What is the center of the image?

Reg G. Willson

3M Engineering Systems and Technology, 3M Center, Building 518-1-01, St. Paul, Minnesota 55144–1000

Steven A. Shafer

School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Received July 22, 1993; revised manuscript received May 31, 1994; accepted June 8, 1994

To model the way that cameras project the three-dimensional world into a two-dimensional image we need to know the camera's image center. First-order models of lens behavior, such as the pinhole-camera model and 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. We present a taxonomy that includes 15 techniques for measuring image center. Several techniques are applied to a precision automated zoom lens, and experimental results are shown.

Key words: Image center, camera modeling and calibration, automated zoom lens, computer vision.

1. CAMERA CALIBRATION AND IMAGE CENTER

The circular construction of lenses used in machine vision gives them imaging properties that are radially symmetric around an image center. To model these properties we need to know where this image center is. Naturally, the accuracy of the model depends on the accuracy of the center.

For a given machine-vision task we may need several models of the camera's imaging properties. For example, we might need a model that describes perspective projection from three dimensions to two dimensions, one that describes the image-plane distortion introduced by the lens, or one that describes the falloff in image intensity as we move toward the edge of the image. Each model describes a specific property of the camera's image formation process, and each may require an image center. In the perspective-projection model the image center may be the normal projection of the center of perspective projection into the sensor plane. In the lens distortion model it may be the center of the distortion pattern. In models of radiometric falloff the image center might be the position of the peak image intensity. The questions are: How do we measure these centers? Are they the same?

Traditionally a camera's imaging properties have been considered to have one image center. For example, in Tsai's camera model[1] the center of radial lens distortion and the center of perspective projection share the same model parameters. Lenz and Tsai[2] described three measurement techniques for calibrating image center for Tsai's model: autocollimating a laser through the lens, changing the lens's focal length, and numerical optimization to find the center that provides the best fit between their camera model and a set of calibration data. An underlying assumption here is that all three techniques will yield the common image center of the two imaging properties being modeled.

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. 1 for one of our camera systems. Thus, image centers are not necessarily interchangeable. Indeed, to model a camera fully we may need several different image centers.

The situation becomes even more complex for an adjustable lens. When camera parameters such as focus and 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 the camera's field of view shifts as a function of the focus and zoom motors of the same camera lens.

Each image-based technique we describe in this paper measures the center of an imaging property. Some measure properties that are seemingly closely related. Some measure the same property with different approaches. And some (e.g., center of the image spot) measure properties for which a model might be of little use. For completeness and comparison we include them all as examples in our taxonomy.

We begin this paper by examining why different imaging 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 measurement techniques of image center based on the number of lens settings that are required for determining the image center. Procedures for measuring 15 different image centers are given, and experimental results are then presented for 10 of these methods. We conclude by examining how image center and field of view move in an automated zoom lens.

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

0740-3232/94/112946-10$06.00

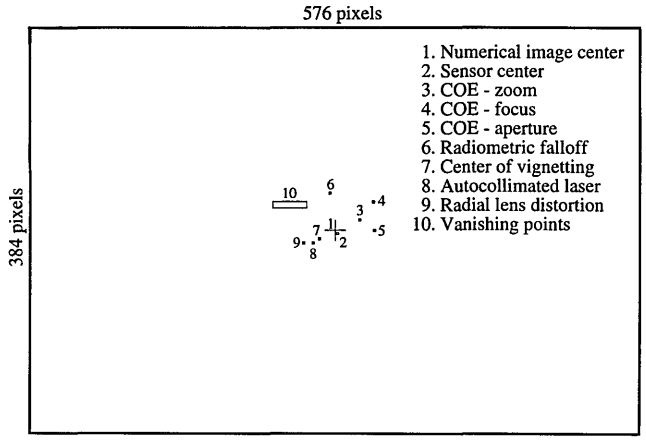

Fig. 1. Different image centers for the Fujinon/Photometrics camera system.

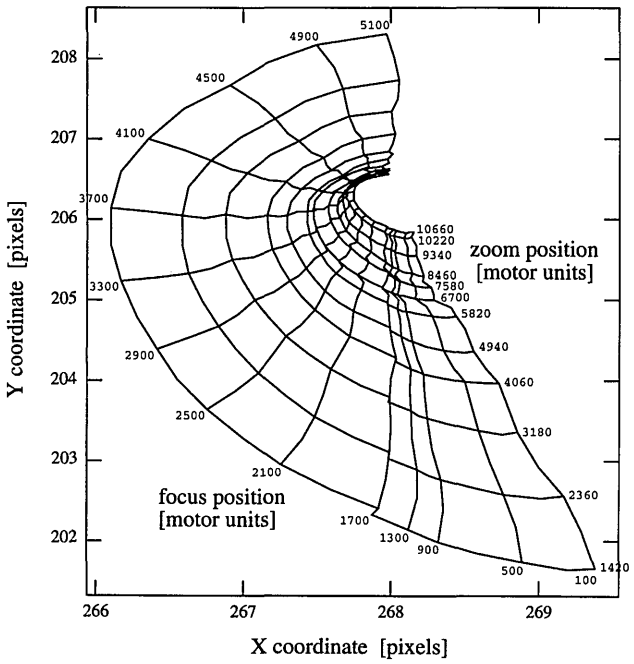

Fig. 2. Shift in image center as a function of the focus and zoom motors.

camera's sensing plane. The optical axis is defined as the line that passes through the centers of the optical components in the lens. In real lenses the optical axis is not so easily defined. The type of complication that arises 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 its imaging properties radially symmetric. In real lenses things are not so easy. For a simple lens element (see Fig. 3) there are actually two axes of symmetry, one optical and one mechanical. The optical axis is defined as the straight line that joins 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 len's edge. Ideally the optical and mechanical axes coincide; in practice, though, they do not. The tolerance between them is called decentration.[3]

In a compound lens two or more lens elements are aligned and mounted together to form a complete lens. Ideally all 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.[4] Thus the different imaging properties of the lens do not necessarily have a common axis of symmetry.

With adjustable lenses one can change the focus and magnification 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 each 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,[5] whereas the sliding of a lens group will cause a translational motion of the len'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. 4, typically all the lens elements are mounted in a single assembly. To vary the lens's focus one changes the separation between the lens assembly and the camera sensor by moving the lens assembly with either a rotational or translational mechanism. A less common focusing method found in newer 35-mm, autofocus lens designs involves the movement of a small, lightweight element within the lens's optics to vary the focus of the image.[6]

In lenses with variable focus and variable focal length (i.e., zoom lenses) (see Fig. 5), one changes the focal length

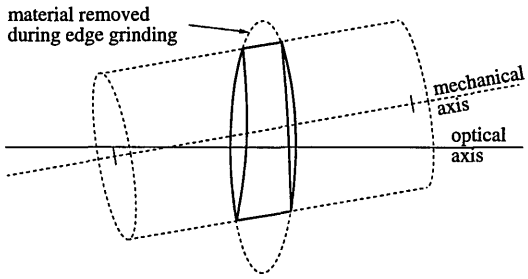

Fig. 3. Decentration for a simple lens.

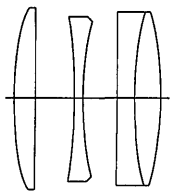

Fig. 4. Fixed focal length lens (from Ref. 10).

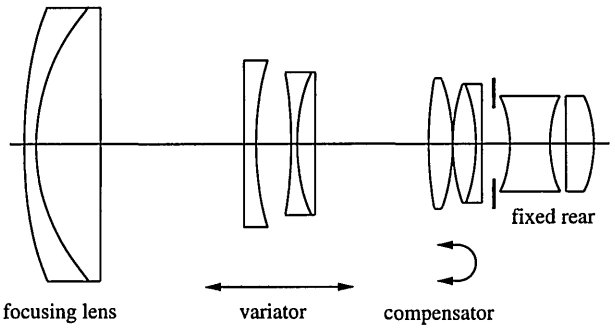

Fig. 5. Variable focal length (zoom) lens (from Ref. 10).

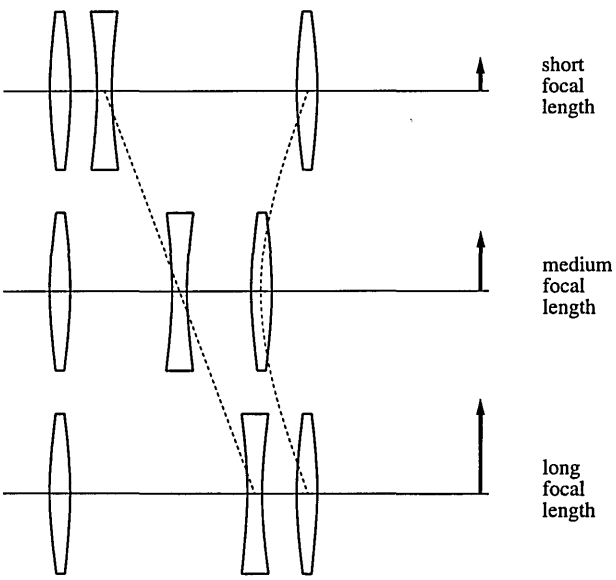

Fig. 6. Nonlinear motion of lens groups during mechanical compensation (from Ref. 3).

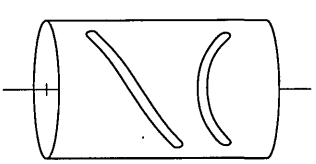

Fig. 7. Mechanical compensation cam (from Ref. 3).

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 nonlinear motion, as illustrated in Fig. 6. Typically this type of mechanical compensation is accomplished with a pin that slides in a rotating cam, such as the one shown in Fig. 7. While the cam moves with a rotating motion during zooming, the lens groups themselves move with a translational motion. One can typically vary the focus in zoom lenses by using a rotational mechanism on the front lens group. For examples of the mechanisms used in several commercial zoom systems, see Horne.[7]

3. 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. Whereas 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 nonimaging 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 the 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 nonfeature based. Feature-based techniques involve the detection of feature points in the image followed by the application of a geometric interpretation of the three-dimensional to two-dimensional projection to yield an image center. The center-of-vanishing points (Subsection 3.B.2) is an example of this type of technique. Nonfeature-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 (Subsection 3.B.5) is an example of this type of technique.

Using this taxonomy, we can describe at least 15 different ways of measuring image center and divide them into the following classes:

Nonimaging
- Numerical center of image/digitizer coordinates (Subsection 3.A.1)
- Center of sensor coordinates (Subsection 3.A.2)

Single image
Feature based
- Center of radial lens distortion (and perspective projection) (Subsection 3.B.1)
- Center of vanishing points (Subsection 3.B.2)
- Center of lines of interpretation (Subsection 3.B.3)
- Center of field of view (Subsection 3.B.4)

Nonfeature based
- Center of an autocollimated laser (Subsection 3.B.5)
- Center of radiometric falloff (Subsection 3.B.6)
- Center of vignetting/image spot (Subsection 3.B.7)
- Center of focus/defocus (Subsection 3.B.8)

Multi-image
Feature based
- Center of expansion (Subsection 3.C.1)
  - From focus
  - From zoom
  - From aperture
  - From color band
- Focus of expansion (Subsection 3.C.2)

A. Nonimaging Techniques

By definition nonimaging techniques do not make use of imaging 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 nonvideo digital output cameras (also called scientific, slow-scan, or pixel-clocked 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 nonvideo 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.

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

$$C_x = \frac{x_{\max} + x_{\min}}{2}, \qquad C_y = \frac{y_{\max} + y_{\min}}{2},$$

where $x_{\max}$, $x_{\min}$, $y_{\max}$, and $y_{\min}$ are the maximum and minimum column and row numbers, respectively. (Throughout this paper 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.)

2. *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

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

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 nonvideo digital output cameras h_x and h_y are integers and $f_{\text{sensor clock}} = f_{\text{digitizer clock}}$. Since neither of the nonimaging techniques makes use of actual imaging properties they are useful only as default values or initial guesses for image center.

B. Single-Image Techniques

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.

1. *Center of Radial Lens Distortion (and Perspective Projection)*

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

$$X_u = (X_d - C_x)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \cdots) + C_x,$$
$$Y_u = (Y_d - C_y)(1 + \kappa_1 r^2 + \kappa_2 r^4 + \cdots) + C_y,$$
$$r = \left\{ \left[\frac{d_x}{s_x} (X_d - C_x) \right]^2 + [d_y (Y_d - C_y)]^2 \right\}^{1/2},$$

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

To determine the center of radial lens distortion and perspective projection we use the model fitting approach described by Lenz and Tsai.[2] We begin by using a rough estimate of image center to calibrate our lens using Tsai's camera model and calibration algorithm.[1] In Tsai's algorithm the image center is a precalibrated constant. We then use nonlinear optimization on the complete camera model (including the image center) to find the model parameters that produce the minimum image-plane error for the calibration data. The details of this procedure can be found in Ref. 8.

Since the fitted camera model's image center parameter is used for both radial lens distortion and perspective projection, the resulting best-fit image center represents a combination of the image centers for these two properties. To obtain accurate centers for the individual properties we would have to fit a model that used separate centers for the two properties.

2. *Center of Vanishing Points*

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 is not parallel with the others 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 one can calculate the center-of-perspective projection for the camera from the three vanishing points, using the formula

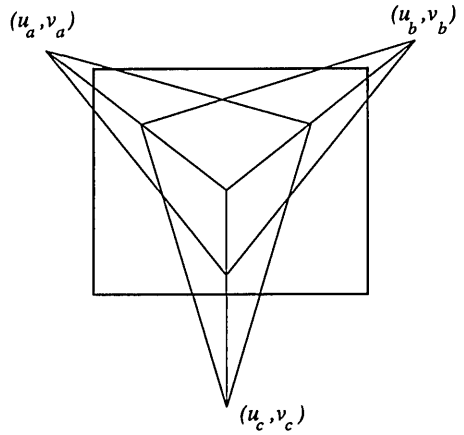

Fig. 8. Vanishing points of a right-angled cube.

presented in Ref. 9:

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

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. 8. Since the center-of-vanishing points relies on the lens's perspective-projection properties, the resulting image center is best suited for perspective-projection models.

3. 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 the lines of interpretation should intersect behind the image plane at one 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 noncoplanar lines of interpretation, L_1, L_2, and L_3, and the two-dimensional image coordinates of their intersection with imaging planes, P_1, P_2, and P_3. The intersection of the lines of interpretation determines the three-dimensional coordinates of the viewing point. The relative two-dimensional 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. 9.

To determine the equations of lines of interpretation we use a target that consists 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 is then used to determine the equation in three-dimensional 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 vanishing-points technique, the lines-of-interpretation technique relies on the lens's perspective-projection properties, and thus the resulting image center is best suited for perspective-projection models. Both techniques use 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, their image centers tend not to be robust.

4. 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. One can then determine the field-of-view center by finding the location of the intersection of the edges in the two superimposed images, as shown in Fig. 10.

Since the field-of-view technique does not directly measure any of the conventionally modeled imaging properties (i.e., perspective projection, radial lens distortion, or radiometric falloff), the image center that it produces is not

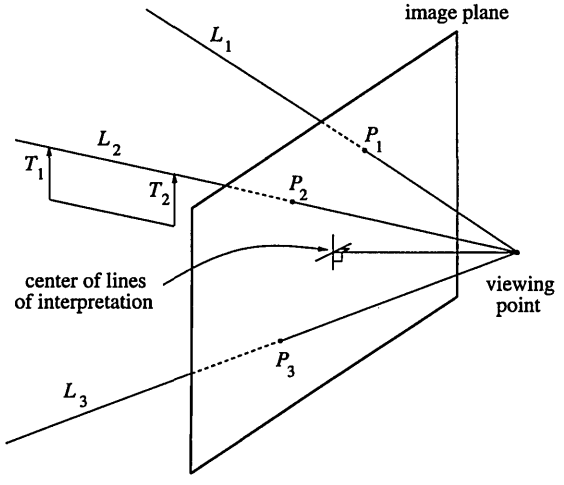

Fig. 9. Center of lines of interpretation.

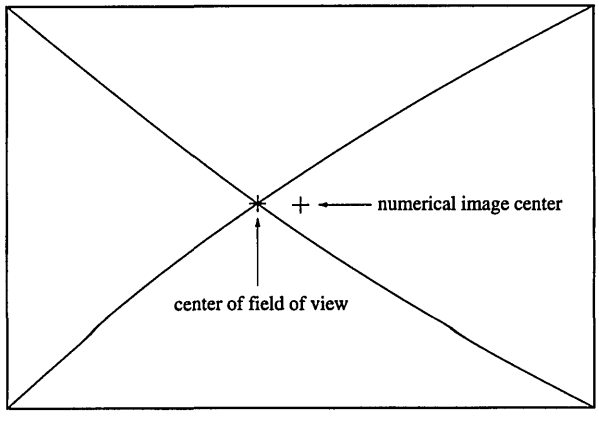

Fig. 10. Center of field of view.

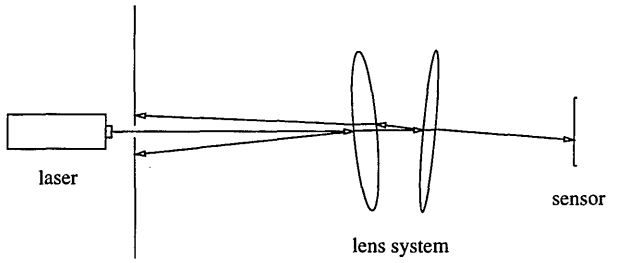

Fig. 11. Center of an autocollimated laser.

strictly applicable in models of these properties. However, the center of the field of view can be useful in simple pointing models for the camera system.

5. 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 that travels 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, because of 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 described by Lenz and Tsai[2] 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. 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, one can roughly line up the reflections coming back from the lens 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 travels 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.

Since an autocollimated laser does not directly measure any of the conventionally modeled imaging properties, the image center that it produces is not strictly applicable in models of these properties. However, the technique's ability to provide image center measurements across wide ranges of focus and zoom makes it useful for tracking the effects of optical alignment changes in variable-parameter lenses.

6. Center of Radiometric Falloff

In a lens system the illumination of the image plane is found to decrease away from the optical axis at least with the fourth power of the cosine of the angle of obliquity with the optical axis.[10] This falloff can be clearly seen in Fig. 12, which shows the profile of a scan line taken from the image of a more-or-less uniform white field (Fig. 13). The abrupt drop in intensity values near the edges is due to vignetting, which is discussed in Subsection 3.B.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 try to measure the intensity across the full field of view at once, instead we 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.

In real lenses many factors contribute to radiometric falloff, and thus a $\cos^{4th}$ model is not necessarily the best (or the easiest) model to fit to the data. The 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 radiometric falloff—is then given by

$$C_x = \frac{a_{01}a_{11} - 2a_{10}a_{20}}{4a_{20}a_{02} - a_{11}^2}, \qquad C_y = \frac{a_{10}a_{11} - 2a_{01}a_{02}}{4a_{20}a_{02} - a_{11}^2}.$$

The polynomial can be fitted to the data in closed form by

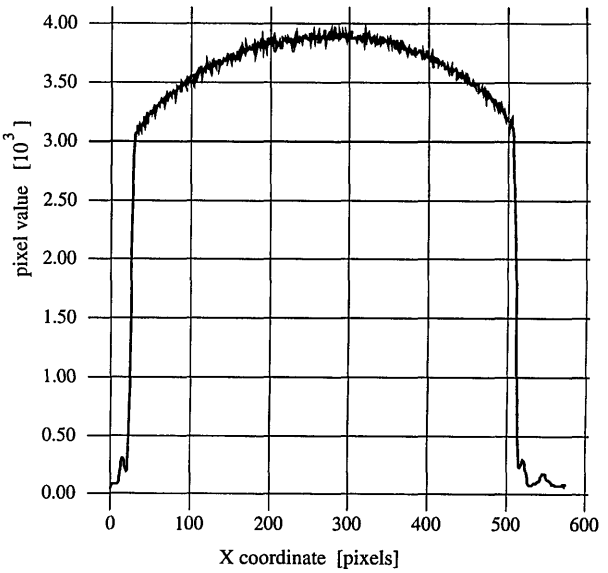

Fig. 12. Pixel intensity profile for row 200 from Fig. 13.

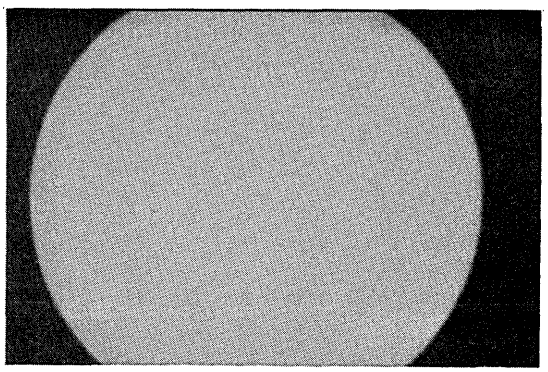

Fig. 13. Image of a uniform white field showing sharply defined vignetting.

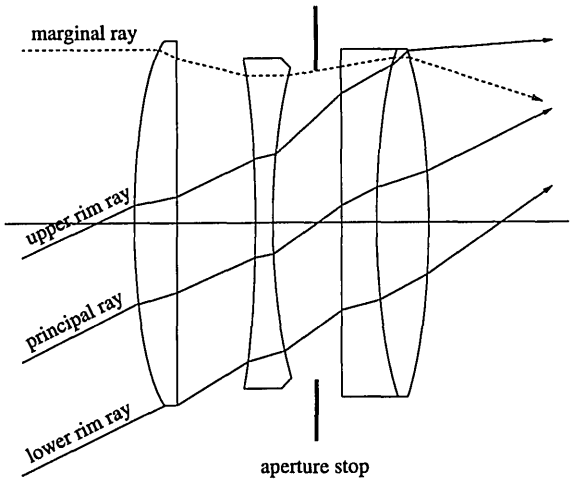

Fig. 14. Vignetting in a lens (Ref. 10).

straightforward linear algebra.[11] The image center from this technique is directly applicable in radiometric falloff models.

7. Center of Vignetting/Image Spot

For angles nearly parallel to the optical axis the edges of the bundle of rays that pass completely through the lens are usually 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 by the aperture stop, as shown in Fig. 14. This phenomenon is known as vignetting and leads to a reduction of the image illumination at increasing distances away from the axis.[10] 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. We then fit a circle 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 be used only when the lens is removed from the camera system or in camera systems for which the image format is smaller than the sensor size.

Since this technique does not directly measure any of the conventionally modeled imaging properties, the image center that it produces is not strictly applicable in models of these properties. Although the properties that this technique directly measures (vignetting and image spot) are not particularly useful to model on their own, for completeness we have included it in our taxonomy.

8. 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 a plane's real image, illustrated in Fig. 15, 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 and Stewart.[12]

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 measure the amount of defocus more accurately we use a target plane that contains 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 that are due to factors such as illumination and radiometric falloff. We then determine the image center by fitting a radially symmetric two-dimensional surface to the resulting pattern of focus and defocus.

Since image defocus is not directly related to any of the conventionally modeled imaging properties, this image center is not strictly applicable in models of these properties. This image center would, however, be applicable in the field curvature models that would be necessary for systems with which we try to compute dense depth from focus accurately.

C. Multi-Image Techniques

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 imaging properties covered by single-image techniques. Thus the image centers produced by these techniques are not strictly applicable in any of the conventionally modeled imaging properties. The centers can, however, be useful in registering images taken at different lens settings.

Changing any lens parameter will cause changes in the image parameters, including, for example, the magnification, focused distance, and intensity of the image. Although any of these imaging properties might be used as the basis of a multiimage definition of image center, image magnification is the easiest to measure.

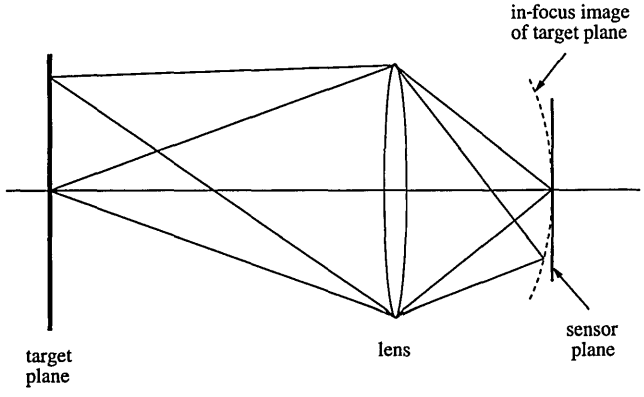

Fig. 15. Field curvature in the image produced by a thin lens.

1. 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 $P_1 \ldots P_n$ in image I_1 and the corresponding points $Q_1 \ldots Q_n$ in image I_2, the center of expansion C satisfies the constraint

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

where

$$k = m_1/m_2 .$$

One can estimate the relative image-plane magnification k from the change in relative separation of the points in each image by evaluating

$$k_{x_{ij}} = \frac{q_{xi} - q_{xj}}{p_{xi} - p_{xj}}, \quad i > j, \quad |q_{xi} - q_{xj}| > \text{threshold},$$

$$k_{y_{ij}} = \frac{q_{yi} - q_{yj}}{p_{yi} - p_{yj}}, \quad i > j, \quad |q_{yi} - q_{yj}| > \text{threshold},$$

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

where n_x and n_y are the number of points in the x and y directions that pass 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 2–3 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 first define the squared error for the center as

$$e_{x_i} = (C_x - p_{x_i}) - k(C_x - q_{x_i}),$$

$$e_{y_i} = (C_y - p_{y_i}) - k(C_y - q_{y_i}),$$

$$e = \sum_{i=1}^{n} ({e_{x_i}}^2 + {e_{y_i}}^2).$$

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

$$C_x = \frac{\sum_{i=1}^{n} (kq_{x_i} - p_{x_i})}{n(k-1)}, \qquad C_y = \frac{\sum_{i=1}^{n} (kq_{y_i} - p_{y_i})}{n(k-1)}.$$

Normally one can change image magnification by varying a len's zoom; however, one can also change magnification by varying the focus, aperture, and color band of the lens.[13] Thus centers of expansion can be defined for all four lens parameters.

For any two images taken at different lens settings the center of expansion and the relative image magnification can be used to scale and register one image with the other.

2. Focus of Expansion

In the focus-of-expansion approach described by Lenz and Tsai[2] and others, 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.

Given the degree to which a camera's alignment can vary with lens parameters (see Fig. 2) and the averaging nature of the focus-of-expansion approach, it is not clear what the image center from this approach would really be useful for.

D. 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.[1] The noncoplanar data used in the calibration were obtained by use of a planar target that contains a total of 225 uniformly spaced reference points (a 15 × 15 grid) mounted on a precision 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 16 shows the mean image-plane error for a range of different image centers used in a Tsai calibration on one data set. For an image center equal to the numerical center of the image at [288, 192] (point 1 in Fig. 1) the mean and standard deviation of the image-plane error are 0.553 and 0.413 pixels, respectively. 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. 1), where the mean and the standard deviation of the error drop to 0.084 and 0.046 pixels, respectively.

To illustrate the variation in the position of image cen-

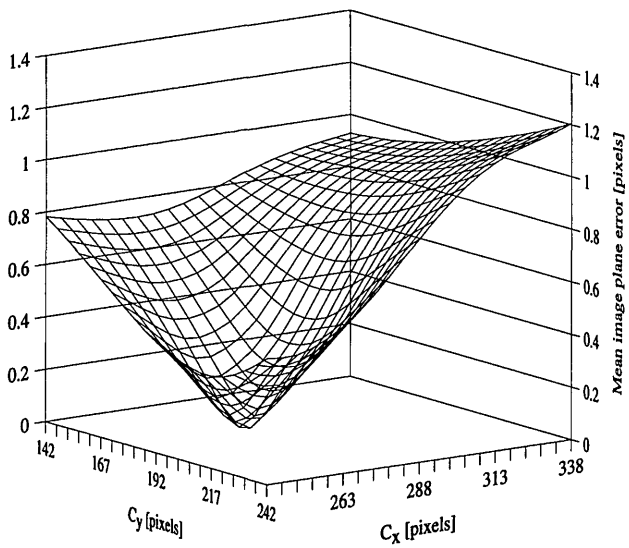

Fig. 16. Mean image-plane error as a function of image center used during calibration.

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 radiometric 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 (and perspective projection)	258.1	203.9
Center of vanishing points	229–261	165–171

ter between different techniques we measured ten different image centers for our automated Fujinon lens. 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 to 2.0 m, an effective focal length of 20 mm, and an aperture of $f/12.5$. The results, drawn to scale in Fig. 1 and listed in Table 1, show variations of more than 90 pixels in the x direction and more than 40 pixels in the y direction (image size is 576 × 384 pixels). The range of values for the center-of-vanishing-points technique represents the results of several trails with slightly different orientations of the right-angled cube that we used to obtain the vanishing points.

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, modeling radially symmetric imaging properties accurately requires knowing the position of the image center. 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 start by autocollimating the lens at one lens setting. We then step through the full range of focus and zoom settings while we record the centroid of the image of the laser. The results, plotted in Fig. 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. For this example the focus was varied from 1 m to ∞ (100–5100 motor units) and the zoom varied from 115 to 15 mm (1420–10,660 motor units). Two observations are worth making here. First, the motion of this 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.

The shape of the plot in Fig. 2 is the product of the cumulative manufacturing tolerances for our particular lens. The plot for another copy of the same lens model would be different.

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. The automation for our lens is provided by highly repeatable digital microstepping motors; thus any error is due primarily to the mechanical and optical properties of the lens itself. Figures 17 and 18 show that the lens has good mechanical repeatability. Figures 19 and 20 show the motion of the laser's image as either the focus or the 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 discontinuity positions of the laser's image in

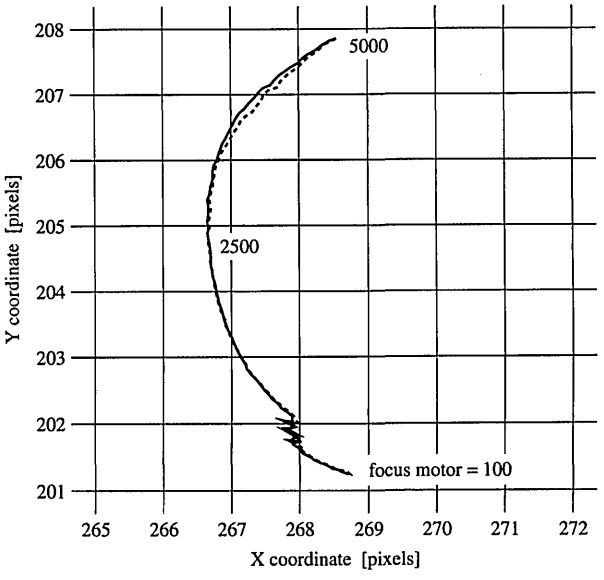

Fig. 17. Mechanical repeatability of shift in laser image that is due to focus motor.

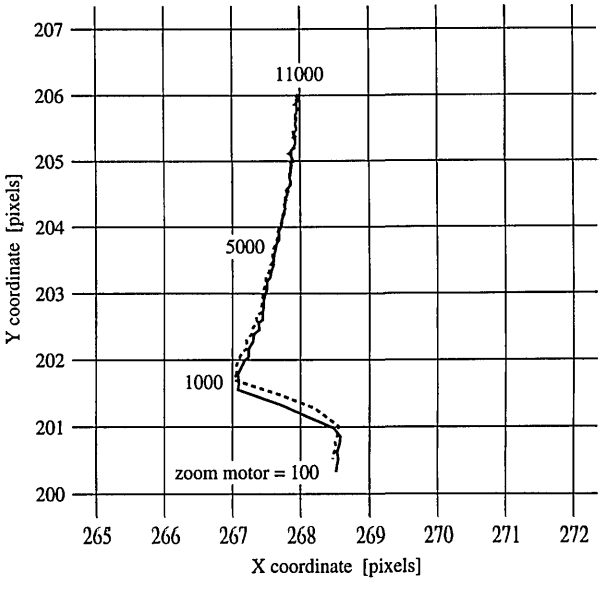

Fig. 18. Mechanical repeatability of shift in laser image that is due to zoom motor.

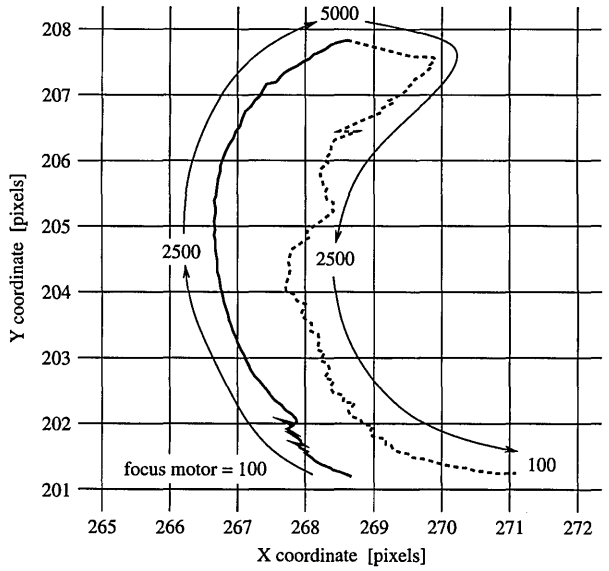

Fig. 19. Mechanical hysteresis in shift in laser image that is due to focus motor.

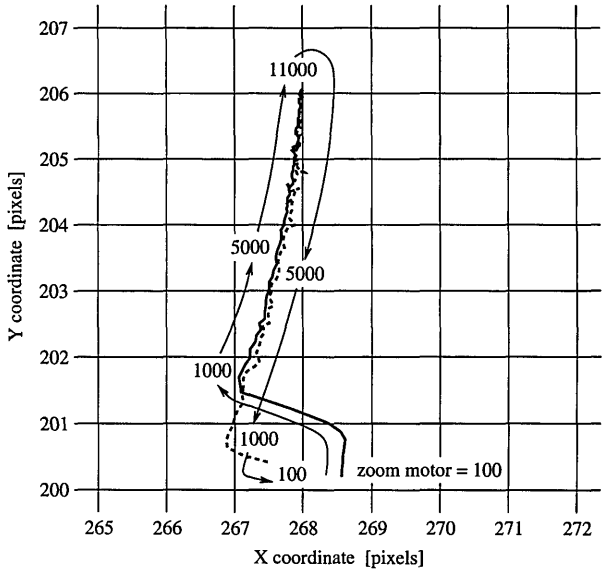

Fig. 20. Mechanical hysteresis in shift in laser image that is due to zoom motor.

Figs. 18 and 20 are due to play in mechanical compensation that is used in the zoom lens. The zoom settings that correspond to the discontinuity mark the point at which the direction of the focus compensation group is reversed.

5. SUMMARY

To model the radially symmetric imaging properties of a camera we need to know their image center. If lenses could be manufactured perfectly they would have perfect symmetry around one well-defined optical axis, which could easily be determined by any one of the 15 methods that we have described in this paper. In practice, however, manufacturing tolerances produce wide variations in the locations of image centers for different imaging properties. Thus different measurements of image centers are not interchangeable.

The image-center calibration problem becomes even more complex in variable-parameter lenses in which 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 (see Willson[8]).

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 measurement technique for each imaging 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.

ACKNOWLEDGMENT

This research is sponsored by the U.S. Department of the Army, Army Research Office, under grant DAAH04-94-G-0006, and by Bell-Northern Research, Ltd. Views and conclusions contained in this paper are those of the authors 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.

REFERENCES

1. R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1986), pp. 364–374.
2. 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* (Institute of Electrical and Electronics Engineers, 1987), pp. 68–75.
3. W. J. Smith, *Modern Optical Engineering, the Design of Optical Systems,* Optical and Electro-Optical Engineering Series (McGraw-Hill, New York, 1966), Chap. 9, p. 243, Chap. 14, p. 420.
4. D. C. Brown, "Decentering distortion of lenses," Photom. Eng. **32,** 444–462 (1966).
5. M. Laikin, *Lens Design* (Dekker, New York, 1991), Chap. 1, p. 17.
6. N. Goldberg, *Camera Technology: the Dark Side of the Lens* (Academic, San Diego, Calif., 1992), Chap. 1, p. 46.
7. D. F. Horne, *Lens Mechanism Technology* (Crane, Russak, New York, 1975), Chap. 2, pp. 25–55.
8. R. G. Willson, "Modeling and calibration of automated zoom lenses," Ph.D. dissertation (Carnegie Mellon University, Pittsburgh, Pa., 1994).
9. L.-L. Wang and W.-H. Tsai, "Computing camera parameters using vanishing-line information from a rectangular parallelepiped," Mach. Vision Appl. **3**(3), 129–141 (1990).
10. R. Kingslake, *Optical System Design* (Academic, New York, 1983), Chap. 6, p. 85, Chap. 7, p. 122, Chap. 15, p. 276.
11. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C,* 2nd ed. (Cambridge U. Press, New York, 1992), Chap. 15, p. 671.
12. H. N. Nair and C. V. Stewart, "Robust focus ranging," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* (Institute of Electrical and Electronics Engineers, New York, 1992), pp. 309–314.
13. R. G. Willson and S. A. Shafer, "Precision imaging and control for machine vision research at Carnegie Mellon University," in *High-Resolution Sensors and Hybrid Systems,* M. M. Blouke, W. Chang, R. P. Khosla, and L. J. Thorpe, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1656,** 297–314 (1992).