Videoendoscopic Distortion Correction and its Application to Virtual Guidance of Endoscopy

James P. Helferty

Chao Zhang

Geoffrey McLennan

William E. Higgins

(JPH), (CZ) and (WEH) Penn State University, Department of Electrical Engineering, University Park, PA 16802
(GM) University of Iowa, Iowa City, IA 52242, USA
(JPH) jph100@psu.edu, (CZ) zhang@aaron.ee.psu.edu, (GM) geoffrey-mclennan@uiowa.edu, (WEH) weh@aaron.ee.psu.edu

This work was partially supported by grant # CA-74325 from the National Cancer Institute of the NIH and by the Olympus Corporation. An unrefereed preliminary version of this work appeared [1].

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 virtual guidance of endoscopic procedures. 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 gives a globally consistent set of expansion coefficients and an optimal image center. 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.

Keywords

endoscopy, distortion correction, barrel distortion, camera calibration, image-guided endoscopy, virtual endoscopy

I. 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 [2]. 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 [3]. 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 [4–6], distortion correction will be necessary to accurately calculate an endoscope's position using the corrected video image registered to a rendered 3D CT image [7]. 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 [8]. 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 [3], 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 [8]. Smith *et al.* found distortion parameters using a manual technique where dot locations were chosen off the screen by hand [9]. Tsai proposed a second-order radial correction for 3D camera calibration using off-the-shelf TV cameras and lenses [10]. Stefansic *et al.* [11] used a second-order correction based on Smith *et al.* [9], 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 [12]. Hein *et al.* used a compressed look-up table to correct distortion from magnetic fields applied to an image intensifier for X-ray CT [13].

Our approach improves upon a method developed by Asari *et al.*, who calculate distortion parameters to fit a rectangular grid of dots to a set of adjacent lines [14]. Their seminal 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 corrected direction (vertical in their paper) may have different slopes and produce a fit where lines 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. In addition we describe an apparatus for taking high-quality calibration images, and provide a method to correct live video from an endoscopic system. Our paper describes the method and provides results for test and real images. Section II gives a mathematical definition of the problem and the experimental set-up for the tests. Section III gives the off-line steps needed to process an input calibration image and retrieve the correction parameters. Section IV describes the method used to display corrected images. Section V gives experimental results on a calibration dot pattern, a rubber bronchial training device, and human videobronchoscopic images. Distortion correction is also shown as a necessary component in a virtual endoscopy system that combines 3D CT images and endoscopic video.

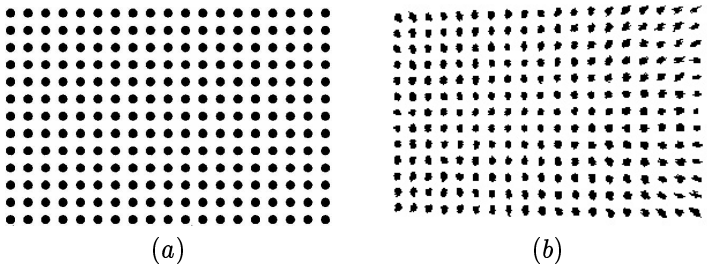


Fig. 1. Results from a previously proposed method [14]. (a) The original test pattern; (b) An imperfect trapezoidal-shaped result. The algorithm can provide somewhat inadequate results, since it does not require the line fits to be parallel. This effect becomes greater as the estimated distortion center gets farther from the actual center.

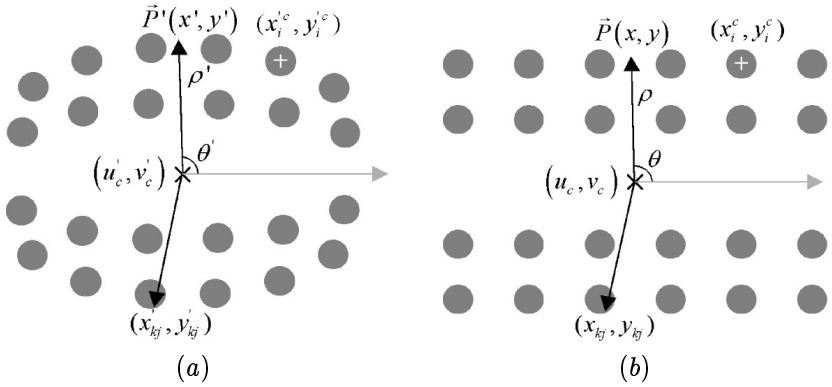


Fig. 2. Variables used to show the geometric dot positions. (a) Distortion space variables; (b) Corrected space variables.

II. Problem Statement and Method

In Section II-A, we mathematically define the problem. Section II-B gives an overview of the numerical methods used to solve the problem. Finally, Section II-C describes the physical set-up for capturing distorted calibration images using a videoendoscopic system.

A. Problem Statement

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, and (x, y) is a pixel in the corrected image. The dot centers in both distorted and corrected 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 distorted image and θ for the corrected

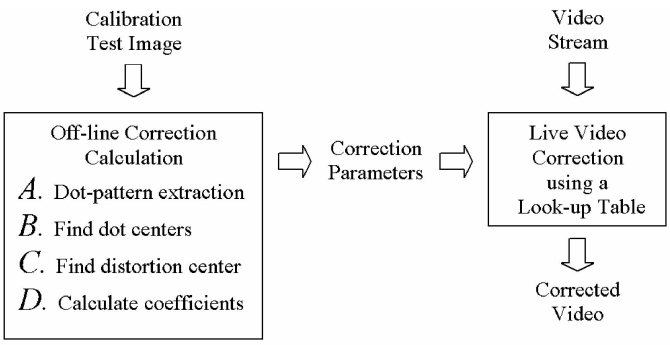


Fig. 3. Complete approach for nonlinear endoscopic distortion correction. The off-line stage takes a calibration test pattern as input and generates the endoscopic correction parameters. Then, these parameters, implemented as a look-up table, are used for live endoscopic video image correction.

image. These variables are calculated as follows:

$$\rho' = \sqrt{(x' - u'_c)^2 + (y' - v'_c)^2},\ \theta' = \arctan\left(\frac{y' - v'_c}{x' - u'_c}\right).$$

The vector P in the corrected image space is (refer to Eq. (2), [14]):

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

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), [14]):

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

The goal of the processing is to calculate the polynomial coefficients, $a_1, \ldots, a_n$, in (1), and the distortion center, (u'_c, v'_c). The correction centers (u_c, v_c) can be chosen by the user, since a change in the correction center position is only a simple translation of the final image.

B. Method

The basic experimental procedure, as suggested by Asari *et al.* [14], is to compute the correction parameters from an endoscopic video image of an image consisting of an equally-spaced array of black dots on a white background. Once this correction is computed, it can be applied to live endoscopic video in real time.

The following is a summary of our proposed calibration procedure and is illustrated in Figure 3. The steps

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

A. Extraction of the 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.

Least-squares line fitting is applied to the dot centers of each grid line. Two steps are needed to calculate the dot centers: (1) extraction of the distorted dot pattern; (2) finding the dot centers (centroids) by using a back-mapping method. Since the test dot pattern conforms to a strict rectangular grid, we assume a single slope b_1^x for grid lines in the x direction and 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. Optimal estimates of the expansion coefficients and image center are obtained through an iterative calculation that minimizes the least absolute-value error.

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 IV.

C. Materials and Preparation

In order to get high quality images and obtain external parameters for further image processing, we built a set of tools to help capture endoscopic calibration images. Figure 4 shows the optical platform for recording an image. The endoscope is placed inside a holding tube and fixed onto an optical stand. A test dot pattern is attached to a *x-y-z* translation stage. The relative position and direction of the stand and translation stage can be easily adjusted and finally screw-fixed to the aluminum base. The system is aligned, so that the endoscope sensor is perpendicular to the test dot plane. The *x-y* direction adjustment centers the test pattern into the viewing area. Experiments show that it is best to use a separate light source instead of the endoscope's built-in light pulse emitter to get uniform brightness over the entire image and to set the best contrast between the black dots and the white background.

The calibration dot pattern is a white picture consisting of a rectangular 13x19 array of black dots. Each dot is 0.1 inches (2.54 mm) in diameter and the distance between the dot centroids is 0.25 inches (6.35

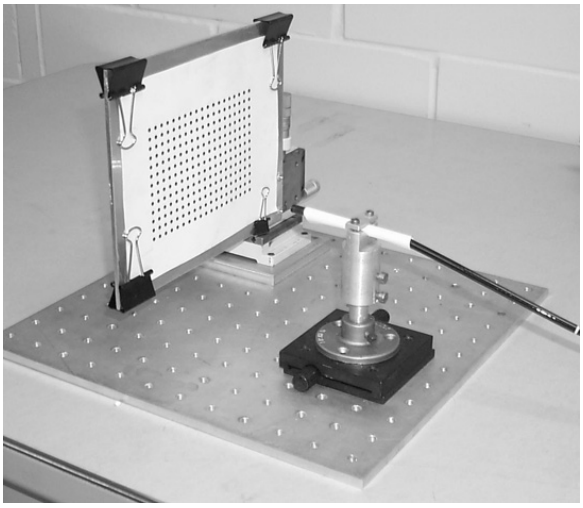


Fig. 4. Optical platform for calibration-pattern capture.

mm). The dot pattern portion of the captured video image was 260x200 pixels. When capturing a test video frame, we positioned the test dot pattern 3 inches (76.2 mm) away from the endoscopic lens. Three inches was a suitable distance, so that the dot pattern would fill the viewing area of the endoscope. Considerable flexibility is possible in this setup, as discussed in the results section.

III. Calculation of Correction Parameters

This section describes the 4-step off-line procedure for calculating the correction parameters (see Figure 3). Section III-A discusses extraction of the distorted dot pattern from a given calibration image. Section III-B overviews dot centroid estimation and correction using the back-mapping method. Section III-C focuses on the estimation of the distortion image center. Finally, Section III-D contains the mathematical details covering calculation of expansion coefficients and error estimation. Table I describes the variables used in the mathematical development appearing later in this section.

A. Dot-Pattern Extraction

The calibration image from the endoscope first needs to be converted to a binary image for dot processing. We emphasize that the calibration image was constructed to make the extraction of the dots straightforward (perfect black dots on a white background; lighting controlled to make sure it is uniform across the pattern). The primary imperfection exhibited by the calibration image is the distortion itself; but since the dots are small, this distortion has virtually no effect on the shape of the dots. Mild imperfections can arise from

TABLE I
DESCRIPTION OF VARIABLES USED TO EXPLAIN THE MATHEMATICAL CONCEPTS.

x_{ij}, y_{ij}	Coordinates of a distorted dot center
x'_{ij}, y'_{ij}	Coordinates of a corrected dot center
ρ_{ij}	Distance of distorted point to distortion center
ρ'_{ij}	Distance of corrected point to distortion center
$a_1, ..., a_n$	Coefficients of correcting polynomial for radial distortion component
u_c, v_c	Center of distortion. All lines through this point are linear in distortion space.
b^x_{0i}, b^x_1	Linear coefficients of a line fit to a set of horizontal dots.
b^y_{0i}, b^y_1	Linear coefficients of a line fit to a set of vertical dots.
K^x_i	Number of dots in horizontal line i of a pattern
K^y_i	Number of dots in vertical line i of a pattern
L_x	Number of horizontal lines in a pattern
L_y	Number of vertical lines in a pattern
F^x_i	Least squares error along horizontal line fit i of a pattern
F^y_i	Least squares error along vertical line fit i of a pattern
E_x	Least absolute error of horizontal line fits after distortion correction is applied
E_y	Least absolute error of vertical line fits after distortion correction is applied

video digitization and imperfect lighting, but the ideal contrast of the image limits these effects.

Since the original calibration dot pattern has basically ideal black-and-white contrast, it is a simple matter to robustly extract the series of dots. Figure 5 illustrates the sequence of steps. The extraction first involves applying a low-pass filter on the original image and subtracting it from the original image. This operation removes intensity variation across the image. Next, the image is thresholded. Simple thresholding works well, since one threshold value is sufficient for the entire image. Also, because of the large contrast between the dots and the background, a wide range of thresholds can be applied to give virtually identical results. The final result is obtained by applying a median filter to the thresholded image to remove any extraneous salt and pepper noise. Other approaches are, of course, possible for extracting the dots. But the ideal form of the image makes dot extraction very easy.

B. Dot Centroid Estimation

The next step of processing is to organize the pixels of the binary image into sets of dots. A "dot" is a region consisting of a collection of 4-connected pixels. The dots are stored in a data structure that holds: the array of pixel values constituting the dots, the center of each dot calculated by averaging the pixel positions, and the row and column index number of each dot on the grid.

The centroid (x^c_i, y^c_i) for dot i is calculated using the N_i pixels making up dot i:

$$x^c_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{ij}, \qquad y^c_i = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}\,. \tag{2}$$

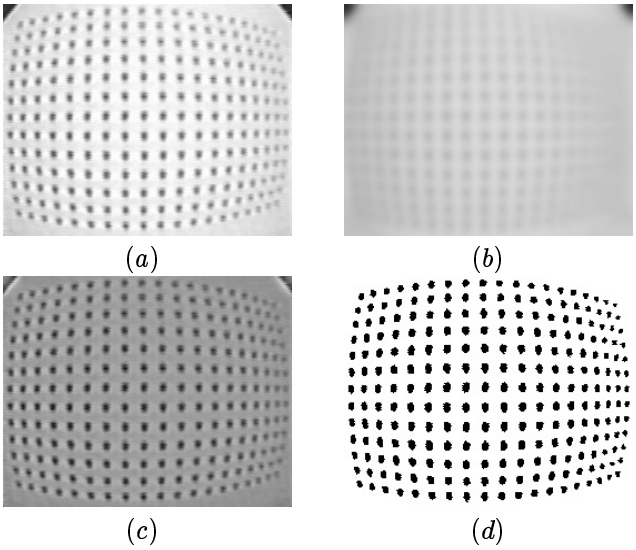


Fig. 5. Overview of dot-extraction method. (a) Original test image from endoscope, captured using the apparatus of Figure 4; (b) Image after a low-pass smoothing filter; (c) Image result from Subtracting image (b) from original image (a); (d) Final extracted binary dot pattern after simple thresholding and application of a median filter.

The image dots are in distortion space. So, the following transformation is necessary for each pixel in order to calculate a corrected dot:

$$x = u_c + \rho\cos\theta = u_c + \rho\cos\theta',$$

$$y = v_c + \rho\sin\theta = v_c + \rho\sin\theta',$$

or

$$x_{ij} = u_c + \left(\sum_{n=1}^{N} a_n \rho'^n\right) \cdot \left(\frac{x'_{ij} - u'_c}{\sqrt{\left(x'_{ij} - u'_c\right)^2 + \left(y'_{ij} - v'_c\right)^2}}\right),$$

$$y_{ij} = v_c + \left(\sum_{n=1}^{N} a_n \rho'^n\right) \cdot \left(\frac{y'_{ij} - v'_c}{\sqrt{\left(x'_{ij} - u'_c\right)^2 + \left(y'_{ij} - v'_c\right)^2}}\right).$$

After averaging, the dot centroids in the corrected space need to be mapped back to the distorted image space. This will require the inverse of the correction polynomial below:

$$\sum_{n=1}^{N} a_n (\rho_i^{c'})^n - \rho_i^c = 0,$$

A Newton-Raphson algorithm solves the nonlinear equation above for the unknown distortion range $\rho_i^{'c}$. The distortion dot centroids are calculated as:

$$x_i^{c'} = u'_c + \rho_i^{c'}\left(\frac{x_i^c - u_c}{\sqrt{\left(x_i^c - u_c\right)^2 + \left(y_i^c - v_c\right)^2}}\right),$$

$$y_i^{c'} = v_c' + \rho_i^{c'} \left(\frac{y_i^c - v_c}{\sqrt{\left(x_i^c - u_c\right)^2 + \left(y_i^c - v_c\right)^2}} \right).$$

Therefore, for any given set of expansion coefficients, centroids of extracted dots are averaged in the corrected image space and then back-mapped to the distorted image space. The back-mapping procedure is necessary, since dot centroids were initially calculated in distortion space and later were adjusted to the corrected space. These dot shapes will be warped and the corrected center will not be the true dot centroid. This is a problem, since the line fits are based on the centers in the corrected space. The back-mapping method applies a nonlinear transformation to ensure that the dot centroids are true in correction space.

C. Distortion Center Estimation

It is possible to calculate a curvature κ for each set of column and row dots as $\kappa = 2c_2$, where $y = c_0 + c_1 x + c_2 x^2$ is a quadratic line fit to the set of horizontal points. For vertical points, the equation becomes $x = c_0 + c_1 y + c_2 y^2$. For each quadratic, the maximum occurs for the x (or y) value that sets the derivative to zero. For a horizontal line, this value is $y_{max} = c_0 - c_1^2/4c_2$ and for a vertical line, $x_{max} = c_0 - c_1^2/4c_2$. The distortion center is first estimated using the same method as Asari *et al.* [14]. This method involves an interpolation of the maximum x and y values where the curvature would be zero.

The above procedure is only useful as an initial approximation. The actual optimal distortion-center is calculated by repeating the coefficient calculation over a rectangular search region of possible center values and saving the result with the minimum error. A deficiency in the method in [14] is that the distortion center calculation was often inaccurate and resulted in a corrected image that had trapezoidal distortion, as shown in Figure 1. The rectangular-region search finds the optimal distortion center and prevents this trapezoidal distortion.

D. Correction Polynomial Coefficient Estimation

Given a possible distortion center, the optimal coefficients are the values that map the distorted grid rows and columns into straight lines. A total error is calculated as

$$E(u_c', v_c') = E_x + E_y.$$

The horizontal row error, E_x, and the vertical column error, E_y, are calculated as

$$E_x = \frac{1}{L_x} \sum_{i=1}^{L_x} \frac{1}{K_i^x} \sum_{j=1}^{K_i^x} \left| \frac{x_{ij} - b_1^x y_{ij} - b_{0i}^x}{\left(1 + (b_1^x)^2\right)^{1/2}} \right|, \tag{3}$$

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

The variables are defined as follows. For horizontal line i, the intercept coefficient is b_{0i}^x. Each line has the same slope b_1^x. The intercept coefficients for vertical line i is b_{0i}^y and the slope value for all the vertical lines is b_1^y. Say a distortion dot image contains a set of L_x horizontal rows and L_y vertical columns. Each horizontal row contains K_i^x dots, and a dot centroid on the grid can be labeled (x'_{ij}, y'_{ij}), where $1 \le j \le K_i^x$ and $1 \le i \le L_x$. Each vertical column contains K_i^y dots, and a dot centroid on the grid can be labeled (x'_{ij}, y'_{ij}), where $1 \le j \le K_i^y$ and $1 \le i \le L_y$. Each error term will be minimized separately given an initial distortion center. The final coefficients will be the average of the optimized coefficient values from the two axes:

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

Each minimization problem is treated separately, with the same minimization procedure used for each. The only difference is that the x and y axes need to be switched, along with the number of lines and number dots per line.

The rest of this section describes how to minimize E_x to find the coefficients $a_1^x, \ldots, a_N^x$. The steps necessary to minimize E_y to find $a_1^y, \ldots, a_N^y$ are the same with the axis changed.

An initial approximation for the coefficients is used. The iteration is used to update the coefficient:

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

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.

Given the line parameters b_1^x and $b_{01}^x, \ldots, b_{0L_x}^x$ in (3), the line equations are

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

The line parameters are calculated by taking derivatives of the above line equations with respect to the line parameters and setting them equal to zero:

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

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

$$b_1^x = \frac{\sum_{ij} x_{ij} y_{ij} - \sum_{i=1}^{L_x} \left(\frac{1}{K_i^x} (\sum_{j=1}^{K_i^x} x_{ij})(\sum_{j=1}^{K_i^x} y_{ij}) \right)}{\sum_{ij} x_{ij}^2 - \sum_{i=1}^{L_x} \left(\frac{1}{K_i^x} (\sum_{j=1}^{K_i^x} x_{ij})^2 \right)}, \tag{4}$$

$$b_{0i}^x = \frac{1}{K_i^x} \left(\sum_{j=1}^{K_i^x} y_{ij} - b_1^x (\sum_{j=1}^{K_i^x} x_{ij}) \right). \tag{5}$$

The derivation of these line coefficients is shown in the Appendix.

The expression of the error gradient is calculated as:

$$\frac{\partial E_x}{\partial a_n} = \frac{1}{L_x} \sum_{i=1}^{L_x} \left(\frac{1}{K_i^x (1 + (b_1^x)^2)^{3/2}} \cdot \sum_{j=1}^{K_i^x} \left| (1 + (b_1^x)^2) \left(\frac{\partial b_{0i}^x}{\partial a_n} + b_1^x \frac{\partial x_{ij}}{\partial a_n} + x_{ij} \frac{\partial b_1^x}{\partial a_n} \right) - (b_{0i}^x + b_1^x x_{ij} - y_{ij}) \left(b_1^x \frac{\partial b_1^x}{\partial a_n} \right) \right| \right).$$

The derivatives $\frac{\partial b_1^x}{\partial a_n}$ and $\frac{\partial b_{01}^x}{\partial a_n}, \ldots, \frac{\partial b_{0L_x}^x}{\partial a_n}$ are given by:

$$\frac{\partial b_1^x}{\partial a_n} = \frac{1}{\sigma_1^2} \cdot \left\{ -2\sigma_2\sigma_3 + \sigma_1 \left[\sum_{ij} \left(\frac{\partial x_{ij}}{\partial a_n} y_{ij} + \frac{\partial y_{ij}}{\partial a_n} x_{ij} \right) - \sum_{i=1}^{L_x} \frac{1}{K_i^x} \left(\sum_{j=1}^{K_i^x} \frac{\partial x_{ij}}{\partial a_n} \cdot \sum_{j=1}^{K_i^x} y_{ij} + \sum_{j=1}^{K_i^x} \frac{\partial y_{ij}}{\partial a_n} \cdot \sum_{j=1}^{K_i^x} x_{ij} \right) \right] \right\},$$

where

$$\begin{aligned}
\sigma_1 &= \sum_{ij} x_{ij}^2 - \sum_{i=1}^{L_x} \frac{1}{K_i^x} (\sum_{j=1}^{K_i^x} x_{ij})^2 \\
\sigma_2 &= \sum_{ij} x_{ij} y_{ij} - \sum_{i=1}^{L_x} \frac{1}{K_i^x} (\sum_{j=1}^{K_i^x} x_{ij})(\sum_{j=1}^{K_i^x} y_{ij}) \\
\sigma_3 &= \sum_{ij} x_{ij} \frac{\partial x_{ij}}{\partial a_n} - \sum_{i=1}^{L_x} \frac{1}{K_i^x} (\sum_{j=1}^{K_i^x} x_{ij})(\sum_{j=1}^{K_i^x} \frac{\partial x_{ij}}{\partial a_n}).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial b_{0i}^x}{\partial a_n} &= \frac{1}{K_i^x} \left[\sum_{j=1}^{K_i^x} \frac{\partial y_{ij}}{\partial a_n} - b_1^x \sum_{j=1}^{K_i^x} \frac{\partial x_{ij}}{\partial a_n} - \frac{\partial b_1^x}{\partial a_n} \sum_{j=1}^{K_i^x} x_{ij} \right], \\
&\quad \text{where } 1 \le i \le L_x.
\end{aligned}$$

(x_{ij}, y_{ij}) is calculated from (2), while $\frac{\partial x_{ij}}{\partial a_n}$ and $\frac{\partial y_{ij}}{\partial a_n}$ are obtained from Eq. (19), [14]:

$$\frac{dx_{ij}}{da_n} = {\rho'}_{ij}^n \cos\theta'_{ij}, \quad \frac{dy_{ij}}{da_n} = {\rho'}_{ij}^n \sin\theta'_{ij}.$$

IV. Live Endoscopic Image Correction

The complete method from initial off-line correction calculation to live endoscopic video correction, as summarized in Figure 3, is straightforward to use in a clinical setting. All of our results use either clinical

equipment or standard computer equipment. The apparatus for capturing the initial calibration pattern (Figure 4) consists of standard optical components and costs less than $100 to assemble. The computation of the calibration function is done on an inexpensive PC and takes only seconds to compute. The only requirement on the PC is that it be equipped with a video capture board, a common PCI-based board available from many manufacturers. With the video capture board, the PC can accept live bronchoscopic video directly and perform the distortion correction at real-time frame rates, with no processing delay. We have applied this method to all of our tests for virtual guidance of bronchoscopy. These results have all used clinically ready bronchoscopy equipment and have spanned the range of cases from phantoms, animals, and humans [15–18].

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:

$$\rho' = a'_1\rho + a'_2\rho^2 + ... + a'_M\rho^M, \tag{6}$$

where $a'_1, a'_2, ..., a'_M$ are the inverse polynomial coefficients and M is the number of coefficients. The elements of the look-up table are calculated as

$$F_x(x, y) = \left(\sum_{i=1}^{M} a'_i\rho^i\right)\frac{x - u_x}{\rho}, \tag{7}$$

$$F_y(x, y) = \left(\sum_{i=1}^{M} a'_i\rho^i\right)\frac{y - v_y}{\rho}, \tag{8}$$

with $\rho = \sqrt{(x - u_x)^2 + (y - u_y^2)}$. The coefficients of (6) are solved using least squares as follows. Substitute

(1) into (6), so that:

$$\rho' = a'_1[\sum_{n=1}^{N} a_n \rho'^n] + ... + a'_M[\sum_{n=1}^{N} a_n \rho'^n]^M. \tag{9}$$

A least-squares equation is formed by subtracting the left side from the right side of (9), squaring the difference, and integrating over the range of values of ρ'. This gives

$$\varepsilon = \int_0^{r_{\max}} \left(\rho' - a'_1 \sum_{n=1}^{N} a_n \rho'^n - ... - a'_M (\sum_{n=1}^{N} a_n \rho'^n)^M \right)^2 d\rho'. \tag{10}$$

Here, $r_{\max}$ is the maximum pixel distance of any screen point to the distortion center. Now, take the derivative (10) with respect to a'_i and set it equal to zero;

$$0 = \frac{d\varepsilon}{da'_i} = \int_0^{r_{\max}} 2(\sum_{n=1}^{N} a_n \rho'^n)^i \left(\rho' - a'_1 \sum_{n=1}^{N} a_n \rho'^n - ... - a'_M (\sum_{n=1}^{N} a_n \rho'^n)^M \right) d\rho'.$$

This leads to the system of equations:

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

$$\begin{bmatrix} \int [\sum_{n=1}^{N} a_n \rho'^n]^2 d\rho' & \cdots & \int [\sum_{n=1}^{N} a_n \rho'^n]^{M+1} d\rho' \\ \vdots & \ddots & \vdots \\ \int [\sum_{n=1}^{N} a_n \rho'^n]^{M+1} d\rho' & \cdots & \int [\sum_{n=1}^{N} a_n \rho'^n]^{2M} d\rho' \end{bmatrix} \begin{bmatrix} a'_1 \\ \vdots \\ a'_M \end{bmatrix}.$$

The limits of the integrals are $(0, r_{\max})$.

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 $a'_1, \ldots, a'_M$. These coefficients are used to calculate the look-up tables $F_x(x, y)$ and $F_y(x, y)$ by applying (7 - 8).

V. Experimental results

We collected test data for two different Olympus CV-200 videoendoscopy systems. One endoscope was used to capture a video sequence from a rubber bronchoscopy training device [7]. Another endoscope was used for capturing a video sequence from a human subject. Prior to collecting the video, an image of the test dot pattern was captured for each endoscope.

All images were captured using a Matrox II frame grabber on a Dell Precision 410 Workstation (a 400 MHz dual Pentium II) with the Windows NT 4.0 Operating System. Image capture and dot-pattern extraction were done using the Matrox Inspector software package. Software for the remaining off-line parameter

calculations and display processing was developed using Microsoft Visual C++ 5.0.

The first set of results gives an error analysis on the test dot pattern, comparing the proposed correction procedure with the Asari *et al.* procedure [14]. The second part of the results section presents live distortion-correction applied to human bronchoscopic images. In the final part of this section, the utility of the procedure is shown in an image-guided virtual bronchoscopic system.

A. Test Pattern Analysis

Figure 6 gives a complete series of results for the test dot pattern; Figure 1a is the original undistorted pattern before endoscopic capture. Figure 6a shows the original input video image digitized directly from the endoscope. The sample image after dot-pattern extraction is shown in Figure 6b. A fifth-order polynomial fit was calculated for the correction parameters using the results of Section III. The corrected test pattern image using our proposed method is shown in Figure 6c. Figure 6d shows the resulting binary image after correction. Notice that the edges of the original distorted image appear too far away and the center of the image appears too close, while the corrected image fixes these difficulties. Figures 6e and 6f show the results of a distortion correction using the Asari *et al.* method in [14]. This method only considers vertical line fits and does not force the lines to be parallel. This causes the "trapezoidal" solution that appears in the result. Our approach forces all the line fits to have the same slope leading to the "rectangular" solution as shown in Figures 6c and 6d.

Table II gives a numerical comparison of the deviation errors in the dot pattern for the approach in this paper and the Asari *et al.* method [14]. Each optimization procedure (Asari *et al.* algorithm and our algorithm) was iterated until it found a minimum and stopped when the next iteration produced a larger error. In order to find the best distortion center, both algorithms iterated over a 20x20 window of possible distortion center values to find the value that gave the minimum error. The Asari *et al.* method calculated the dot's distance from the fitted vertical line with each line having an independent slope. Our method calculated the error from summing the dot's distance from both the vertically fitted line and the horizontally fitted line with parallel lines having the same slope. The optimizations resulted in a set of coefficients. In order to normalize the errors, the final coefficients were scaled, so that the resulting corrected dot patterns would have the same perimeter. The final corrected images each had a perimeter of 1000 pixels.

The errors in the Table II are given in pixel units based on this normalizing procedure. The errors presented in the table are defined as follows:

$\sqrt{x^2+y^2}$ error - Euclidean distance of a corrected dot center's location from its known actual location.

$|x|$ error - Absolute value of the horizontal (x) direction error of a corrected dot center's location from its actual location.

$|y|$ error - similar to $|x|$ error, but vertical (y) direction considered.

For these errors, the average, the standard deviation, and maximum values were computed over the entire test pattern.

As table II shows, our proposed method reduces the overall dot placement error ($\sqrt{x^2+y^2}$) by nearly 50%. Further, the maximum (worst) dot placement error is reduced by nearly 40%. The Asari *et al.* approach gives slightly better x-direction performance, but their method focuses only on this direction. Our method gives a factor of three reduction on vertical (y) errors. In all cases, our approach gives the smallest maximum error.

We point out that the endoscopic image is rectangular. Thus, the original distortion errors are greatest in the horizontal direction. This accounts for the differences in the individual x- and y- direction performances. The placement of the endoscope during test-pattern capture has a minor impact. We placed the endoscope, so that the captured image would fill the field of view and not be blurred. This makes dot-pattern extraction more precise, leading to overall more accurate distortion-correction calculation. The 3 inch (76.2 mm) placement of the endoscope from the test pattern roughly corresponds to the endoscope's intrinsic focal length. Regarding the dot pattern, the dots have to be large enough and sufficiently spread apart to permit successful extraction. Certainly, considerable variation is acceptable in this pattern as long as there is sufficient information (known dot centers) that span the field of view of the endoscope to run the optimization. Lastly, we point out that our approach performs distortion correction on endoscopic video at real-time rates. More specifically, we corrected 228x228 color images at 30 frames per second.

B. Videobronchoscopy Correction

Figure 7 shows the results of distortion correction on images captured from a rubber bronchoscopy 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. Figure 7 shows views of the left main bronchus subdivision. Notice in Figure 7a that points in the outside of the image are squeezed together and points in the center are spread apart. This leads to a false distance perception with the outside area appearing too far from the camera and the center image area appearing too close. Figure 7b shows the image in proper perspective for a more consistent perception of range.

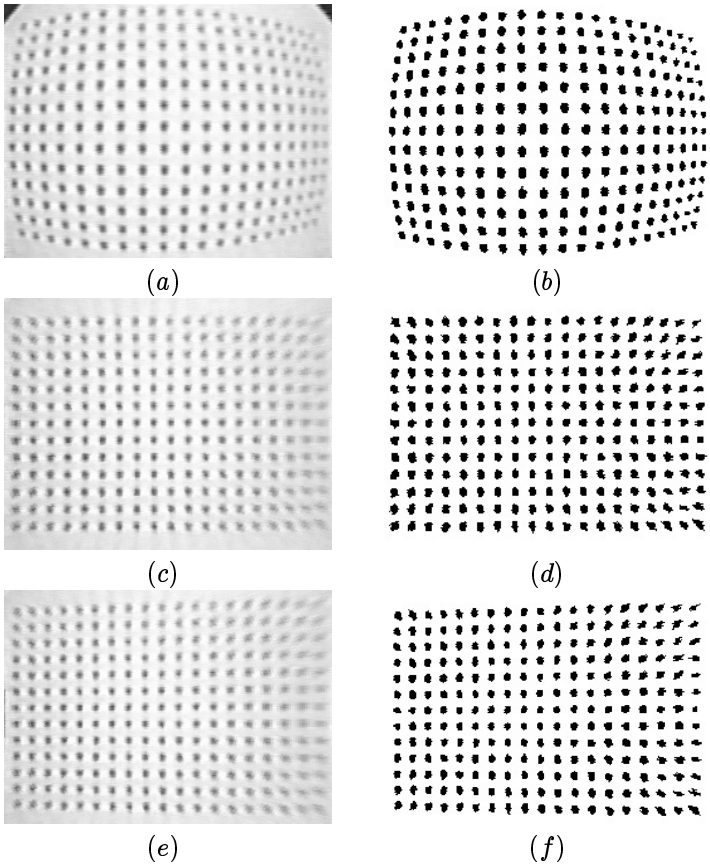
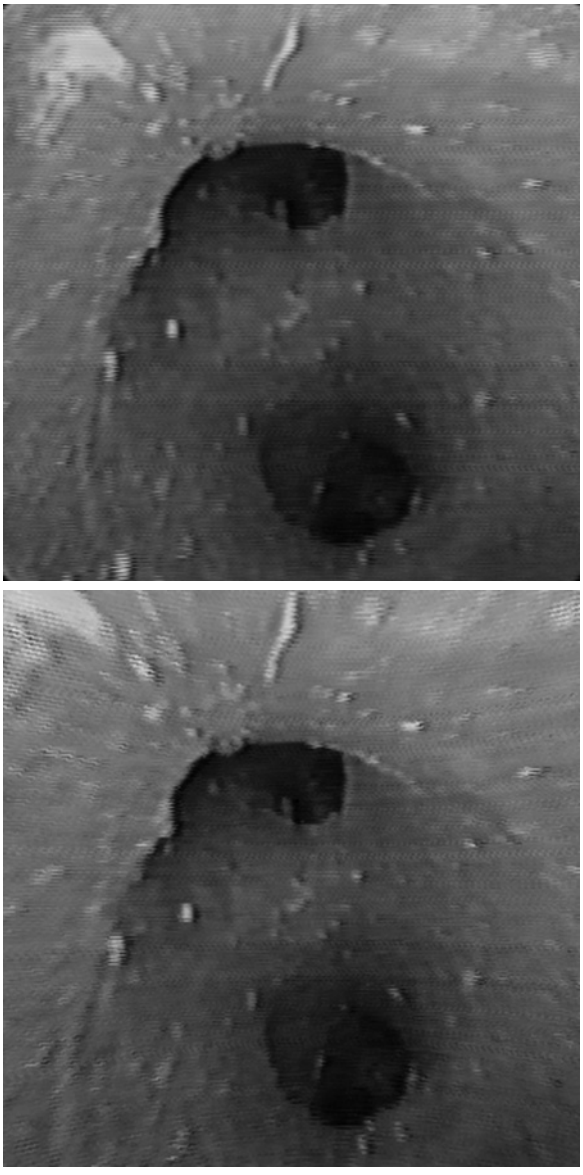


Fig. 6. Complete results for a test pattern. (a) Original endoscopic video image; (b) Image after dot-pattern extraction; (c) Corrected image of test pattern using our proposed method; (d) Corrected dot-pattern image using our proposed method; (e) Corrected image of test pattern using Asari *et al.* method; (f) Corrected dot-pattern image using Asari *et al.* method.

TABLE II

Error analysis for test pattern correction. Errors are given in terms of normalized pixel units.

Error Measure	Proposed	Asari *et al.* [14]
Average $\sqrt{x^2+y^2}$ error	0.038298	0.072920
Standard Deviation $\sqrt{x^2+y^2}$ error	0.000678	0.003095
Maximum $\sqrt{x^2+y^2}$ error	0.162132	0.261759
Average $\|x\|$ error	0.029547	0.021487
Standard Deviation $\|x\|$ error	0.000678	0.000550
Maximum $\|x\|$ error	0.162079	0.186303
Average $\|y\|$ error	0.017838	0.063572
Standard Deviation $\|y\|$ error	0.000276	0.003363
Maximum $\|y\|$ error	0.102137	0.222951



(a)

(b)

Fig. 7. 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 is corrected.

Figure 8 shows results for human videobronchoscopic images. This test used a different endoscope from the previous example. The dot pattern calibration procedure was repeated and a separate set of parameters was applied to correct the image. Notice the areas in the center of the image appear closer in the distorted image than in the corrected image, and areas on the outer edge of the distorted image seem further away. This effect can be seen in both the carina view in Figure 8a, and the bifurcation view in Figure 8c. Figures 8b and 8d show the corrected images with the distortion removed. Figure 8e is a bifurcation image with grid lines to

better show the effect of distortion. Figure 8f is the corrected bifurcation image with grid lines showing the distortion removed. Also, comparing Figures 8c-d, the airway wall in the lower right appears significantly wider in the undistorted image (d) than in the distorted image (c).

C. Virtual Guidance of Bronchoscopy

An emerging application in the field of virtual endoscopy is to use a virtual graphics-based system to guide endoscopic procedures [7]. The endoscope cannot give information beyond the walls of hollow passages. Virtual data, be it extraluminal slice images or other forms of rendering, can provide such data. If the virtual world and real endoscopic world can be registered during a live procedure, then the endoscopist's "vision" can be greatly augmented, resulting in more successful needle biopsies and other procedures. Figure 9 presents an example from our recent ongoing work in virtual bronchoscopy [7]. A 3D CT scan of the bronchoscopy training device was first made using an Imatron EBCT scanner. Then, during live operation, a bronchoscope was inserted into the device, giving video data. Figure 9 shows an endoluminal rendering of the 3D CT data matched to a corresponding video frame; reference [7] discusses this matching procedure. The video distortion needs to be removed before matching, so that the video pixels will correctly correspond to the same points in the rendered image. Also, the distance of the endoscope to the test pattern (See Figure 4) is needed for proper calibration of the endoscope and virtual system's fields of view. In Figure 9, a view of the carina is matched, displaying the CT location as a cross-hair on a Coronal (x-z) CT slice. With such a match done between the virtual (CT) world and the live videobronchoscopy world, other visualization tools, such as local cross-sectional images, sliding thin-slab images, and straightened "tube" views, can now conceivably be used for guiding bronchoscopy procedures [4, 7].

VI. Discussion

This paper described a complete procedure for correcting barrel distortion in videoendoscopic images. The results show that this procedure improves upon the performance of an existing method. Complete details were given for the correction calculation, calibration apparatus, and live videoendoscopy correction. The correction procedure runs at interactive video frame rates on a standard PC. The method was demonstrated on a variety of images. The corrected images show how the corrected procedure puts the imaged scene in proper depth perspective and leads to a better perception of range. With the corrected videoendoscopy images, more accurate measurements of stenotic airways, for example, could conceivably be done, permitting

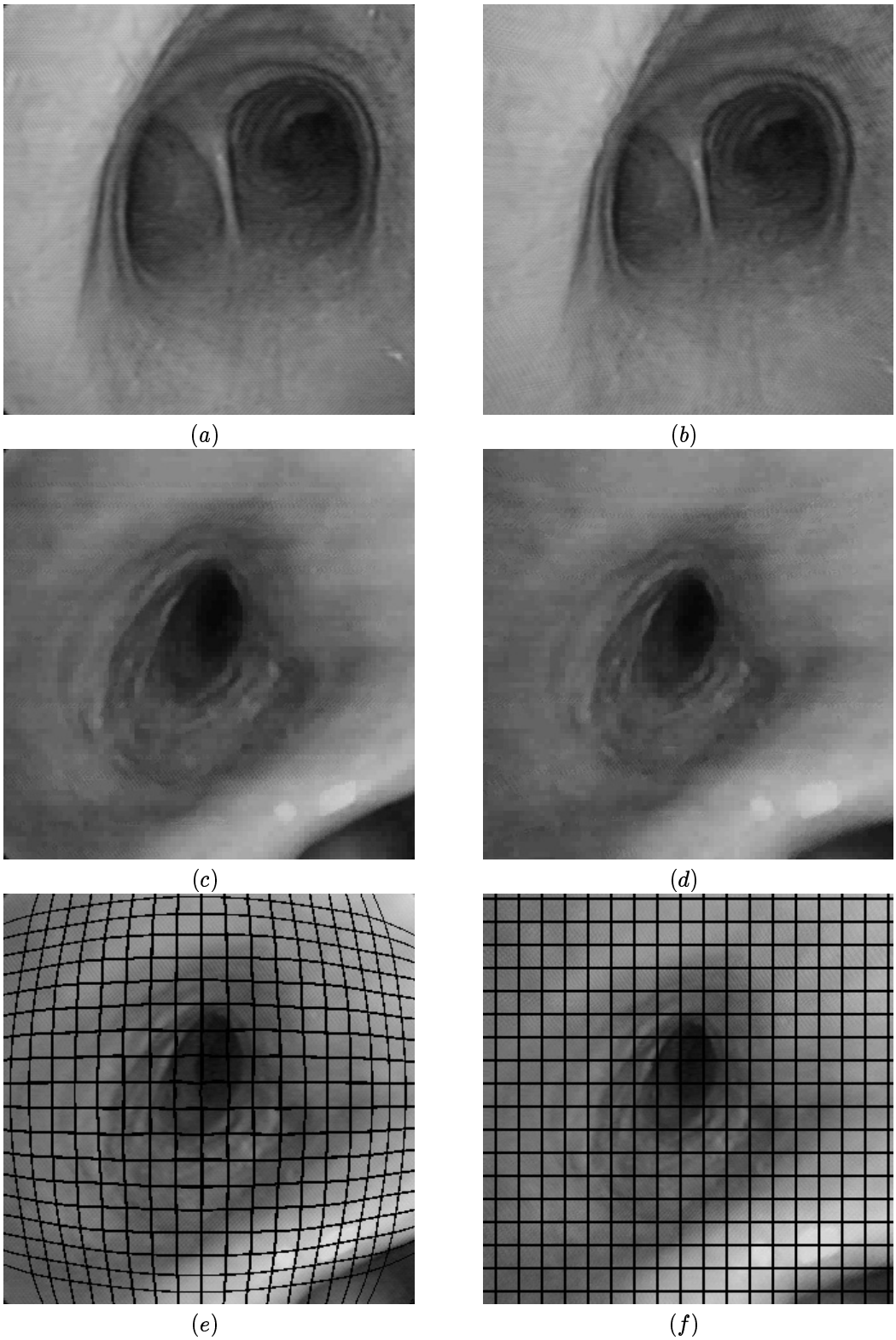


Fig. 8. Results from two video frames captured with an endoscope inside a human subject. (a) Original (distorted) videobronchoscopic image of human carina; (b) Corrected video image of carina; (c) Original (distorted) videobronchoscopic image of a bifurcation; (d) Corrected video image of bifurcation. (e) Original (distorted) videobronchoscopic image with grid lines to show the effect of distortion; (f) Corrected video image with grid lines.

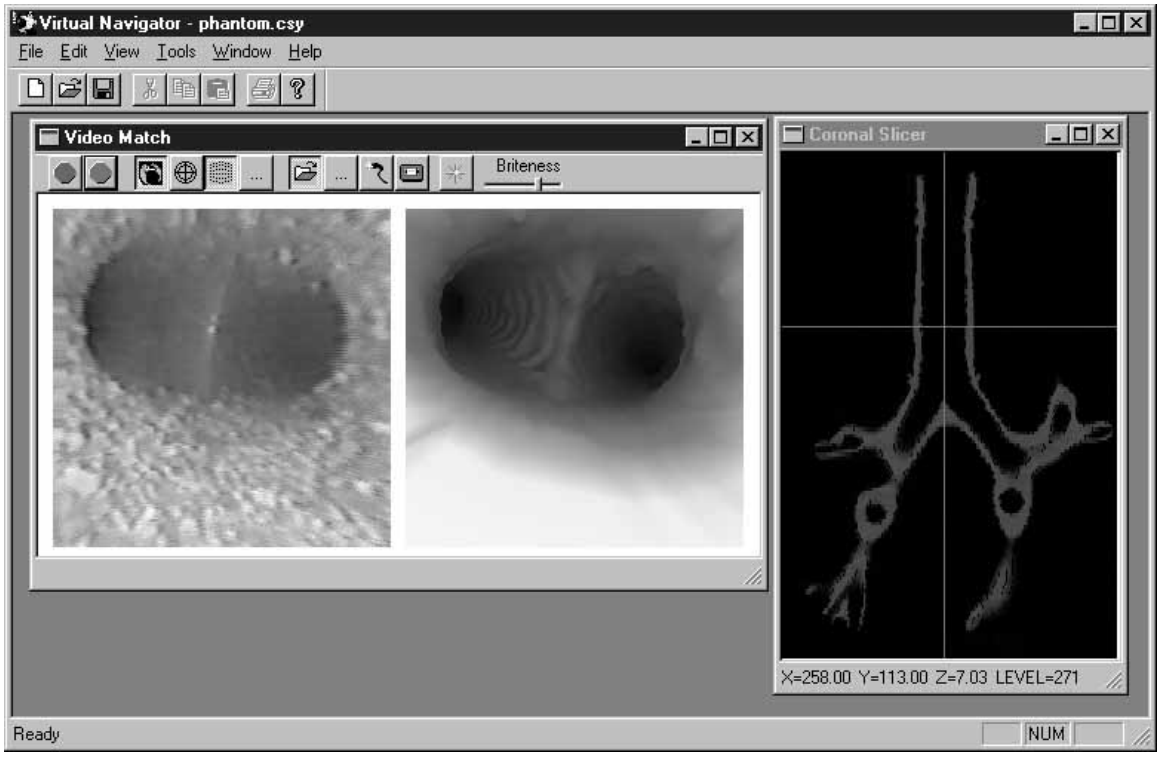


Fig. 9. Example from matching a rendered endoluminal CT view to a corresponding video shot for the rubber bronchoscopy training device. The complete software package runs on windows-based PC's and performs virtual (CT) and real world matching [7]. In the "Video Match" frame, the left view is a video frame near the carina. The right view is the corresponding matching rendered CT view. The view labeled "Coronal Slicer" illustrates the x-y-z position of the endoscope and virtual world on a 2D coronal (x-z) CT slice. The (x, y, z) dimensions of the CT scan were 450 x 154 x 160 voxels, with voxel sample intervals given by $\Delta x = \Delta y = 0.35mm$ and $\Delta z = 1.5mm$.

more accurate stent design. Finally, the necessity of the video correction was illustrated for an emerging application involving the merging of virtual (CT-based) endoscopy and videobronchoscopy.

References

[1] C. Zhang, J. P. Helferty, G. McLennan, and W. E. Higgins, "Nonlinear distortion correction in endoscopic video images," *IEEE ICIP-2000*, vol. 34, Sept. 10-13, 2000.

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

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

[4] W. E. Higgins, K. Ramaswamy, R. Swift, G. McLennan, and E. A. Hoffman, "Virtual bronchoscopy for 3D pulmonary image assessment: State of the art and future needs," *Radiographics*, vol. 18, no. 3, pp. 761–778, May-June 1998.

[5] G. D. Rubin, C. F. Beaulieu, V. Argiro, H. Ringl, A. M. Norbash, J. F. Feller, M. D. Dake, S. Napel, R. B. Jeffrey, and S. Napel, "Perspective volume rendering of CT and MR images: applications for endoscopic imaging," *Radiology*, vol. 199, no. 2, pp. 321–330, May 1996.

[6] D. J. Vining, K. Liu, R. H. Choplin, and E. F. Haponik, "Virtual bronchoscopy: relationships of virtual reality endobronchial simulations to actual bronchoscopic findings," *Chest*, vol. 109, no. 2, pp. 549–553, Feb. 1996.

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.
[12] 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.
[13] 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.
[14] 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.
[15] W. E. Higgins, A. J. Sherbondy, J. P. Helferty, A. P. Kiraly, G. McLennan, E. A. Hoffman, and J. Z. Turlington, "Virtual bronchoscopy for 3D CT assessment and endoscopic guidance," *Radiology*, vol. 217(P), p. 706, Nov. 2000.
[16] J. P. Helferty, A. J. Sherbondy, E. A. Hoffman, G. McLennan, and W. E. Higgins, "Experiments in virtual-endoscopy guidance of bronchoscopy," *SPIE Medical Imaging 2001: Physiology and Function from Multidimensional Images, C. Chen and A. V. Clough, eds.*, vol. 4321, Feb 18-22 2001.
[17] W. E. Higgins, J. P. Helferty, E. A. Hoffman, A. P. Kiraly, A. F. Ross, and G. McLennan, "In vivo validation CT-guided bronchoscopic biopsy," *to appear, American Thoracic Society 2001*, May 18-23 2001.
[18] G. McLennan, J. P. Helferty, E. A. Hoffman, A. J. Sherbondy, J. S. Ferguson, and W. E. Higgins, "Virtual endoscopic guidance of bronchoscopy," *to appear, American Thoracic Society 2001*, May 18-23 2001.

Appendix

This appendix details the derivation of the least-squares line-fit parameters, as given in (4 - 5). Since the lines are parallel, they have the same slope and different intercept values. The line fits are necessary to determine how close the corrected dot patterns match parallel lines. The parameters are obtained by taking the derivative of the squared error between the corrected dot centroids and the line, and afterward, solving the differential equations for the fit parameters.

Solving the differential equation

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

with respect to the line parameters gives the system of linear equations gives

$$\begin{bmatrix} K_1 & \cdots & 0 & \sum_{j=1}^{K_1} x_{1j} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & K_{L_x} & \sum_{j=1}^{K_{L_x}} x_{L_x j} \\ \sum_{j=1}^{K_1} x_{1j} & \cdots & \sum_{j=1}^{K_{L_x}} x_{L_x j} & \sum_{ij} x_{ij}^2 \end{bmatrix} \begin{bmatrix} b_{01}^x \\ \vdots \\ b_{0L_x}^x \\ b_1^x \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{K_1} y_{1j} & \cdots & \sum_{j=1}^{K_{L_x}} y_{L_x j} & \sum_{ij} x_{ij} y_{ij} \end{bmatrix}^T.$$

Each of the first L_x equations in the system needs to be multiplied by $-(1/K_i^x) \sum_{j=1}^{K_i^x} x_{ij}$ with $1 \leq i \leq L_x$.

This gives the matrix equation:

$$\begin{bmatrix} -\sum_{j=1}^{K_1^x} x_{1j} & \cdots & 0 & -\frac{1}{K_1^x}(\sum_{j=1}^{K_1^x} x_{1j})^2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -\sum_{j=1}^{K_{L_x}^x} x_{L_x j} & -\frac{1}{K_{L_x}^x}(\sum_{j=1}^{K_{L_x}^x} x_{L_x j})^2 \\ \sum_{j=1}^{K_1^x} x_{1j} & \cdots & \sum_{j=1}^{K_{L_x}^x} x_{L_x j} & \sum_{ij} x_{ij}^2 \end{bmatrix} \begin{bmatrix} b_{01}^x \\ \vdots \\ b_{0L_x}^x \\ b_1^x \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{K_1^x}\sum_{j=1}^{K_1^x} y_{1j} \sum_{j=1}^{K_1^x} x_{1j} \\ \vdots \\ -\frac{1}{K_{L_x}^x}\sum_{j=1}^{K_{L_x}^x} y_{L_x j} \sum_{j=1}^{K_{L_x}^x} x_{L_x j} \\ \sum_{ij} x_{ij} y_{ij} \end{bmatrix}.$$

Summing all of the equations in the system together and solving for b_1^x gives:

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

Using the above result, each of the first L_x equations can be solved for b_{0i}^x to get:

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