What is the Center of the Image?

Reg G. Willson and Steven A. Shafer

April 1993
CMU-CS-93-122

School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

To appear in shortened form in the *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 15-17, 1993, New York, NY.

This research was sponsored by the Avionics Laboratory, Wright Research and Development Center, Aeronautical Systems Division (AFSC), U.S. Air Force, Wright-Patterson AFB, OH 45433-6543 under Contract F33615-90-C-1465, ARPA Order No. 7597. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of DARPA or the U.S. Government.

Keywords:

image center, camera calibration, automated zoom lens, computer vision, Calibrated Imaging Lab.

Contents

List of Figures

List of Tables

Abstract

People need to calibrate camera systems in order to determine the relationship between the positions of features in object space and their corresponding positions in the image. Part of camera calibration is the determination of image center. But, what is the image center? Ideally, the image center is considered to be the point of intersection of the camera's optical axis with the camera's sensing plane. In fact there are many possible definitions of image center, and in real lenses most do not have the same coordinates. In addition, the image centers move as lens parameters are changed. In this paper we examine why image centers are not necessarily the same for different image properties and why they vary with lens parameters. We then provide a taxonomy of 16 different image centers and describe procedures for measuring them. Finally we examine the calibration of image center for a variable parameter lens. Several techniques are applied to a precision automated zoom lens and experimental results are shown. We conclude that the accuracy of the image center can be an important factor in the accuracy of the overall camera calibration, and that the large variation in the position of the image center across different definitions and different lens settings makes the calibration problem much more complex than is conventionally believed. With proper modeling, by using correct definitions for all image centers in a system, we can improve the accuracy of our camera calibration.

1 Camera Calibration and Image Center

Camera calibration involves modeling the relationship between the positions of features in 3D object space and their corresponding positions in the 2D image. Precise camera models have many terms that describe the properties of the imaging process, and some of these terms are used to account for properties that vary with 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 that we use.

In an ideal lens there would be one image center which could be used in modeling any of the radially varying imaging properties. In practice the manufacturing tolerances for lenses result in different imaging properties having centers in different places, as shown in Figure 1 for our camera and lens. So, 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 2 shows how the position of a fixed point at the center of our camera's field of view shifts as a function of the focus and zoom motors of the camera lens.

We start this paper 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 16 different image centers are given and experimental results are then presented for ten of the methods. We conclude by examining how image center and field of view move in a variable parameter zoom lens.

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.

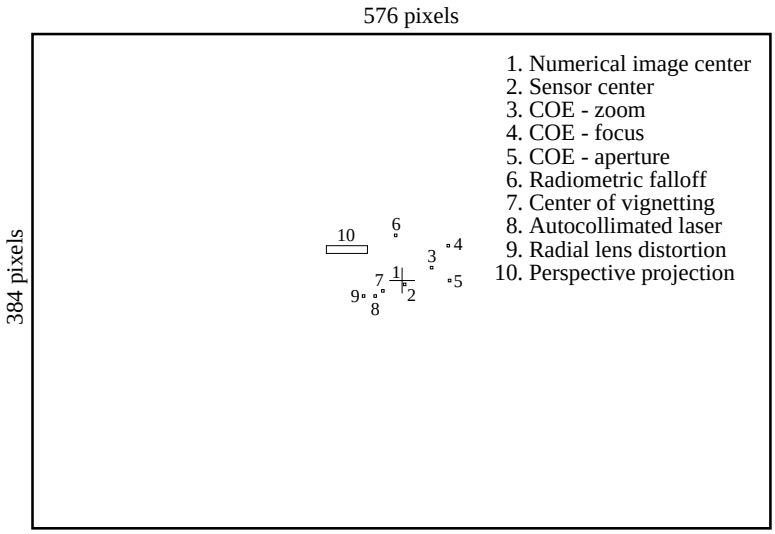

Figure 1: Different image centers for the same camera system.

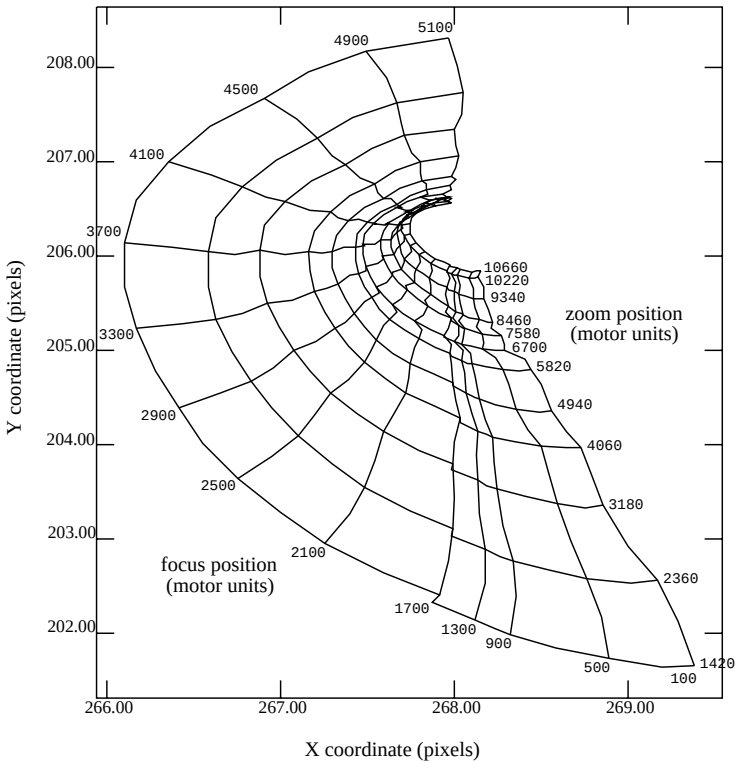

Figure 2: Shift in image center as a function of focus and zoom motors.

2 Real Lenses

Traditionally a camera's image center is considered to be the point of intersection of the lens' 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 practice the optical axis is not so easily defined for real lenses. 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 lens system the optical axis is defined as the straight line passing through all of the radii of curvature of the lens elements. The rotational symmetry of the system naturally leads to imaging properties that are radially symmetric around the optical axis. In a real lens system things are not so simple. For a simple lens element like that shown in Figure 3 there are actually two axes of symmetry, one optical and one mechanical. The optical axis of the lens is defined as the straight line joining the centers of curvature of the two surfaces of the lens. The mechanical axis of the lens is determined during manufacture by the centerline of the machine used to grind the lens' edge. Ideally the optical and mechanical axes would coincide. Practically though they won't. The tolerance between them is called *decentration* [6].

In a compound lens two or more lens elements are aligned and mounted together to form the complete lens, as illustrated in Figure 4. Ideally all of the elements would be aligned along a common optical axis, but this is not always possible 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. In fact, the decentration and misalignment can produce tangential lens distortion and asymmetric radial lens distortion [1]. As a result, the different imaging properties of the lens will not necessarily have a common axis of symmetry.

With adjustable lenses the focus and zoom are changed by varying the positions of the lens elements within the lens body. 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 then be rotated around the lens body to move the group along the axis of the lens. In the second method the lens elements are mounted on slides or rails which can then be translated along the axis of the lens body using internal cams. As the spacing of the lens elements is changed so is any misalignment between their mechanical and optical axes. Generally the rotation of a lens group will cause a rotational drift in the position of the lens' optical axis [4], while the sliding of a lens group will cause a translational motion of the lens' 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 variable focus fixed focal length lenses typically all of the lens elements are mounted together in a single fixed assembly. To vary the lens' focus the separation between the optics and the camera sensor is changed by moving the lens assembly with either a rotational or translational type mechanism. A less common focusing method found in some newer 35 mm autofocus lens designs involves the movement of a small lightweight element within the lens' optics to vary the focus of the image [2].

In variable focal length (zoom) lenses, such as the one illustrated in Figure 5, the focal length is changed by moving groups of lens elements relative to one another. Typically this is done by using a translational type mechanism on one or more of the internal groups. The lens' focus is often varied by using a rotational mechanism on the front lens group.

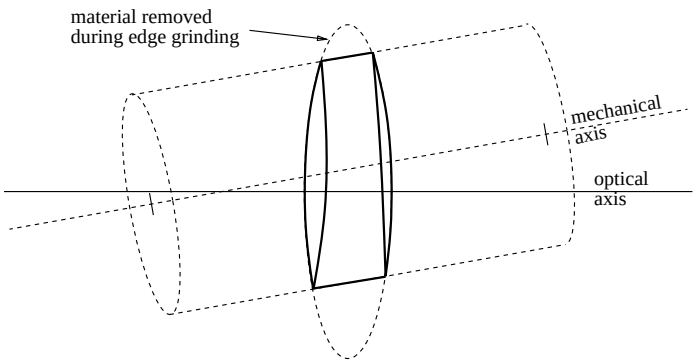

Figure 3: Decentration for a simple lens.

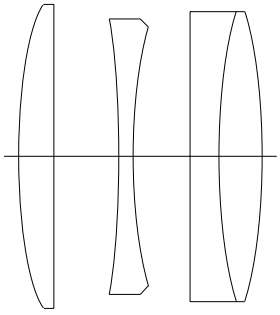

Figure 4: Fixed focal length lens. (From [3])

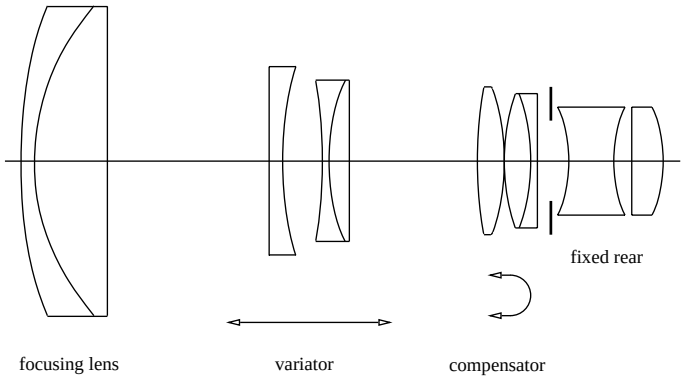

Figure 5: Variable focal length (zoom) lens. (From [3])

3 A Taxonomy of Image Centers

In machine vision the most commonly used definitions of image center are the *focus of expansion* and the *center of perspective projection*. The *numerical center* of the image (i.e., digitizer) coordinates is also commonly used, but unlike the other two it does not involve the measurement of a system's actual imaging properties. We base our taxonomy on the number different lens settings that must be used in order to establish the image center. The center of image coordinates belongs to the class of techniques that we call *non-imaging* which require no image measurements. 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 camera setting. The focus of expansion approach belongs to a third class that we call *multi-image techniques* that measure properties that occur between two or more images taken at different camera settings. In this last class the image center is defined in terms of the differences between images taken at different lens settings and should not be confused with simply tracking one of the single image techniques over different lens settings.

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 3.2.2) 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 3.2.5) is an example of this type of technique.

We can name at least 16 different definitions of image center under this taxonomy. By class they are:

Non-imaging

- Numerical Center of Image/Digitizer Coordinates (section 3.1.1)
- Center of Sensor Coordinates (section 3.1.2)

Single image

Feature based

- Center of Radial Lens Distortion (section 3.2.1)
- Center of Perspective Projection (section 3.2.2)
- Center of Lines of Interpretation (section 3.2.3)
- Center of Field of View (section 3.2.4)

Non-feature based

- Center of an Autocollimated Laser (section 3.2.5)
- Center of cos^{4th} Radiometric Falloff (section 3.2.6)
- Center of Vignetting/Image Spot (section 3.2.7)
- Center of Focus/Defocus (section 3.2.8)

Multi-image

Feature based

- Center of Expansion (section 3.3.1)
  - From Focus
  - From Zoom
  - From Aperture
  - From Image Band
- Focus of Expansion (section 3.3.2)

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. They are video output cameras (also called closed circuit television or CCTV cameras) and non-video digital output cameras (also called scientific, slow scan or pixel clocked cameras).

In video output cameras each row of the CCD is scanned off of the sensor and converted to a continuous analog signal. The continuous analog signal is then 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, and depends on the synchronization of the digitizer to the start of the analog signal for each row and on the relative sampling 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.

3.1.1 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.[1]

[1] Throughout this paper we specify the image center in xy image coordinates, in pixels, where x corresponds to column number in the image and y corresponds to row number.

3.1.2 Center of Sensor Coordinates

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

$$\begin{aligned} c_x &= (c_{x\ sensor} - k_x) \times \frac{f_{sensor\ clock}}{f_{digitizer\ clock}} \\ c_y &= c_{y\ sensor} - k_y \end{aligned}$$

where

$c_{x\ sensor}$	is the center of the sensor in pixels in the x direction,
$c_{y\ sensor}$	is the center of the sensor in pixels in the y direction,
k_x	is the number of sensor columns skipped over before digitizing starts,
k_y	is the number of sensor rows skipped over before digitizing starts,
$f_{sensor\ clock}$	is the frequency that sensor elements are clocked off of the CCD and
$f_{digitizer\ clock}$	is the frequency at which the digitizer samples.

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

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 unautomated or even fixed.

3.2.1 Center of Radial Lens Distortion

Lens distortion is the displacement of the image of a point away from the position that is predicted by a perfect perspective projection by the camera. Displacements that are along radial lines from the center of an image are called radial lens distortions. For 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

$$\begin{aligned} x_u &= (x_d - c_x)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \cdots) + c_x \\ y_u &= (y_d - c_y)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \cdots) + c_y \end{aligned}$$

$$r = \sqrt{\left[\frac{d_x}{s_x}(x_d - c_x)\right]^2 + \left[d_y(y_d - c_y)\right]^2}$$

where d_x, d_y and s_x are camera constants and κ_i are the distortion coefficients.

In Tsai's camera calibration approach [7] the data used to calibrate the camera system consists of the world coordinates (x_w, y_w, z_w) and the measured image coordinates (x_i, y_i) for a set of feature points. The image plane error is the difference between the image coordinates of a feature point given by the calibrated camera model and the actual measured image coordinates (x_i, y_i) given in the calibration data. We define the center of radial lens distortion as the image center that produces the minimum average image plane error for the calibration data. To determine this center we follow Tsai's camera calibration algorithm with a non-linear optimization of all of the calculated model parameters plus the c_x and c_y parameters.

3.2.2 Center of Perspective Projection

Under perspective projection the images of two lines that are parallel in object space but not parallel to the camera's sensing plane will appear to intersect at a location (u, v), called a vanishing point. In the image of 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 [8],

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

An image of three sets of parallel lines that are mutually orthogonal can be easily obtained by imaging the corner of a right angled cube and using the cube's nine visible edges, as shown in Figure 6.

3.2.3 Center of Lines of Interpretation

In a camera each pixel lies on a line-of-sight or line of interpretation through object space. All lines of interpretation intersect behind the image plane at a location called the *viewing point* for the camera. 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[2] 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 Figure 7.

To determine the equations of lines of interpretation we use a target consisting of two points, T_1 and T_2, mounted on the ends of a rod. The rod is manipulated manually until the two points 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 points defines the interception point of the line of interpretation with the image plane.

3.2.4 Center of Field of View

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

To measure the center of the field of view we position a straight edge target such that the target's edge 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 center of field of view is then determined by finding the location of the intersection of the edges in the two superimposed images, as shown in Figure 8.

[2] As with the center of perspective projection, the lines of interpretation technique determines the image center based on a limited number of image measurements, generally without regard to underlying phenomena like radial lens distortion. As a result the image centers from these techniques tend not to be very robust.

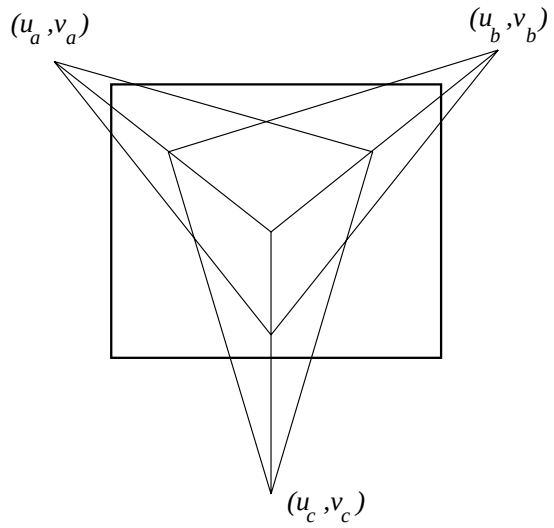

Figure 6: Vanishing points of a right angled cube.

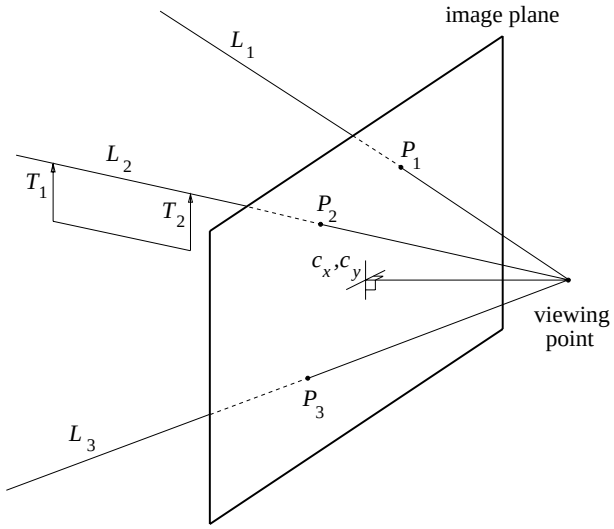

Figure 7: Center of lines of interpretation.

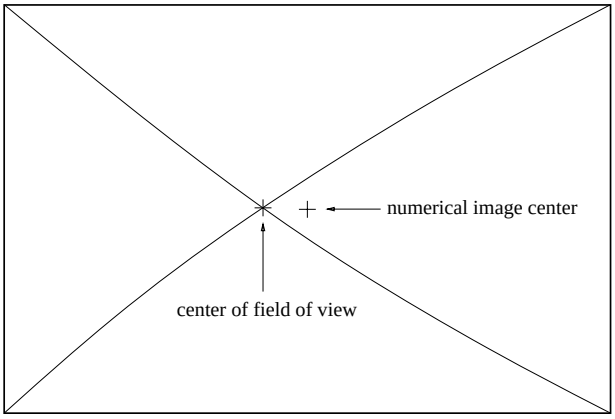

Figure 8: Center of field of view.

3.2.5 Center of an Autocollimated Laser

In an ideal lens the centers of the radii of curvature for all of the lens elements would 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 radii of curvature for the lens elements do not fall on a line. Instead manufacturing tolerances result in decentering and tilting of the lens elements relative to one another. As a result there is no single path which will have all of the reflected light return directly along the same path; the reflected light 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 Figure 9. The laser beam serves as an intense highly collimated light ray. As the beam travels down the lens the lens elements reflect part of it back out the lens and onto the white screen. By manipulating the position and orientation of the laser and the lens the reflections coming back from the lens can be roughly lined up with the hole that the laser is being passed through. At this point the laser is said to be autocollimated, with the laser beam traveling along the "best" optical axis for the lens. An image taken of the laser in this configuration yields the image center for an autocollimated laser.

3.2.6 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 [3]. This falloff can be clearly seen in Figure 10 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 is the subject of section 3.2.7.

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 patch of pixels at the center

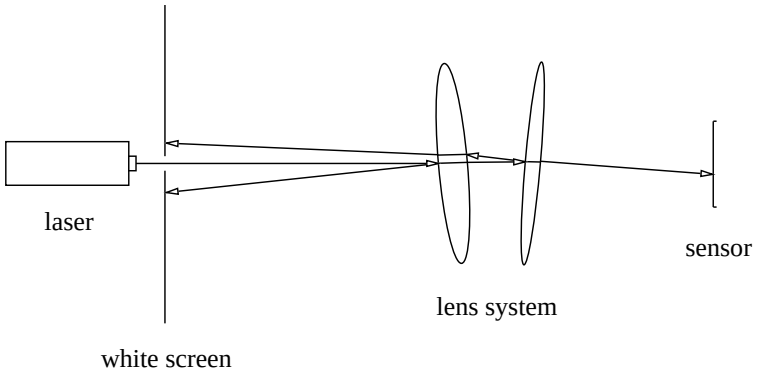

Figure 9: Center of an autocollimated laser.

of the image 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 and thus 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 can be done in closed form for the polynomial.

3.2.7 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 Figure 11. This phenomenon is known as vignetting and leads to a reduction of the image illumination at increasing distances away from the axis [3]. Figure 12 shows sharply defined vignetting in an image of a uniform white field.

To determine the center of the vignetting for 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.

Note: In virtually all commercial camera systems the size of the lens' image spot (the *image format*) is chosen to 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.

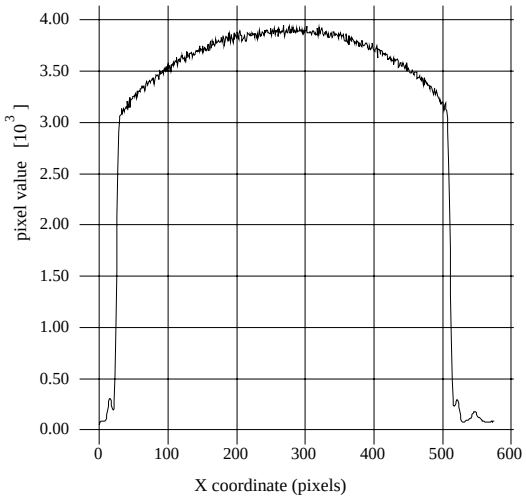

Figure 10: Pixel intensity profile for row 200 from Figure 12.

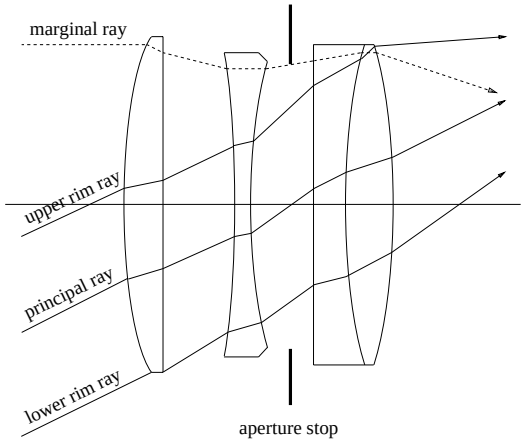

Figure 11: Vignetting in a lens. (From [3])

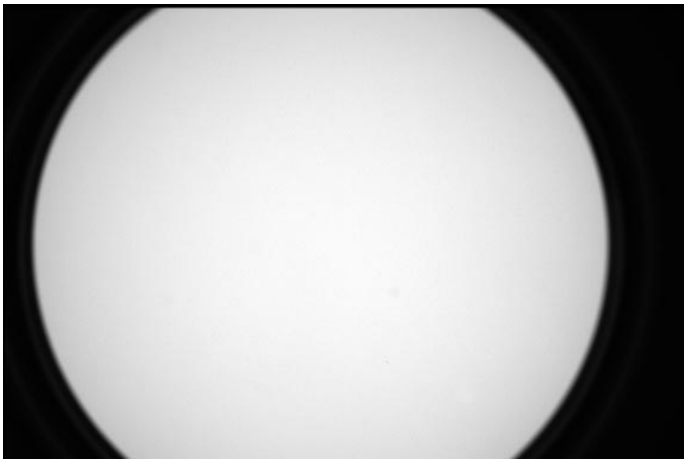

Figure 12: Image of a uniform white field showing sharply defined vignetting.

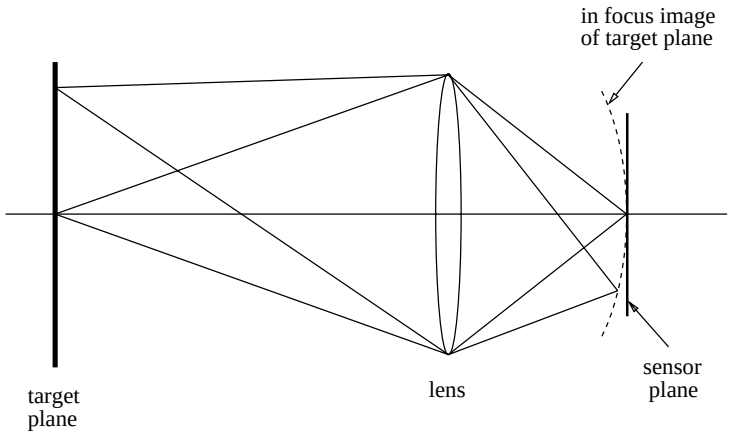

Figure 13: Field curvature in a lens.

3.2.8 Center of Focus/Defocus

With an ideal lens a planar target in front of the lens would produce an image of the target behind the lens that would also be planar. With real lenses though 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 Figure 13, and its ideal planar image is known as the field curvature of the lens. In practical terms field curvature means that the focussed distance of the lens varies across the field of view of the lens, as has been shown in [5].

To measure the center of focus or defocus we start by imaging a plane that is nearly perpendicular to the optical axis and parallel to the sensing plane of the camera. The field curvature of the lens introduces local defocusing in the image of the plane. If the target plane is nearly perpendicular to the optical axis then 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 attenuate the effect of the low frequency variations in the image intensity due to factors such as illumination and the cos^{4th} law. The image center is then determined by fitting a radially symmetric model to the resulting pattern of focus and defocus.

3.3 Multi-image Definitions

The last class in our image center taxonomy is 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 as the lens parameters are varied than they do about about the image properties covered by the previous single image techniques.

Changing any lens parameter will cause changes in the image parameters, including for example the magnification, focussed 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.

3.3.1 Center of Expansion for Focus, Zoom, Aperture and Image Band

Given two images taken at different magnifications exactly one position in the scene in both images will remain in the same place in 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 $P_1 \ldots P_n$ in image I_1 and the corresponding points $Q_1 \ldots Q_n$ in image I_2, then there exists a center of expansion C that satisfies the constraint

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

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

$$k_{x_{ij}} = \frac{q_{x_i} - q_{x_j}}{p_{x_i} - p_{x_j}}, \quad i > j, \quad | q_{x_i} - q_{x_j} | > threshold$$

$$k_{y_{ij}} = \frac{q_{y_i} - q_{y_j}}{p_{y_i} - p_{y_j}}, \quad i > j, \quad | q_{y_i} - q_{y_j} | > threshold$$

$$k = \frac{\Sigma k_{x_{ij}} + \Sigma k_{y_{ij}}}{n_x + n_y}$$

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 2 to 3 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 start by defining 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 minimizes the squared error we differentiate e with respect to c_x and c_y and set the results equal to zero, yielding

$$\begin{aligned} c_x &= \frac{\sum_{i=1}^{n} (k q_{x_i} - p_{x_i})}{n(k-1)} \\ c_y &= \frac{\sum_{i=1}^{n} (k q_{y_i} - p_{y_i})}{n(k-1)} \end{aligned}$$

Normally image magnification is changed by varying a lens' zoom, however magnification can also be changed by varying the focus, aperture and color band of the lens [9]. Accordingly centers of expansion can be defined for all four lens parameters.

Table 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

3.3.2 Focus of Expansion

In what is known as the 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.

3.4 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 [7]. The non-coplanar data used in the calibration was obtained using a planar target containing a total of 225 uniformly spaced reference points (a 15×15 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 14 shows the mean value of the image plane error for Tsai's technique for a range of different image centers. For an image center equal to the numerical center of the image at [288, 192] (point 1 in Figure 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 Figure 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.[3] The results, drawn to scale in Figure 1 and listed in Table 1, show variations of over 90 pixels in the x direction and over 40 pixels in the y direction (image size is 576×384 pixels).

[3]The first nine measurements were made with a focussed 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 focussed distance varying from 1.2 - 2.0 m, an effective focal length of 20 mm and an aperture of $f/12.5$.

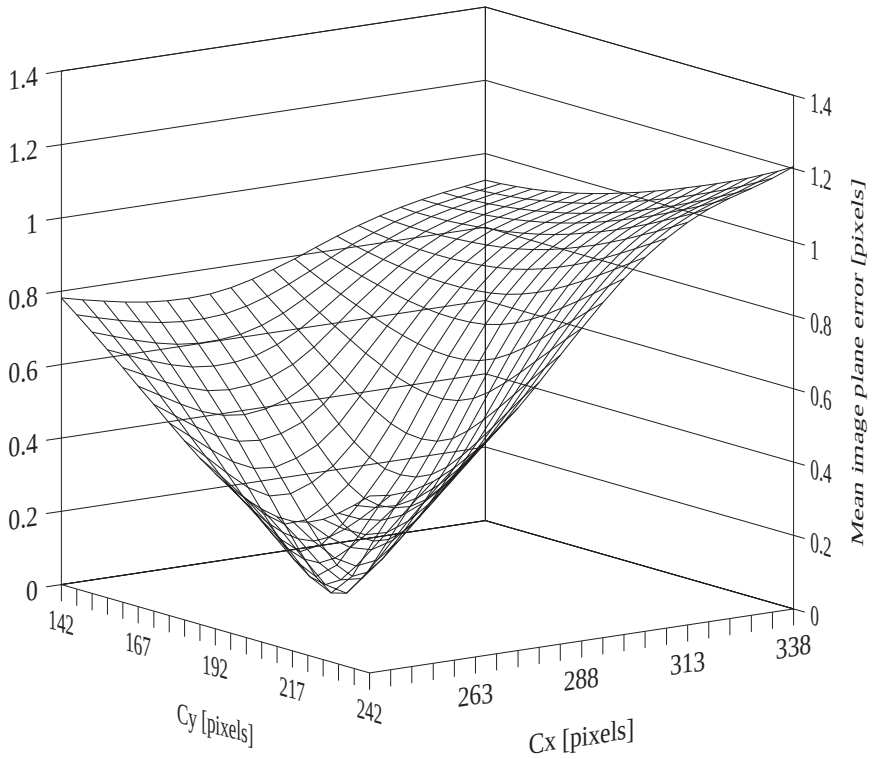

Figure 14: Mean image plane error as a function of image center used during calibration.

4 Image Center in Variable Parameter Lenses

Varying the focus and zoom of a camera lens changes the alignment of the lens components and causes the image center and 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 for any given lens setting. But knowing how an image center shifts can also be important for other aspects of the camera's calibration. In Tsai's camera model the image center parameters c_x and c_y are used both as the center of radial lens distortion and as the point of intersection of the camera's z axis with the image plane. To maintain model calibration as the lens parameters are varied the coordinates of the model's z axis intercept must be adjusted to compensate for shifts in the camera's field of view. We note that there is no reason that the shifting of the z axis intercept should coincide with the shifting of the center of radial lens distortion. For variable parameter lenses a more accurate approach would be to use two separate image centers to model the two properties.

For our study of image center 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 start by autocollimating the lens at one lens setting. We then step through the full range of focus and zoom settings while the centroid of the image of the laser is recorded. The results, plotted in Figure 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, also 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.[4] Figures 15 and 16 show that the lens has good mechanical repeatability. Figures 17 and 18 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 always approaching a given lens setting from one direction.

[4] The automation for our lens is provided by highly repeatable digital microstepping motors [9], thus all error is due primarily to the mechanical and optical properties of the lens itself.

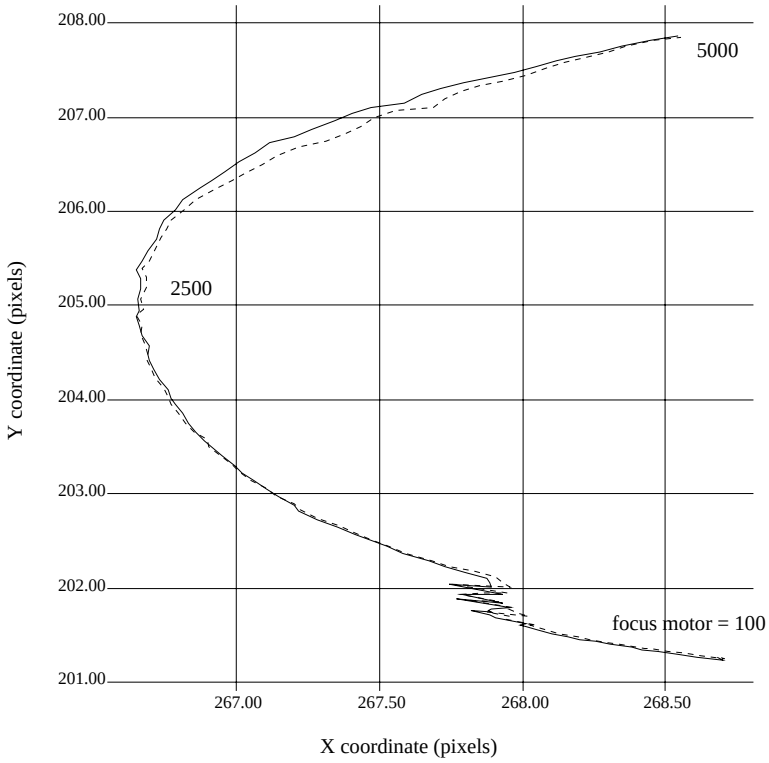

Figure 15: Mechanical repeatability of shift in laser image due to focus motor.

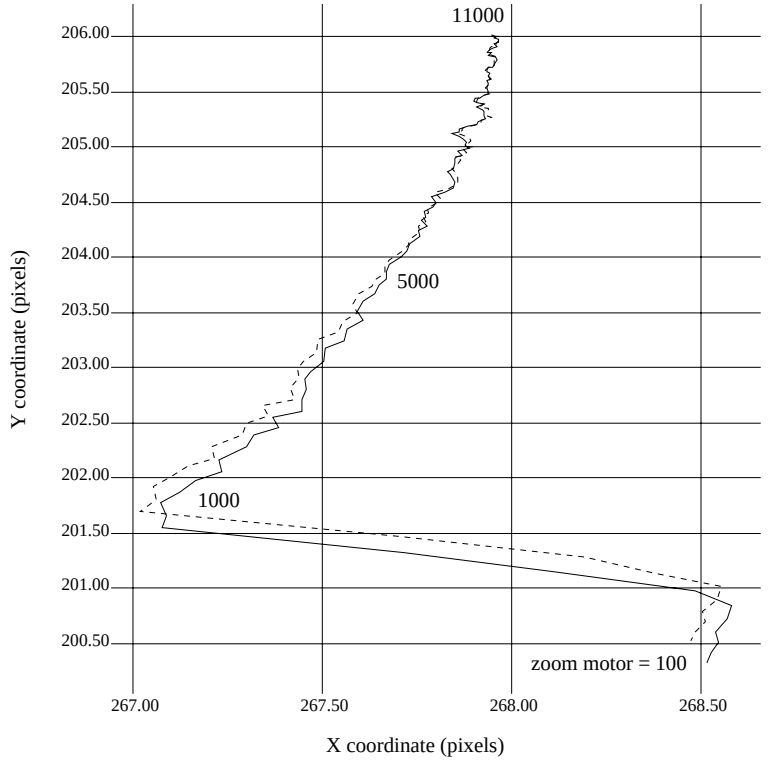

Figure 16: Mechanical repeatability of shift in laser image due to zoom motor.

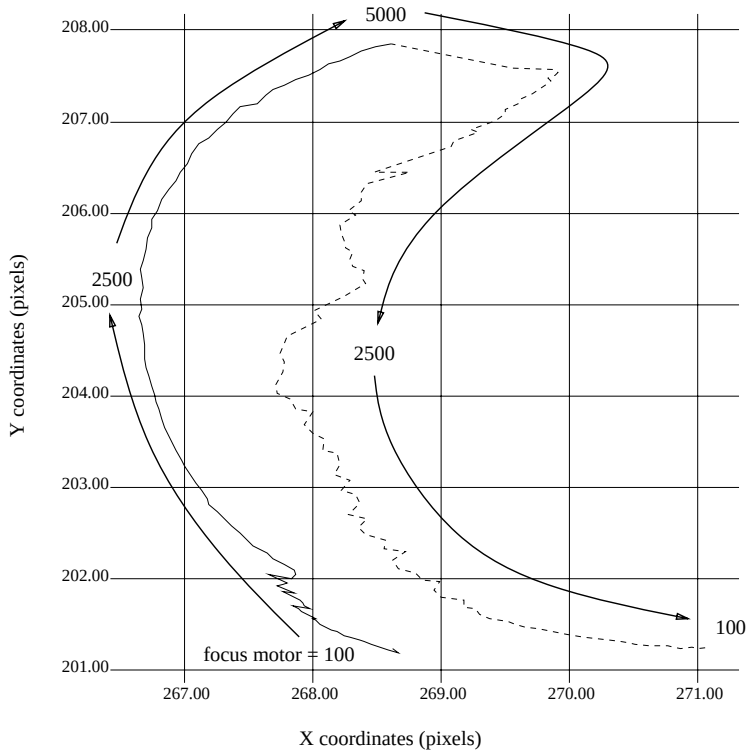

Figure 17: Mechanical hysteresis in shift in laser image due to focus motor.

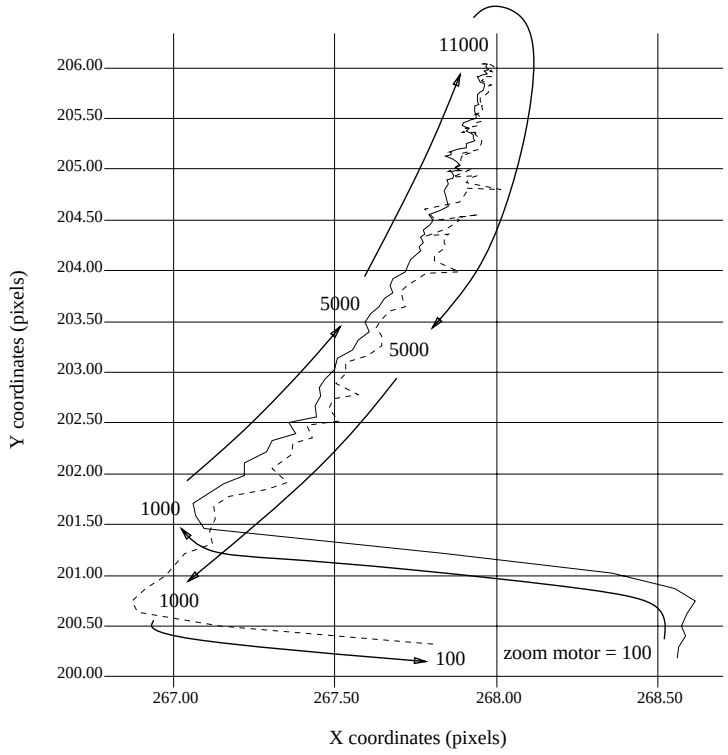

Figure 18: Mechanical hysteresis in shift in laser image due to zoom motor.

5 Conclusions

For an ideal lens camera calibration would involve modeling a 3D to 2D projection through a simple 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 were at least manufactured perfectly they would have perfect radial symmetry around one well defined optical axis which could easily be determined by any one of the 16 methods that we have described in this paper. In practice however, lens manufacturing tolerances result in a wide variation in the location of the image centers for different image properties. Thus 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, the motion is for the most part regular and repeatable and can be modeled and compensated for.

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 calibration accuracy of our camera models.

6 Acknowledgments

This research was sponsored by the Avionics Lab, Wright Research and Development Center, Aeronautical Systems Division (AFSC), U.S. Air Force, Wright-Patterson AFB, OH 45433-6543 under Contract F33615-90-C-1465, ARPA Order No. 7597. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

References

[1] D. C. Brown. Decentering distortion of lenses. *Photometric Engineering*, 32(3):444–462, 1966.

[2] N. Goldberg. *Camera Technology: The Dark Side of the Lens*. Academic Press, San Diego, CA, 1992.

[3] R. Kingslake. *Optical System Design*. Academic Press, New York, NY, 1983.

[4] M. Laikin. *Lens Design*. Marcel Dekker, New York, NY, 1991.

[5] 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, 1992.

[6] W. J. Smith. *Modern Optical Engineering, The Design of Optical Systems*. Optical and Electro-Optical Engineering Series. McGraw-Hill, New York, NY, 1966.

[7] 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.

[8] 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.

[9] R. G. Willson and S. A. Shafer. Precision imaging and control for machine vision research at Carnegie Mellon University. In *Proceedings of Conference on High-Resolution Sensors and Hybrid Systems*, volume 1656, pages 297–314, San Jose, CA, February 1992. SPIE.