Modeling and Calibration of Automated Zoom Lenses

Reg G. Willson

3M Engineering Systems and Technology
3M Center, Building 518-1-01, Saint Paul, MN 55144-1000 USA
E-mail: rgwillson@mmm.com

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 lens settings 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.

In this paper we present a methodology for producing accurate camera models for systems with automated, variable-parameter lenses. To demonstrate our methodology's effectiveness we applied it to produce an "adjustable," perspective-projection camera model based on Tsai's fixed camera model. Our model was calibrated and tested on an automated zoom lens where it operated across continuous ranges of focus and zoom with an average error of less than 0.11 pixels between the predicted and measured positions of features in the image plane.

Keywords: camera modeling, camera calibration, zoom lens, computer vision

1 FIXED VERSUS ADJUSTABLE CAMERA MODELS

Conventionally camera models have been used to capture the imaging properties of fixed-parameter lenses. For fixed lenses the image-formation process is static, and thus the camera model's terms are constants. In variable-parameter lenses the image-formation process is an adjustable function of the lens control settings, and thus the terms in the camera models must also be variable. The question is, "How do the terms in the camera models vary with the lens settings?" 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 model terms and the lens's physical configuration is not direct. Second, the relationship between the lens's physical configuration and the control settings is complex and typically we have very little a priori knowledge about the underlying mechanisms involved. We have no good theoretical basis for these relationships. Since every model term is potentially a function of every lens control, the actual relationships between the model terms and the lens controls must be determined empirically.

Unlike the calibration of fixed-parameter lenses, the calibration of variable-parameter lenses requires measurements over ranges of lens settings. This raises several challenges. First, the dimensionality of the data is the same as the number of controls that are to be concurrently modeled. Even if we just took 10 measurements across the ranges of focus, zoom, and aperture controls, 1000 settings 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. As the lens is zoomed in (i.e. the focal length is increased) the number of features 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 may become too small and/or crowded to be accurately measured. As a result, different calibration setups may be required to cover the full range of zoom. Similar problems can occur with the focus and aperture controls.

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

1. Collection of calibration data for the fixed camera model across ranges of lens settings.
2. Calibration of the fixed camera model at each measured lens setting.
3. Characterization of the relationships between the fixed camera model's parameters and the lens settings.

The equations for the fixed camera model plus the calibrated parameter models constitute an adjustable camera model.

1.1 Collecting calibration data

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 where m_f is the focus motor setting). 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 settings to its imaging 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 lens's control space to find approximate limits.

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 lens control so that the underlying variations in the model parameters can be accurately characterized. Since we start with little or no a priori information about the relationships between the lens controls and the camera model parameters the sampling strategy must be determined empirically.

1.2 Characterizing variations in the fixed model's parameters

If we take the parameter values from the calibrated fixed camera models and just store them in lookup tables then we need to make no assumptions about how they vary with the lens settings. However, if we want to use a more compact, algebraic form for the parameter values or interpolate between the sampled lens settings, we must determine expressions for the individual parameter models. For our camera systems we find that simple polynomials work well. We choose the polynomial orders based on design objectives for the final adjustable model and on an examination of the data.

Having chosen the form of the parameter models we next need to fit them to the their respective data. Instead of fitting all of the parameter models independently and in one step, we work with the parameters one at a time. In our approach we fit one polynomial model to the data for one parameter, 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. For a given set of parameter models the final level of error generally depends on the sequence in which the models are fit. We 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 ascending-polynomial-order, greedy-within-order sequencing.

After the all parameter models have been fit we cycle through the parameters again, 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 error between the adjustable camera model and the calibration data.

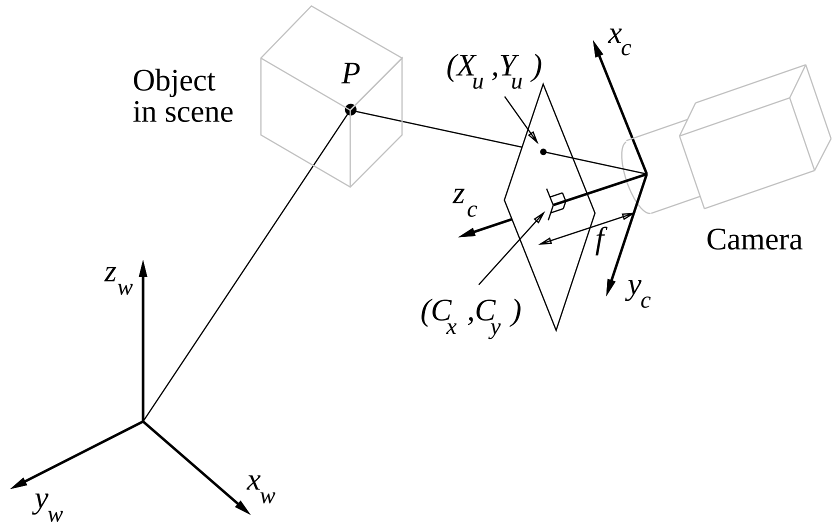

Figure 1: Fixed perspective-projection camera model geometry

2 TSAI'S FIXED PERSPECTIVE-PROJECTION CAMERA MODEL

The basis for our adjustable camera model is the 3D to 2D perspective-projection model described by Tsai [2].

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. For a fixed lens 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.

In Tsai's model, illustrated in Fig. 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 pinhole camera's effective focal length.

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

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{1}$$

where

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \tag{2}$$

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

$$R = \mathrm{Rot}(R_x)\mathrm{Rot}(R_y)\mathrm{Rot}(R_z) \tag{3}$$

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

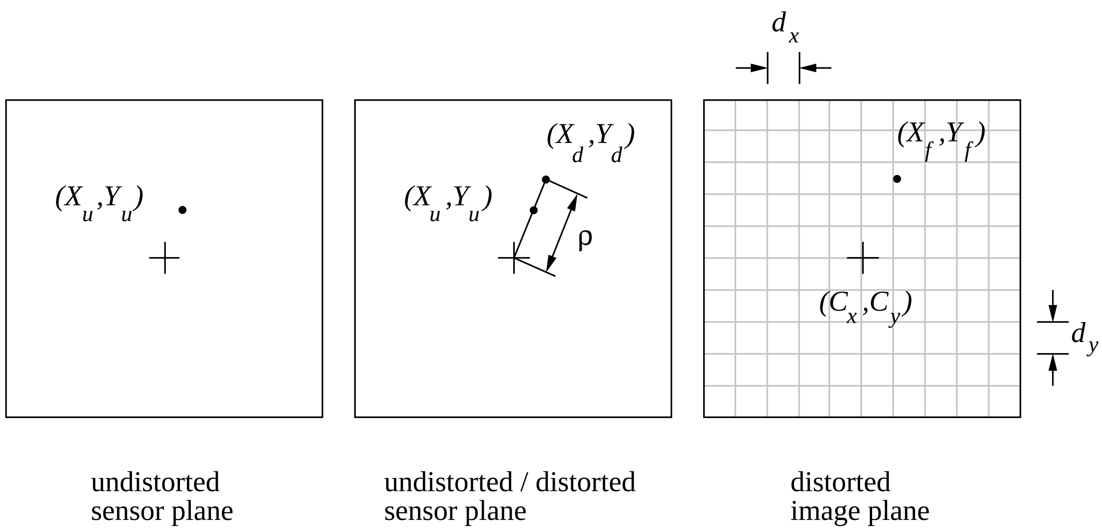

Figure 2: Transformation from undistorted sensor to distorted frame coordinates

coordinates to the position of its image in undistorted sensor-plane coordinates, (X_u, Y_u). This transformation is described by

$$X_u = f\frac{x_c}{z_c} \tag{4}$$

and

$$Y_u = f\frac{y_c}{z_c} \tag{5}$$

The third transformation, illustrated in Fig. 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

$$X_u = X_d(1 + \kappa_1\rho^2), \tag{6}$$
$$Y_u = Y_d(1 + \kappa_1\rho^2) \tag{7}$$

and

$$\rho = \sqrt{X_d^2 + Y_d^2} \tag{8}$$

where κ_1 is the coefficient of radial lens distortion.

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

$$X_f = d_x^{-1} X_d s_x + C_x \tag{9}$$

and

$$Y_f = d_y^{-1} Y_d + C_y \tag{10}$$

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.

2.1 Fixed camera model performance metrics

One of the first questions we have about any camera model is how accurately it captures the camera's 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

$$\text{UIPE} = \sqrt{(\Delta X_{f_u})^2 + (\Delta Y_{f_u})^2} \tag{11}$$

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 (6), (7), and (8), while (X_{u_1}, Y_{u_1}) are calculated from the 3D coordinates of the point (x_w, y_w, z_w) using (1), (4), and (5). The algorithms that we use to calibrate the camera model minimize the sum-of-squared error in the undistorted image plane for the calibration data.

Our estimation of the unknown parameters in Tsai's 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 paper we used a planar calibration target mounted on a translation stage. The calibration target contained 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 was known from their 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 reference points in the image plane were measured to sub-pixel accuracy using the procedure described in [3].

When calibrating our fixed camera models we assumed that the six exterior orientation parameters $(R_x, R_y, R_z, T_x, T_y, T_z)$ and five interior orientation parameters $(f, \kappa_1, C_x, C_y, s_x)$ were all unknown and had to be estimated from the calibration data. We calibrated the fixed camera model in two steps. First we used Tsai's algorithm to obtain approximate estimates for nine of the model's 11 parameters and then we used iterative, non-linear optimization to refine all 11 parameters.

2.2 Fixed camera model calibration example

To demonstrate the calibration of a fixed camera model we calibrated our camera system for one lens setting. The calibration data for the model came from two images of the calibration target taken with sensor-to-target ranges of 1.5 m

Parameter	Value	Units
f	60.013	mm
C_x	267.198	pixels
C_y	255.040	pixels
κ_1	-0.000103	1/mm^2
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

Table 1: Example of a calibrated fixed camera model

and 2.5 m. 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 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.

3 AN ADJUSTABLE PERSPECTIVE-PROJECTION CAMERA MODEL

Before we proceed in developing our adjustable camera model we introduce the following notation.

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

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

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.

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

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

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

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

$$g_P(S) = polynomial(S; a_0, \ldots, a_n)$$

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.

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

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.

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

3.1 Adjustable camera model performance metrics

Our objective has been to develop a model of the camera's imaging behavior that "holds calibration" across ranges of lens settings. By "holds calibration" we mean that the model maintains an acceptable level of accuracy at any setting. 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" at the settings used for the 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 earlier. To be able to measure progress during calibration and 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 lens 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,

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

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,

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

For the following adjustable model 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.

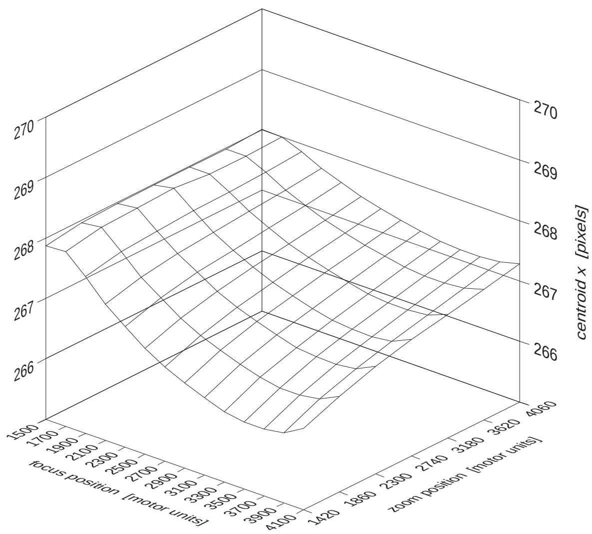

Figure 3: Variation in X coordinate of autocollimated laser's image with focus and zoom motors

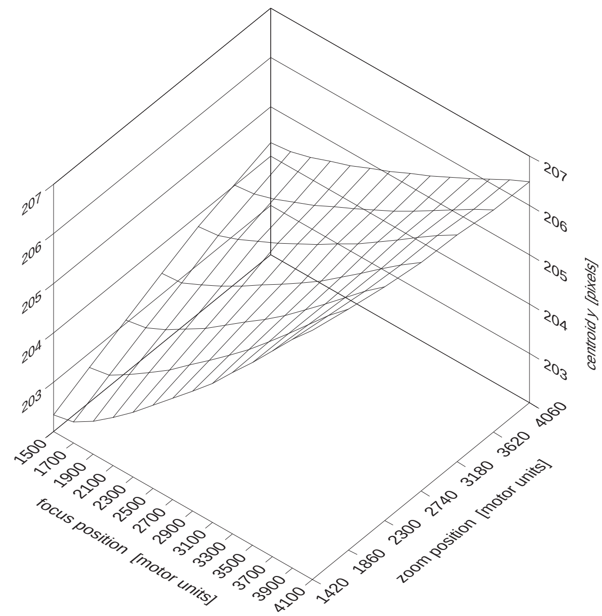

Figure 4: Variation in Y coordinate of autocollimated laser's image with focus and zoom motors

3.2 Example - an adjustable camera model for focus and zoom

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 $f/16$.

Figures 3 and 4 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 11×11 sampling of focus and zoom settings for a total of 121 separate settings $(S_1, \ldots, S_{121})$ across the operating space for our camera model. Had the plots revealed discontinuities in the lens's imaging properties different sampling strategies and parameter model formulations would have to be used.

Calibration data for the adjustable camera model was obtained using the target and translation stage described earlier. 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 the 121 different lens settings $(S_1, \ldots, S_{121})$ we obtained 121 sets of calibration data $(D_1, \ldots, D_{121})$. Each set contained between 110 and 429 calibration data points.

For each set of data $(D_1, \ldots, D_{121})$ we calibrated fixed camera models $(M_{f_1}, \ldots, M_{f_{121}})$. Figures 5 through 16 show the values for the 11 fixed model parameters and the M_UIPE plotted against the focus and zoom settings. Despite the apparent noise in many of the model parameters, the M_UIPE for the individual fixed camera models 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.

For the 11 parameter models we used bivariate polynomial functions with the same model order for each independent variable. The largest bivariate polynomial that can be fit to the 121 data points is 14th order (120 coefficients).

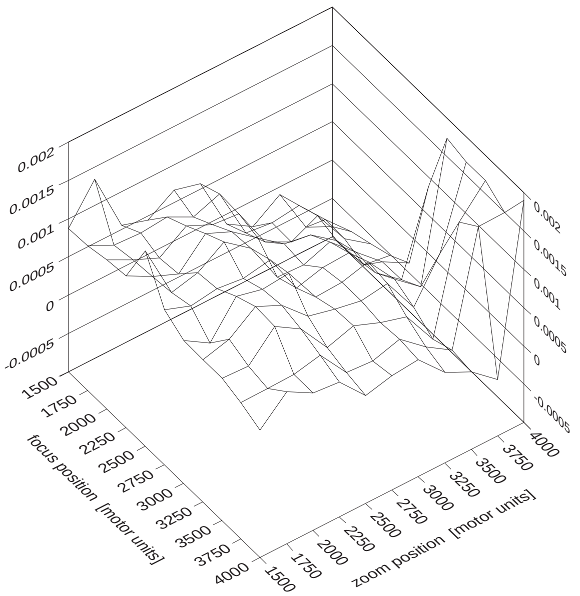

Figure 5: Fixed camera model R_x versus focus and zoom motors

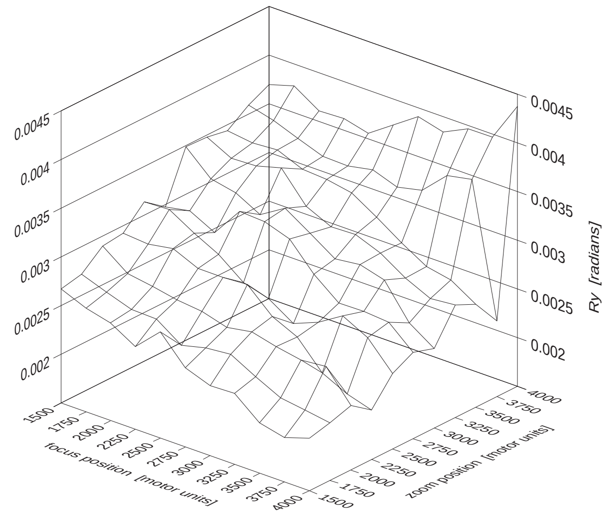

Figure 6: Fixed camera model R_y versus focus and zoom motors

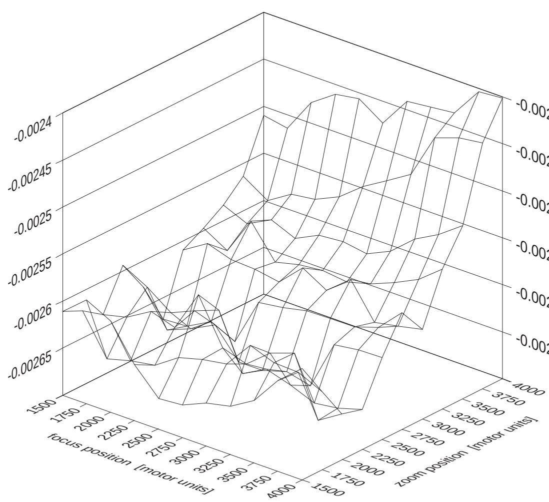

Figure 7: Fixed camera model R_z versus focus and zoom motors

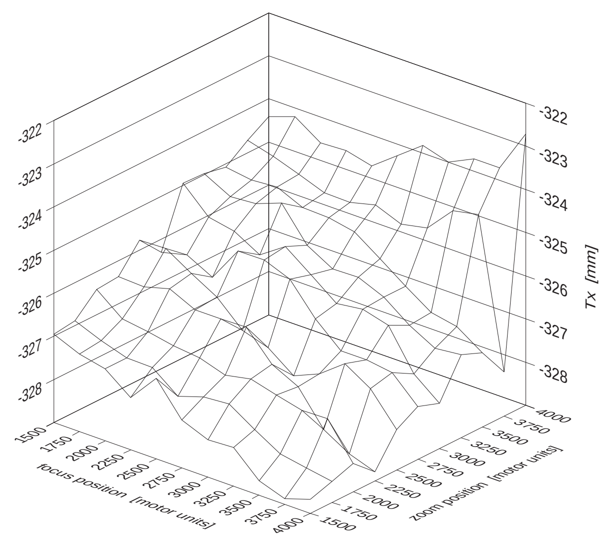

Figure 8: Fixed camera model T_x versus focus and zoom motors

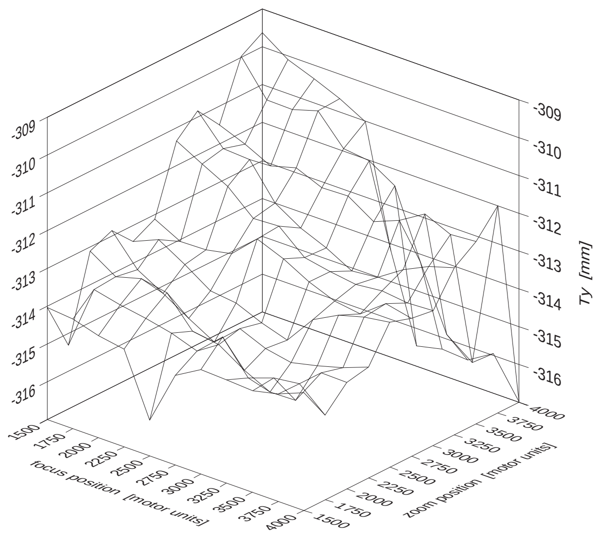

Figure 9: Fixed camera model T_y versus focus and zoom motors

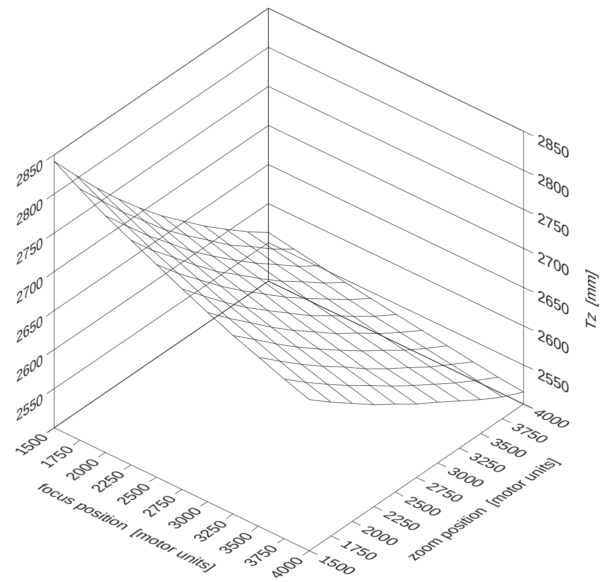

Figure 10: Fixed camera model T_z versus focus and zoom motors

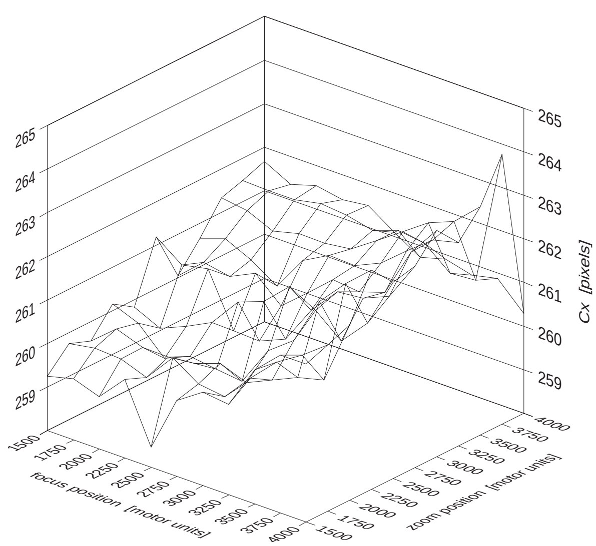

Figure 11: Fixed camera model C_x versus focus and zoom motors

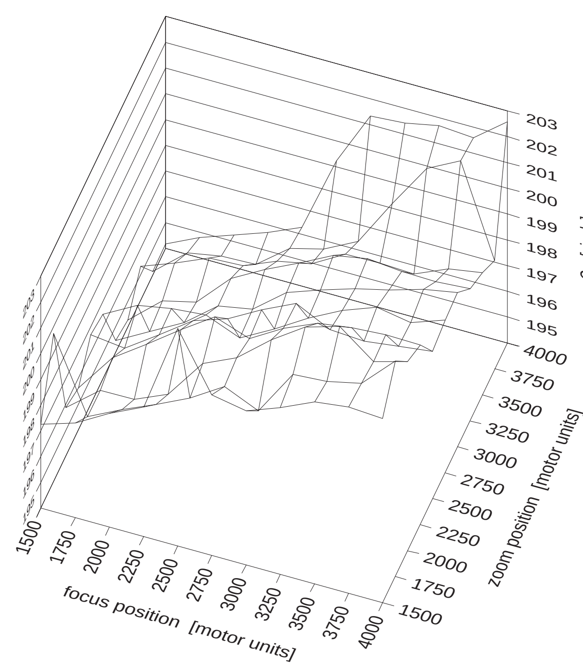

Figure 12: Fixed camera model C_y versus focus and zoom motors

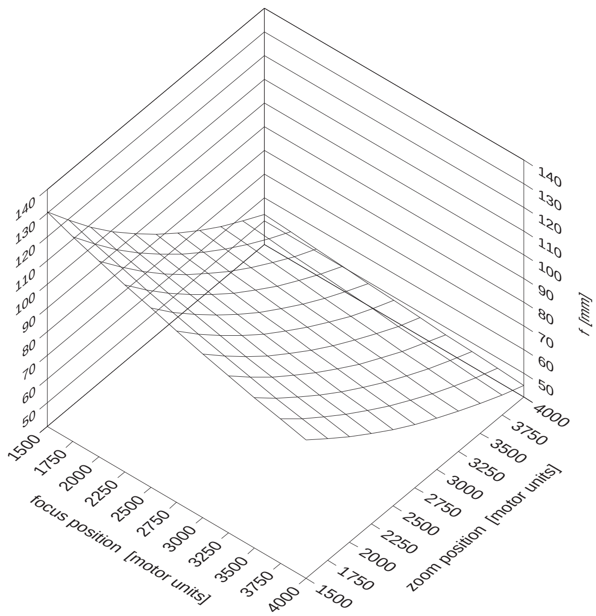

Figure 13: Fixed camera model f versus focus and zoom motors

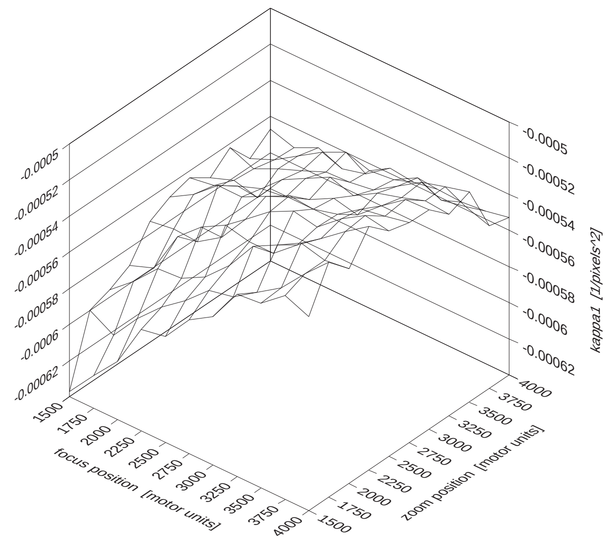

Figure 14: Fixed camera model κ_1 versus focus and zoom motors

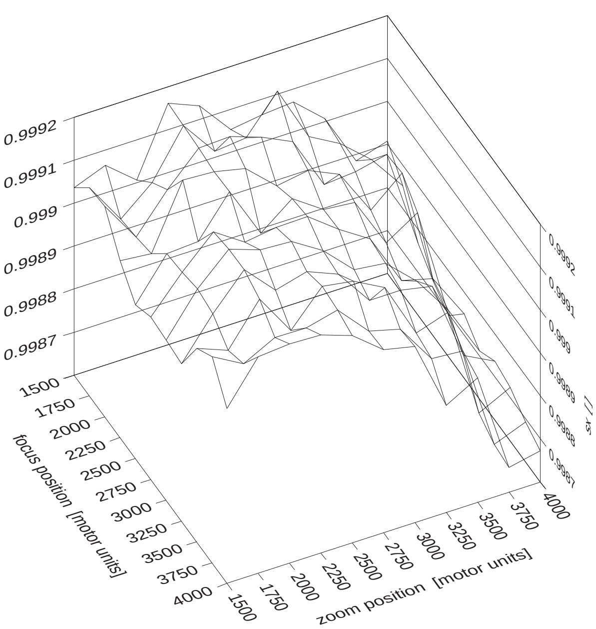

Figure 15: Fixed camera model s_x versus focus and zoom motors

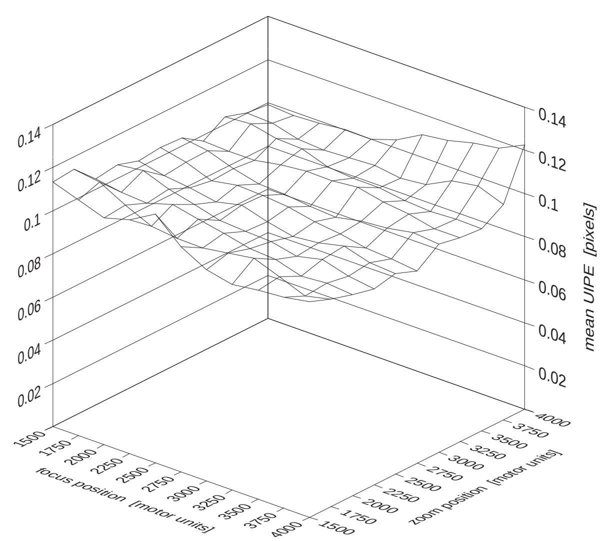

Figure 16: Fixed camera model M_UIPE versus focus and zoom motors

To fit the polynomial functions to the parameter values we used least-squared-error fitting. 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_i the fixed camera model 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 fixed model parameter values 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-squared-error fitting is not a maximum likelihood estimator for the models. Even so, since the least-squared-error fitting can be accomplished with a direct, non-iterative approach this is our preferred fitting method. Tests using much slower but more robust fitting techniques using local M-Estimates [1] showed no significant improvement in the performance of the final adjustable camera model in this example.

The second implication of the non-Gaussian noise is that we cannot use a Chi-Squared test to determine how high a polynomial order to use in each parameter model. Instead we chose the model order based on design requirements for the adjustable camera model and on an examination of the data. To partially decouple the intrinsic from the extrinsic parameters and make the 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 [3] we discuss alternate strategies for choosing parameter model orders.

Table 2 summarizes the parameters, the orders chosen for their parameter models, and the rationale for the choice of order. The final adjustable camera model required a total of 96 coefficients for the 11 parameter models.

To fit the parameter models to the calibration data we used ascending-polynomial-order, greedy-within-order sequencing. Table 3 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 initial fixed camera models. Steps 1 to 11 are for the individual parameter model fits. Steps 12 and 13 are for iterative refinement. The last entry is for the final adjustable camera model.

Figures 17 through 21 show the final adjustable camera model surfaces for the parameters having second- and fifth-order polynomials (the zero-order models are constants). While the final f, T_z, and κ_1 models are all similar in shape to their original unfitted parameters, the remaining models are all rather different than their original data.

Figure 22 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 individual fixed camera models. Figure 23 shows the difference between the M_UIPE for the final adjustable camera model and the M_UIPE for the individual 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 individual fixed camera models.

If we were to calibrate this 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 κ_1 surfaces for both lenses would be similar, as would the positions of any image property discontinuities in the lens's control space.

3.3 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[1]. 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

[1] The adjustable camera model can only be guaranteed to be accurate over the range of distances and camera parameters that the calibration data covered.

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.
κ_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 2: Choice of polynomial orders for the parameter models

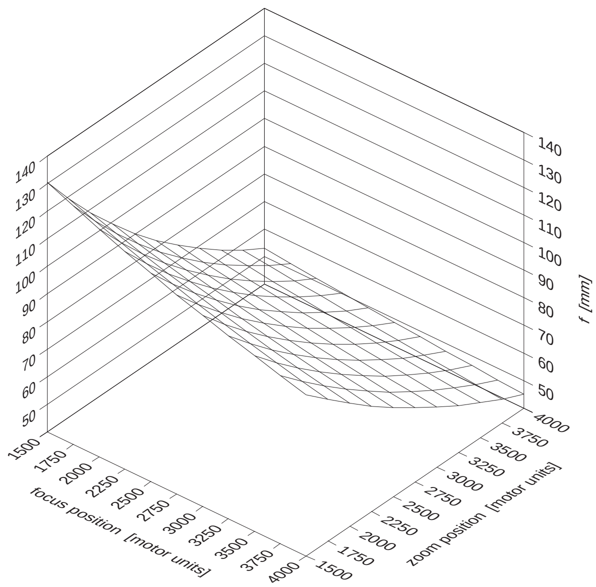

Figure 17: Adjustable camera model f versus focus and zoom motors

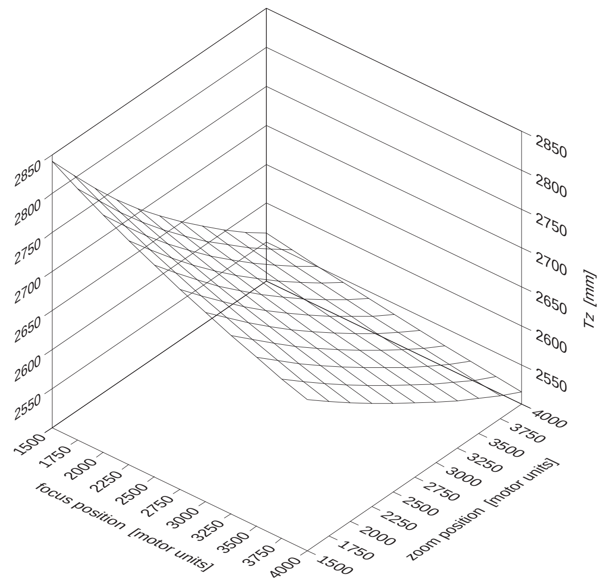

Figure 18: Adjustable camera model T_z versus focus and zoom motors

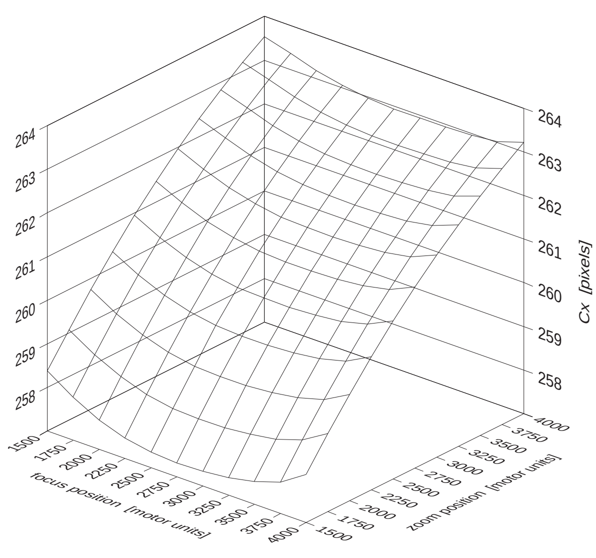

Figure 19: Adjustable camera model C_x versus focus and zoom motors

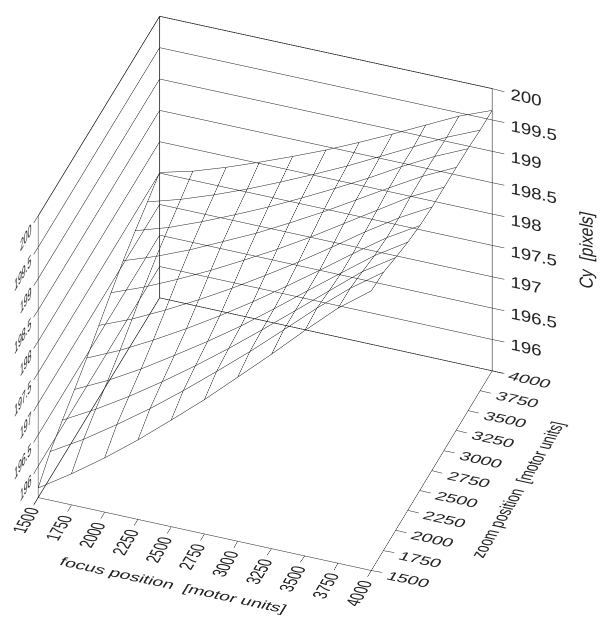

Figure 20: Adjustable camera model C_y versus focus and zoom motors

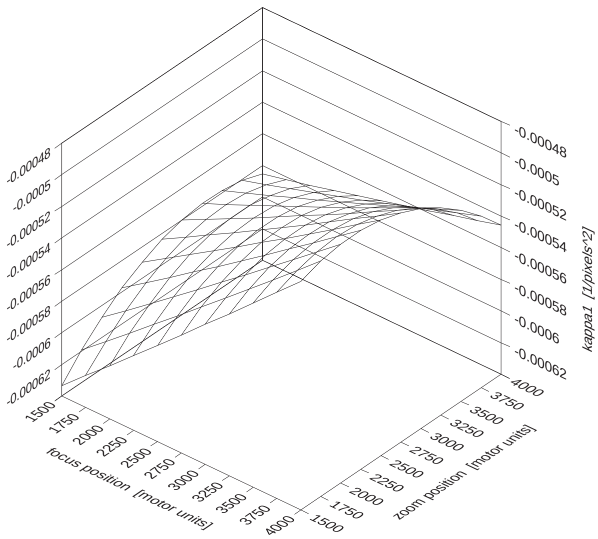

Figure 21: Adjustable camera model κ_1 versus focus and zoom motors

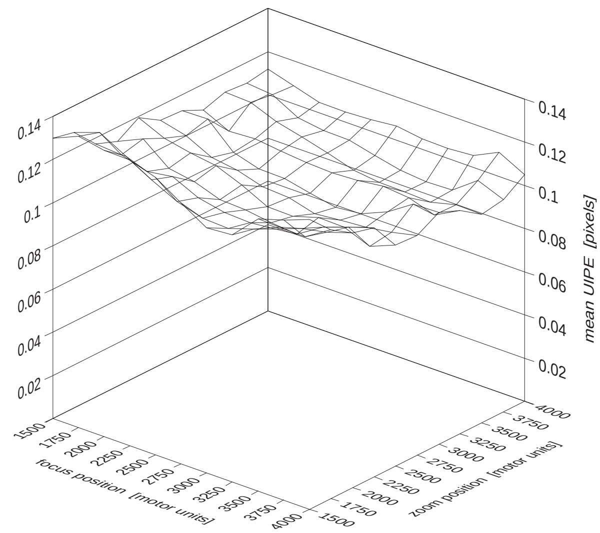

Figure 22: Adjustable camera model M_UIPE versus focus and zoom motors

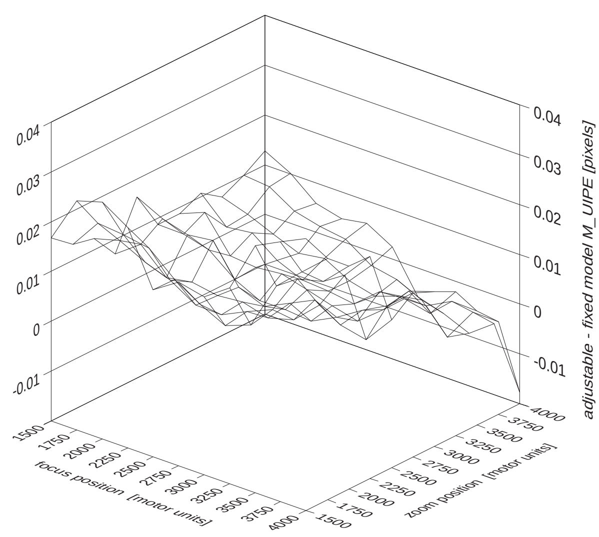

Figure 23: Difference between the final adjustable model M_UIPE and the initial fixed model M_UIPE

Fitting Step	Parameter	Polynomial Order	MM_UIPE [pixels]	max UIPE [pixels]	SSS_UIPE [pixels^2]
individual 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	κ_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 3: Fitting sequence for parameter models

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, ΔT_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, ΔT_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 24 illustrates these relationships in the 2D xz camera coordinate plane.

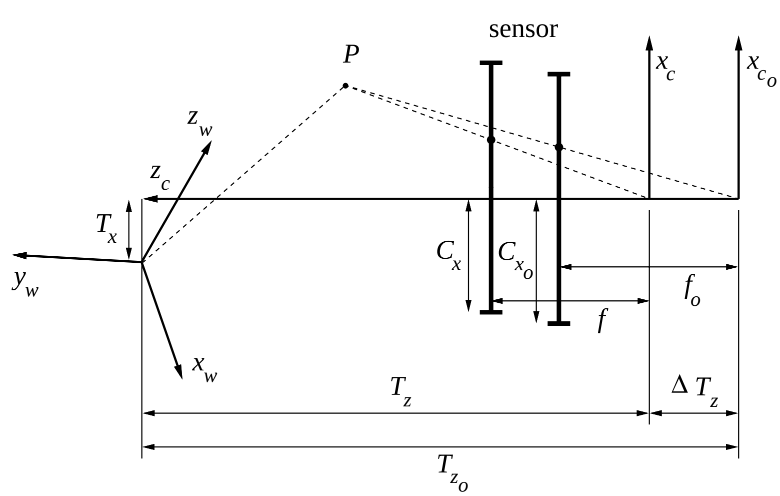

Figure 24: Extrinsic (and intrinsic) parameter changes with lens settings

4 SUMMARY

In this paper we have presented a methodology for empirically building camera models for systems with variable-parameter lenses. The methodology involves first calibrating a conventional fixed camera model at a number of settings spanning the desired range of lens settings for the adjustable model. We then characterize how the parameters of the fixed model vary with 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 settings. The resulting adjustable 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.

This approach makes no a priori assumptions about the dependencies between the fixed camera model parameters and the lens settings. It is general and can be applied to produce an adjustable camera model from any fixed one and allows any number of lens controls to be incorporated. The degree of accuracy and complexity, and consequently the required calibration effort, can be chosen arbitrarily.

5 ACKNOWLEDGMENTS

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. Additional support for publication of this paper has been provided by 3M.

REFERENCES

[1] 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.

[2] 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.

[3] R. G. Willson. *Modeling and Calibration of Automated Zoom Lenses.* PhD thesis, Carnegie Mellon University, January 1994.