NONLINEAR DISTORTION CORRECTION IN ENDOSCOPIC VIDEO IMAGES

Chao Zhang[1] *James P. Helferty*[2] *Geoffrey McLennan*[3] *William E. Higgins*[4]

Penn State University, Department of Electrical Engineering, University Park, PA 16802
[1]zhang@aaron.ee.psu.edu, [2]jph100@psu.edu, [3]geoffrey-mclennan@uiowa.edu, [4]weh@aaron.ee.psu.edu

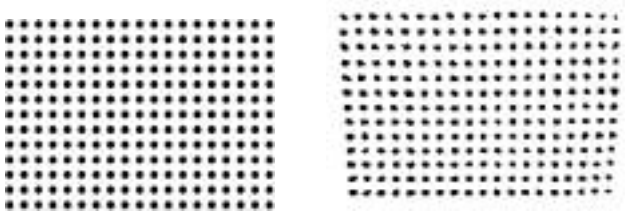
ABSTRACT

Modern video-based endoscopes offer physicians a wide-angle field of view for minimally-invasive procedures. Unfortunately, inherent barrel distortion prevents accurate perception of range. This makes measurement and distance judgment difficult and causes difficulties in emerging applications, such as 3D medical-image registration. Such distortion also arises in other wide field-of-view camera circumstances. This paper presents a distortion-correction technique that can automatically calculate correction parameters, without precise knowledge of horizontal and vertical orientation. The method is applicable to any camera-distortion correction situation. Based on a least-squares estimation, our proposed algorithm considers line fits in both field-of-view directions and global consistency that gives the optimal image center and expansion coefficients. The method is insensitive to the initial orientation of the endoscope and provides more exhaustive field-of-view correction than previously proposed algorithms. The distortion-correction procedure is demonstrated for endoscopic video images of a calibration test pattern, a rubber bronchial training device, and real human circumstances. The distortion correction is also shown as a necessary component of an image-guided virtual-endoscopy system that matches endoscope images to corresponding rendered 3D CT views.

1. INTRODUCTION

Endoscopes are an invaluable tool in pulmonary medicine, urology, orthopedic surgery and gynecology. They permit minimally invasive procedures, involving little or no injury to healthy organs and tissues [1]. Endoscopes, which operate close to their subjects, are designed to have "barrel" distortion to allow greater detail in the center of the display combined with a wide-angle field of view [2]. Such a system puts detail where it is needed most, but results in inconsistent measurements of distance and range due to nonlinear spatial spreading. In an emerging scenario entailing virtual endoscopy, distortion correction will be necessary to accurately calculate an endoscope's position using the corrected video image registered to a rendered 3D CT image [3]. With the two sources registered, a CT image-guided system can be used conceivably to assist a physician in performing more precise endoscopic procedures. Finally, many wide field-of-view camera situations produce distorted images that may require geometric correction.

In general, methods that correct "barrel" distortion must calculate a distortion center and correct both radial and tangential components [4]. It is typically assumed in video camera systems that the distortion is strictly radial, and tangential correction is unnecessary. Geometric calibration to reduce camera-induced geometric distortion in images is well known [2], and the distortion correction can be found in many ways. Shah *et al.* used a low-power laser beam that is adjusted to pass through the optical axis onto an electronic CCD array to find the optical center [4]. Smith *et al.* found distortion parameters using a manual technique where dot locations were chosen off the screen by hand [5]. Tsai proposed a second-order radial correction for 3D camera calibration using off-the-shelf TV cameras and lenses [6]. Stefansic *et al.* [7] used a second-order correction based on Smith *et al.* [5], which described a method for calculating expansion coefficients of orthogonal Chebyshev polynomials. Sawhney *et al.* calibrated a lens with a third-order correction in the creation of video mosaics [8]. Hein *et al.* used a compressed look-up table to correct distortion from magnetic fields applied to an image intensifier for X-ray CT [9].



(*a*) (*b*)

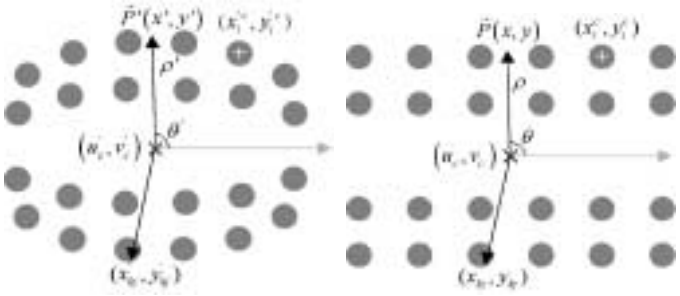
Figure 1. Results from a previously proposed method [10]. (a) The original test pattern; (b) An imperfect trapezoidal-shaped result. The algorithm can provide somewhat inadequate results, since it does not require the line fits to be parallel. This effect becomes greater as the estimated distortion center gets farther from the actual center.

Our approach improves upon on a method developed by Asari *et al.*, who calculate distortion parameters to fit a rectangular grid of dots to a set of adjacent lines [10]. In addition, we describe an apparatus for taking high-quality calibration images and provide a method to correct live video from an endoscopic system. Their method offers two advantages in that the correction coefficients are calculated in an automatic process and the calibration pattern does not need to be exactly vertical or horizontal. Thus, precise mechanical alignment is unnecessary. Their results are encouraging, but deficiencies remain. Based on their assumptions, each corrected grid line at the chosen direction (vertical in their paper) may have a different slope and produce line fits which are not parallel. For example, given a test pattern shown in Figure 1a, the effect is shown in Figure 1b. This effect gets more pronounced as the distortion center is further away from its proper location.

We propose a technique that finds solutions where all lines have the same slope, so that they are parallel. Unlike previous approaches, both horizontal and vertical line fits are necessary in order to find the optimal distortion center and proper overall solution. Our paper describes the method and provides results for test and real images. A more detailed report of this work has been submitted to a refereed journal for review [11].

2. PROBLEM STATEMENT AND METHOD

Figure 2 shows variables for distortion space and correction space. The goal is to derive a method for transforming the distorted image into a corrected image. The distortion center is (u_c', v_c') and correction center is (u_c, v_c). (x', y') represents the position of a pixel in the distortion image,



(a) (b)

Figure 2. Variables used to show the geometric dot positions. (a) Distortion space variables; (b) Corrected space variables.

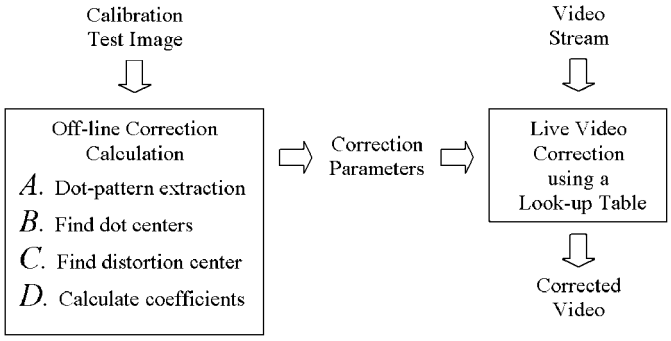


Figure 3. Complete approach for nonlinear endoscopic distortion correction. The off-line stage takes a calibration test pattern as input and generates the endoscopic correction parameters. Then, these parameters, implemented as a look-up table, are used for live endoscopic video image correction.

and (x, y) is a pixel in the corrected image. The dot centers in both distorted and correction space are defined as $(x_i^{'c}, y_i^{'c})$ and (x_i^c, y_i^c). The distance from distortion center to a pixel is ρ' for distorted space and ρ for the corrected image. The angles between the image pixels and distortion center are θ' for the distortion image and θ for the corrected image. These variables are calculated as follows:

$$\rho' = \sqrt{(x' - u_c')^2 + (y' - v_c')^2}, \qquad \theta' = \arctan\left(\frac{y' - v_c'}{x' - u_c'}\right).$$

The vector P in the corrected image space is (refer to Eq. (2), [10]):

$$\rho = \sqrt{(x - u_c)^2 + (y - v_c)^2}, \qquad \theta = \arctan\left(\frac{y - v_c}{x - u_c}\right).$$

Since only radial distortion is expected for an endoscope, the mapping from distortion space to correction space is

$$\rho = \sum_{n=1}^{N} a_n \rho'^n, \quad \theta = \theta'. \tag{1}$$

After correction of the image, the new pixel location in the corrected image space that maps the pixel in the distorted image can be obtained as (refer to Eq. (4), [10]):

$$x = u_c + \rho\cos\theta', \quad y = v_c + \rho\sin\theta'.$$

Least-squares line fitting is applied to the dot centroids of each grid line. Two steps are needed to calculate the dot centroids: (1) extraction of the distorted dot pattern and (2) finding the dot centroids by using a back-mapping method. Since the test dot pattern conforms to a strict rectangular grid, we assume a horizontal slope b_1^x for grid lines in the x direction and one vertical slope b_1^y for grid lines in the y direction of corrected space. Each grid line will have a different intercept value. The total error is from the sum of the distances from each dot to the corresponding grid line fit. Errors are calculated in both the x and y directions. A best estimate of the expansion coefficients and image center is obtained through an iterative calculation that minimizes the least absolute-value error. The following is a summary of the calibration procedure as diagrammed in Figure 3:

The steps below, described more fully in Section 3, constitute the off-line processing used to calculate the correction parameters and distortion center:

A. Extraction of distorted dot pattern.
B. Calculation of dot centroids in the corrected space and back-mapping to the distorted space.
C. Estimation of initial distortion center and corrected image center.
D. Iterative optimal calculation of expansion coefficients and image center.

After the correction parameters are computed, they can be used for real-time videoendoscopic image correction. The corrected images are calculated using a look-up table from correction pixel values to the input distortion pixels (see Figure 3). This step is detailed in Section 4.

3. EVALUATE CORRECTION PARAMETERS

The calibration image from the endoscope first needs to be converted to a binary image for dot processing. This is done by applying a low-pass filter on the original image and subtracting it from the original image. These image operations remove intensity variation across the image. This makes thresholding easier, since only one threshold value is necessary for the entire image. The final result is shown after thresholding and applying a median filter to remove salt and pepper noise.

A distortion center is first estimated using the same method as Asari *et al.* by interpolating the point where curvature is zero, based on calculations of curvature from column-wise and row-wise distorted lines [10]. This is only useful as an initial approximation and the optimal distortion centers are calculated by repeating the full parameter calculation over a rectangular search region of possible center values and saving the result with the minimum error.

The expansion coefficients are calculated by finding the $\{a_i\}$ coefficients of the polynomial transform that orients the centroids of grid dots to a straight line along x and y direction:

$$y_i = b_1^x x_{ij} + b_{0i}^x, \ 1 \le j \le K_i, 1 \le i \le L_x.$$

Here, K_i is the number of dots in row i and L_x is the number of rows. The variable F_i^x is defined similar to (refer to Eq. 7, [10]):

$$F_i^x = \sum_{j=1}^{K_i} (y_{ij} - b_1^x x_{ij} - b_{0i}^x)^2,$$

Similarly, the variables $F_j^y, b_1^y, \{b_{0j}^y\}$ denote error and line coefficients on the vertical direction. The unknowns, $b_1^x, \{b_{0i}^x\}, b_1^y, \{b_{0j}^y\}$, and (u_c, v_c) are chosen to minimize the function F:

$$F = \sum_{i=1}^{L_x} F_i^x + \sum_{j=1}^{L_y} F_j^y, \quad \frac{\partial F}{\partial b_{\alpha\beta}} = 0.$$

where $b_{\alpha\beta}$ represents all the b's. $\frac{\partial F}{\partial b_{\alpha\beta}}$ can be considered mathematically independent for x and y. Therefore, the minimization of F_x and F_y can be considered separately. The least-squares estimation discussed below focuses on the x direction. The results at the y direction will be easily obtained after that.

$b_1^x, \{b_{0i}^x\}$ can be calculated from:

$$\frac{\partial F_x}{\partial b_{\alpha\beta}} = \frac{\sum_{i=1}^{L_x}\sum_{j=1}^{K_i} \partial\left(y_{ij} - b_1^x x_{ij} - b_{0i}^x\right)^2}{\partial b_{\alpha\beta}} = 0.$$

The resulting equation for finding the line fit parameters through the horizontal dot centroids, assuming all lines have the same slope, is:

$$b_1^x = \frac{\sum_{ij} x_{ij}y_{ij} - \sum_{i=1}^{L_x}\left(\frac{1}{K_i^x}(\sum_{j=1}^{K_i^x} x_{ij})(\sum_{j=1}^{K_i^x} y_{ij})\right)}{\sum_{ij} x_{ij}^2 - \sum_{i=1}^{L_x}\left(\frac{1}{K_i^x}(\sum_{j=1}^{K_i^x} x_{ij})^2\right)},$$

$$b_{0i}^x = \frac{1}{K_i}\left(\sum_{j=1}^{K_i} y_{ij} - \left(\sum_{j=1}^{K_i} x_{ij}\right)\cdot b_1^x\right).$$

The error function E_x is defined as the normalized sum of the perpendicular distances from each dot centroid point to its corresponding fitting line:

$$E_x = \sum_{i=1}^{L_x}\frac{1}{K_i}\sum_{j=1}^{K_i}\left|\frac{y_{ij} - b_1^x x_{ij} - b_{0i}^x}{(1+(b_1^x)^2)^{1/2}}\right|.$$

E_x is a function of the expansion coefficients since x_{ij} and y_{ij} are a function of the corrected range ρ from (1). Since E_x is defined as an error function, E_x has a positive value which decreases as the distortion reduces. The main task now is to find a set of expansion coefficients that minimize the total error E_x (x direction discussed so far). A global minimization strategy is used in an iterative procedure to create new expansion coefficients. The iterative procedure calculates an error $E_x(\Delta)$ for each set of new coefficients, and from the mathematical model we apply the following recursive relationship:

$$a_n^x(\Delta+1) = a_n^x(\Delta) + \alpha n^\beta E_x(\Delta)\frac{1}{\left(\frac{\partial E_x(\Delta)}{\partial a_n}\right)},$$

for $n = 1, \ldots, N$ where α is the convergence rate parameter, β is the expansion index, and $\frac{\partial E_x(\Delta)}{\partial a_n}$ is the error gradient. The final coefficient will be the average of the optimized coefficient values from the different axes:

$$a_n = \frac{a_n^x + a_n^y}{2}, \quad n = 1, \ldots, N.$$

4. LIVE ENDOSCOPIC IMAGE CORRECTION

Given the set of correction parameters, live endoscopic video image correction can now be done (recall Figure 3). In our application, a look-up table maps points in the corrected space to points in distortion space to efficiently calculate corrected images. This is done as follows. Given pixel values (x, y), the distortion values are stored in two arrays:

$$x' = F_x(x,y) \quad y' = F_y(x,y).$$

This look-up table gives real-valued coordinates corresponding to pixel values in the distorted image. The associated intensity values can be calculated using either a bilinear transformation or, more simply, finding the nearest neighbor.

The look-up table is calculated from an inverse mapping of the polynomial in (1). The inverse mapping is defined as

$$\rho' = b_1\rho + b_2\rho^2 + ... + b_M\rho^M, \tag{2}$$

where $b_1, b_2, ..., b_M$ are the inverse polynomial coefficients and M is the number of coefficients. The elements of the look-up table are calculated as

$$F_x(x,y) = \left(\sum_{i=1}^{M} b_i\rho^i\right)\frac{x-u_x}{\rho}, \tag{3}$$

$$F_y(x,y) = \left(\sum_{i=1}^{M} b_i\rho^i\right)\frac{y-v_y}{\rho}, \tag{4}$$

with $\rho = \sqrt{(x-u_x)^2 + (y-u_y^2)}$. The coefficients of (2) are solved using least squares as follows. Substitute (1) into (2), so that:

$$\rho' = b_1[\sum_{n=1}^{N} a_n\rho'^n] + ... + b_M[\sum_{n=1}^{N} a_n\rho'^n]^M. \tag{5}$$

A least-squares equation is formed by subtracting the left side from the right side (5), squaring the difference, and integrating over the range of values of ρ'. Now, take the derivative with respect to b_i and set it equal to zero; This leads to the system of equations:

$$\begin{bmatrix} \int\rho'[\sum_{n=1}^{N} a_n\rho'^n]^1 d\rho' \\ \vdots \\ \int\rho'[\sum_{n=1}^{N} a_n\rho'^n]^M d\rho' \end{bmatrix} =$$

$$\begin{bmatrix} \int[\sum_{n=1}^{N} a_n\rho'^n]^2 d\rho' & \cdots & \int[\sum_{n=1}^{N} a_n\rho'^n]^{M+1} d\rho' \\ \vdots & \ddots & \vdots \\ \int[\sum_{n=1}^{N} a_n\rho'^n]^{M+1} d\rho' & \cdots & \int[\sum_{n=1}^{N} a_n\rho'^n]^{2M} d\rho' \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix}.$$

The limits of the integrals are $(0, r_{\max})$. Here, $r_{\max}$ is the maximum pixel distance of any screen point to the distortion center. The integral expressions are solved numerically using the trapezoidal rule and partitioning the range space over small increments. Solving the system of equations gives the inverse coefficients $b_1, \ldots, b_M$. These coefficients are used to calculate the look-up tables $F_x(x,y)$ and $F_y(x,y)$ by applying (3 - 4).

5. EXPERIMENTAL RESULTS

The distorted images were captured using two electronic videoendoscopy systems (Olympus CV-200). Figure 1a shows the test dot patterns used for the endoscopic images in Figure 4. Figure 4a shows the original input video image digitized directly from the endoscope. The sample image after dot extraction is shown in Figure 4b. A fifth-order polynomial fit was calculated for the correction parameters. The corrected test pattern image is shown in Figure 4c. The

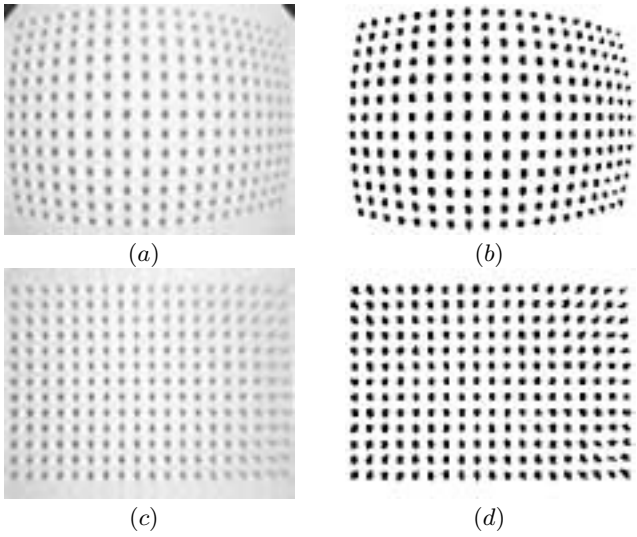


Figure 4. Complete results for a test pattern. (a) Original endoscopic video image; (b) Image after dot extraction; (c) Corrected image of test pattern; (d) Corrected dot extraction image.

resulting dot extraction image after correction is shown in Figure 4d. Notice that the edges of the distorted image appear too far away and the center of the image appears too close. The corrected image doesn't have these problems.

Figure 5 shows the results of distortion correction on an endoscopic image of a rubber bronchoscopic training device. This device is a rubber mold of the airways of the human chest. The device is used to train physicians in bronchoscopy procedures. Notice in Figure 5a that points in the outside of the image appear "too far" from the camera and points in the center appear "too close." Figure 5b shows the image in proper perspective for a more consistent perception of range. Figure 6 shows the distortion correction procedure as a necessary component of an image-guided system for registering an endoscope image to a rendered CT image and calculating the endoscope position [3]. This system will help physicians find suspect biopsy sites that can not be normally seen past the bronchial walls.

6. ACKNOWLEDGEMENTS

This work was partially supported by NIH grant # CA-74325 and by the Olympus Corporation.

REFERENCES

[1] K. Leggett, "Endoscopy: Many plusses and still a few minuses," *Biophotonics International*, vol. 6, no. 2, pp. 50–53, March/April 1999.

[2] K. R. Castleman, *Digital Image Processing.* Englewood Cliffs, NJ: Prentics Hall, 1996.

[3] A. J. Sherbondy, A. P. Kiraly, A. L. Austin, J. P. Helferty, S. Wan, J. Z. Turlington, E. A. Hoffman, G. McLennan, and W. E. Higgins, "Virtual bronchoscopic approach for combining 3D CT and endoscopic video," *SPIE Medical Imaging 2000: Physiology and Function from Multidimensional Images, A. Clough and C.T. Chen, eds.*, vol. 3978, pp. 104–116, Feb. 12-17, 2000.

[4] S. Shah and J. K. Aggarwal, "Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation," *Patt. Recogn.*, vol. 29, no. 11, pp. 1775–1788, Nov. 1996.

[5] W. E. Smith, N. Vakil, and S. A. Maislin, "Correction of distortion in endoscope images," *IEEE Transaction on Medical Imaging*, vol. 11, no. 1, pp. 117–122, March 1992.

[6] R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision," *Proc. IEEE Comput. Vision Patt. Recogn.*, June 1986.

[7] J. D. Stefansic, A. J. Herline, W. C. Chapman, and R. L. Galloway, Jr, "Endoscopic tracking for use in interative, image-guided surgery," *SPIE Medical Imaging: Image Display,* Y. Kim and S. Mun, Eds., vol. 3335, pp. 208–215, 1998.

[8] H. S. Sawhney and R. Kumar, "True multi-image alignment and its application to mosaicing and lens distortion correction," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 235–243, March 1999.

[9] I. A. Hein, M. D. Silver, and S. Oishi, "Distortion correction table compression for volume X-ray CT applications," *SPIE Medical Imaging 2000: Physics of Medical Imaging, J. T. Dobbins and J. M. Boone, eds.*, vol. 3977, February 12-17, 2000.

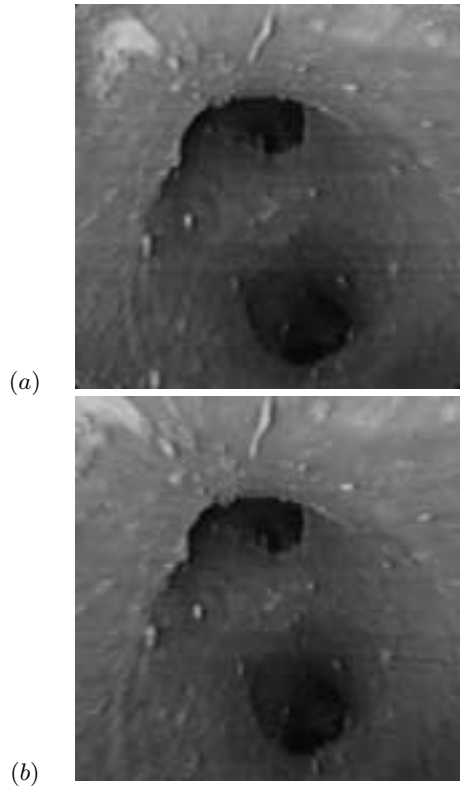


Figure 5. Results from a video frame captured with the endoscope inside a rubber bronchoscopic training device. (a) Original (distorted) endoscopic video image; (b) Corrected video image. The white streak on the top-left of image (a) appears too far away and the center of the bifurcation appears too close to the camera due to the distortion. Image (b) is in proper perspective since the distortion was corrected.

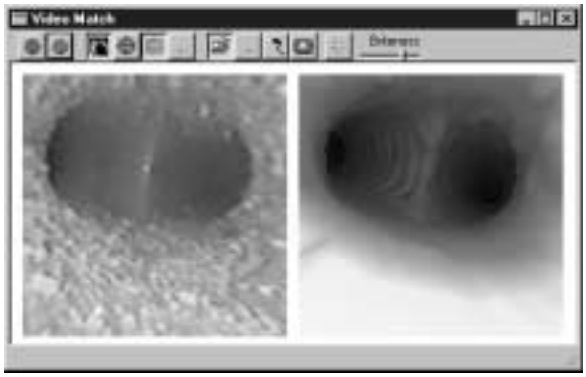


Figure 6. Corrected image of the rubber training device used in an image guided system. The distortion correction is necessary to match the endoscope image to a rendered view from a 3D CT Scan of the rubber device [3].

[10] K. V. Asari, S. Kumar, and D. Radhakrishnan, "A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation," *IEEE Transaction on Medical Imaging*, vol. 18, no. 4, pp. 345–354, April 1999.

[11] C. Zhang, J. P. Helferty, G. McLennan, and W. E. Higgins, "Nonlinear distortion correction in endoscopic video images," *submitted to IEEE Transactions on Medical Imaging*, 2000.