashemi. Finding the aspect-ratio of an imaging system. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 122–156, Maui, HI, June 1991.
- [4] H. A. Beyer. Linejitter and geometric calibration of CCD-cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, 45(1):17–32, 1990.
- [5] H. A. Beyer. Accurate calibration of CCD-cameras. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 96–101, Champaign, IL, June 1992.
- [6] H. A. Beyer. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System*. PhD thesis, Institute of Geodesy and Photogrammetry, Federal Institute of Technology (ETH) Zurich, Switzerland, May 1992.
- [7] T. E. Boult and G. Wolberg. Correcting chromatic aberrations using image warping. In *Proceedings of Image Understanding Workshop*, San Diego, CA, January 1992. DARPA.
- [8] G. E. P. Box. A useful method for model-building. *Technometrics*, 4(3):301–318, August 1962.
- [9] G. E. P. Box and W. G. Hunter. The experimental study of physical mechanisms. *Technometrics*, 7(1):23–42, February 1965.
- [10] R. Bracho, J. F. Schlag, and A. C. Sanderson. POPEYE: A grey-level vision system for robotics applications. Technical Report CMU-RI-TR-83-6, Carnegie-Mellon University, May 1983.
- [11] D. C. Brown. Decentering distortion of lenses. *Photometric Engineering*, 32(3):444–462, 1966.
- [12] D. C. Brown. Close-range camera calibration. *Photometric Engineering*, 37(8):855–866, 1971.

- [13] CANON. J55x information brochure.
- [14] G. Champleboux, S. Lavallee, P. Sautot, and P. Cinquin. Accurate calibration of cameras and range imaging sensors : the NPBS method. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1552–1557, Nice, France, May 1992.
- [15] J. S. Chen and G. Medioni. Detection, localization, and estimation of edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2):191–198, February 1989.
- [16] J. A. Cox. Evaluation of peak location algorithms with subpixel accuracy for mosaic focal planes. In *Proceedings of SPIE Conference on Processing of Images and Data from Optical Sensors*, volume 292, pages 288–299, 1981.
- [17] T. Darrell and K. Wohn. Pyramid based depth from focus. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 504–509, Ann Arbor, MI, June 1988.
- [18] S. Das and N. Ahuja. A comparative study of stereo, vergence, and focus as depth cues for active vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 194–199, New York, NY, June 1993.
- [19] B. Funt and J. Ho. Color from black and white. *International Journal of Computer Vision*, 3(2):109–117, June 1989.
- [20] D. B. Gennery. Camera calibration including lens distortion. Technical Report JPL D-8580, Jet Propulsion Laboratory, May 1991.
- [21] N. Goldberg. *Camera Technology: The Dark Side of the Lens*. Academic Press, San Diego, CA, 1992.
- [22] K. D. Greban, C. E. Thorpe, and T. Kanade. Geometric camera calibration using systems of linear equations. In *Proceedings of Image Understanding Workshop*, pages 820–825, Cambridge, MA, April 1988. Morgan Kaufmann.
- [23] B. K. P. Horn. Focusing. Memo 160, Massachusetts Institute of Technology, May 1968.
- [24] R. Kingslake. *Applied optics and optical engineering*, volume 2. Academic Press, New York, NY, 1965.
- [25] R. Kingslake. *Optical System Design*. Academic Press, New York, NY, 1983.
- [26] G. J. Klinker. *A Physical Approach to Color Image Understanding*. PhD thesis, Carnegie Mellon University, May 1988.
- [27] E. P. Krotkov. *Exploratory Visual Sensing for Determining Spatial Layout with an Agile Stereo Camera System*. PhD thesis, University of Pennsylvania, April 1987. Technical Report MS-CIS-87-29.

- [28] M. Laikin. *Lens Design*. Marcel Dekker, New York, NY, 1991.
- [29] J.-M. Lavest, G. Rives, and M. Dhome. Three-dimensional reconstruction by zooming. *IEEE Journal of Robotics and Automation*, 9(2):196–207, April 1993.
- [30] R. K. Lenz and R. Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 68–75, Raleigh, NC, March 1987.
- [31] I. Maalen-Johansen. On the precision of subpixel measurements in videometry. In *Proceedings of Second Conference on Optical 3-D Measurement Techniques*, pages 169–178, Switzerland, October 1993. Institute of Geodesy and Photogrammetry, Federal Institute of Technology (ETH) Zurich.
- [32] R. M. Malone, B. C. Frogget, D. A. Grafton, and F. R. Mitchell. Design of a twenty-element long focal length zoom lens. In *Proceedings of SPIE Conference on Current Developments in Optical Engineering IV*, volume 1334, pages 43–48, San Diego, CA, July 1990.
- [33] H. N. Nair and C. V. Stewart. Depth from focus. Technical Report 90-10, Rensselaer Polytechnic Institute, 1990.
- [34] H. N. Nair and C. V. Stewart. Robust focus ranging. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 309–314, Champaign, IL, June 1992.
- [35] S. K. Nayar and Y. Nakagawa. Shape from focus: An effective approach for rough surfaces. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 218–225, Cincinnati, OH, May 1990.
- [36] C. L. Novak. *Estimating Scene Properties by Analyzing Color Histograms with Physics-Based Models*. PhD thesis, Carnegie Mellon University, December 1992.
- [37] K. Pahlavan. *Active Robot Vision and Primary Ocular Processes*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, May 1993.
- [38] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, New York, NY, second edition, 1992.
- [39] J. Rissanen. Complexity of models. Technical Report RJ 6981 (66518), IBM, August 1989.
- [40] M. V. Rodin. *Stereovision Axiale: Calibrage, Appariement et Reconstruction Tridimensionnelle (in French)*. PhD thesis, Institut National Polytechnique de Toulouse, October 1993.
- [41] J. F. Schlag, A. C. Sanderson, C. P. Neuman, and F. C. Wimberly. Implementation of automatic focusing algorithms for a computer vision system with camera control. Technical Report CMU-RI-TR-83-14, Carnegie-Mellon University, August 1983.

- [42] C. C. Slama, C. Theurer, and S. W. Henriksen. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Church, VA, fourth edition, 1980.
- [43] W. J. Smith. *Modern Optical Engineering, The Design of Optical Systems*. Optical and Electro-Optical Engineering Series. McGraw-Hill, New York, NY, 1966.
- [44] I. E. Sobel. *Camera Models and Machine Perception*. PhD thesis, Stanford, May 1970. Memo AIM-121.
- [45] I. E. Sobel. On calibrating computer controlled cameras for perceiving 3-D scenes. *Artificial Intelligence*, 5(4):185–198, 1974.
- [46] G. Surya and M. Subbarao. Depth from defocus by changing camera aperture: A spatial domain approach. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 61–67, New York, NY, June 1993.
- [47] E. C. Tam. Smart electro-optical zoom lens. *Optics Letters*, 17(5):369–371, March 1992.
- [48] K. A. Tarabanis. *Sensor Planning and Modeling for Machine Vision Tasks*. PhD thesis, Columbia University, 1991.
- [49] K. A. Tarabanis, R. Y. Tsai, and D. S. Goodman. Modeling of a computer-controlled zoom lens. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1545–1551, Nice, France, May 1992.
- [50] A. Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam, Netherlands, 1987.
- [51] J. M. Tenenbaum. *Accommodation in Computer Vision*. PhD thesis, Stanford, September 1970. Memo AIM-134.
- [52] R. Y. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 364–374, Miami Beach, FL, 1986.
- [53] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, August 1987.
- [54] L.-L. Wang and W.-H. Tsai. Computing camera parameters using vanishing-line information from a rectangular parallelepiped. *Machine Vision and Applications*, 3(3):129–141, 1990.
- [55] A. G. Wiley. *Metric Aspects of Zoom Vision*. PhD thesis, University of Illinois at Urbana-Champaign, June 1991.
- [56] R. G. Willson and S. A. Shafer. Precision imaging and control for machine vision research at Carnegie Mellon University. In *Proceedings of SPIE Conference on High-Resolution Sensors and Hybrid Systems*, volume 1656, pages 297–314, San Jose, CA, February 1992.

- [57] G. Wolberg. *Digital Image Warping*. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [58] Y. Yakimovsky and R. Cunningham. A system for extracting three-dimensional measurements from a stereo pair of TV cameras. *Computer Graphics and Image Processing*, 7:195–210, 1978.