ELSEVIER

Computer-based system for the virtual-endoscopic guidance of bronchoscopy

J.P. Helferty [a], A.J. Sherbondy [b], A.P. Kiraly [c], W.E. Higgins [d],*

[a] *Lockheed-Martin Corporation, King of Prussia, PA, USA*
[b] *Department of Radiology, Stanford University, Stanford, CA, USA*
[c] *Siemens Corporate Research Center, Princeton, NJ, USA*
[d] *Department of Electrical Engineering, Penn State University, University Park, PA 16802, USA*

Abstract

The standard procedure for diagnosing lung cancer involves two stages: three-dimensional (3D) computed-tomography (CT) image assessment, followed by interventional bronchoscopy. In general, the physician has no link between the 3D CT image assessment results and the follow-on bronchoscopy. Thus, the physician essentially performs bronchoscopic biopsy of suspect cancer sites blindly. We have devised a computer-based system that greatly augments the physician's vision during bronchoscopy. The system uses techniques from computer graphics and computer vision to enable detailed 3D CT procedure planning and follow-on image-guided bronchoscopy. The procedure plan is directly linked to the bronchoscope procedure, through a live registration and fusion of the 3D CT data and bronchoscopic video. During a procedure, the system provides many visual tools, fused CT-video data, and quantitative distance measures; this gives the physician considerable visual feedback on how to maneuver the bronchoscope and where to insert the biopsy needle. Central to the system is a CT-video registration technique, based on normalized mutual information. Several sets of results verify the efficacy of the registration technique. In addition, we present a series of test results for the complete system for phantoms, animals, and human lung-cancer patients. The results indicate that not only is the variation in skill level between different physicians greatly reduced by the system over the standard procedure, but that biopsy effectiveness increases.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Virtual endoscopy; Image-guided surgery; 3D imaging; Image registration; Image fusion; Lung cancer; CT imaging; Bronchoscopy

1. Introduction

Lung cancer is the most common cause of cancer death in the United States, with roughly 170,000 new cases diagnosed each year [1]. It accounts for nearly 30% of all cancer deaths and has a five-year survival rate under 15%. The diagnosis of lung cancer occurs in two stages: (1) three-dimensional (3D) computed-tomography (CT) image assessment; and (2) bronchoscopy [2–4].

During Stage-1 3D CT Image Assessment, the physician manually "reads" a patient's 3D CT chest scan to identify and plan bronchoscopic biopsy. This reading is done by either examining a film series of the 3D image data on a view panel or by manually scrolling through the 3D image data on a computer console. Either way, the physician relies on experience and medical knowledge to mentally reconstruct the complex 3D anatomy. Manual reading has become especially impractical with the advent of modern multi-detector CT (MDCT) scanners, which typically produce several hundred submillimeter-resolution two-dimensional (2D) slice images per scan [5].

Next, during Stage-2 Bronchoscopy, the physician attempts to maneuver the bronchoscope through the airways to each preplanned biopsy site. The bronchoscope provides a real-time video stream of the airway interior

---

* Corresponding author. Fax: +1 814 863 5341.
   E-mail address: weh2@psu.edu (W.E. Higgins).

to assist in this maneuver. Unfortunately, the physician must make judgments based on the patient's anatomy depicted in the 3D CT image data; this is difficult as the CT data differs greatly in form from the bronchoscopic video. In addition, the physician must essentially perform the procedure blindly, since the target biopsy sites, be they lymph nodes or suspect cancer nodules, are not visible in the local airway video. Ancillary devices, such as fluoroscopy or CT fluoroscopy, are available, but these only provide limited projection or 2D thick-slice views [6]. Thus, physicians vary greatly in their skill level in bronchoscopy and the success rate of bronchoscopic biopsy tends to be very low [3,7,8].

We describe a computer-based system that improves the accuracy of bronchoscopy and reduces the skill-level variation between different physicians. The system enables detailed 3D CT-based procedure planning and follow-on image-guided bronchoscopy. During both the planning and bronchoscopy stages, the system greatly augments the physician's vision of the patient's anatomy. By using computer graphics and other computer-vision methods, far greater use is made of the 3D CT data during procedure planning. During bronchoscopy, the system gives direct image guidance by employing (a) image fusion of the 3D CT data and bronchoscopic video and (b) 3D navigation paths to the preplanned biopsy sites.

Our system has been partly motivated by recent efforts in image-guided surgery [9–14]. These systems add image-based guidance during surgery to improve the procedure success rate and enable difficult procedures. All of these systems require registration of preoperative medical imaging data and the live 3D physical surgical space. To facilitate registration, these systems often employ additional devices, such as fiducial markers attached to the patient's skin [10,12], markers attached to the surgical device [11–13], and either an optical or electromagnetic tracking system for measuring the marker positions [10,12,11,13,14]. Our system uses image fusion between the 3D CT image data and bronchoscopic video to perform registration between the preoperative image data and live 3D surgical space, similar to [9]. Hence, we do not require other devices.

Our system also has been partly motivated by the new field of virtual endoscopy, which has developed for more exhaustive 3D radiologic image assessment [15–17]. When applied to the chest, virtual endoscopy is usually referred to as virtual bronchoscopy (VB) [3,4,8,16,18–21]. In VB, a high-resolution 3D CT chest image serves as a "virtual environment" representing the chest anatomy. Endoluminal (interior) renderings of the airways, generated from computer processing of the CT data, act as views from a "virtual bronchoscope." In this way, unlimited exploration of the 3D anatomy can be made, with no risk to the patient.

Several recent efforts have drawn upon VB to assist bronchoscopy [3,4,8,22], but these efforts either: (a) did not offer direct image-guided bronchoscopy; (b) only gave information at airway-branch junctions; (c) required far too much computation to be usable during a live procedure; and/or (d) were not tested during live procedures. Notably, though, two of these efforts proposed methods that registered the 3D CT volume (the "*Virtual World*") to the bronchoscopic video (the "*Real World*") [8,22]—this enables potentially advantageous fusion of the two image sources without employing a supplemental guidance device. Our system uses fusion of preoperative image data with the 3D physical space.

Sections 2 and 3 of this paper describe our system and specific mathematical details. Section 4 gives detailed validation results for the CT-video registration method, which is pivotal to the system's functionality. Section 5 provides three sets of test results for the complete system, while Section 6 offers concluding comments.

2. System overview

We first overview the system. The system is used in two stages, per the standard lung-cancer assessment protocol, as illustrated in Fig. 1. It is integrated on a Windows-based PC (Dell Precision 620 workstation PC, dual-933MHz Pentium-III, 2GB RAM, Windows 2000). A Matrox Meteor-II frame grabber board is used for real-time video capture, while a GeForce4 Ti4600, with 128MB video memory on board, is used for the video card. The software is written
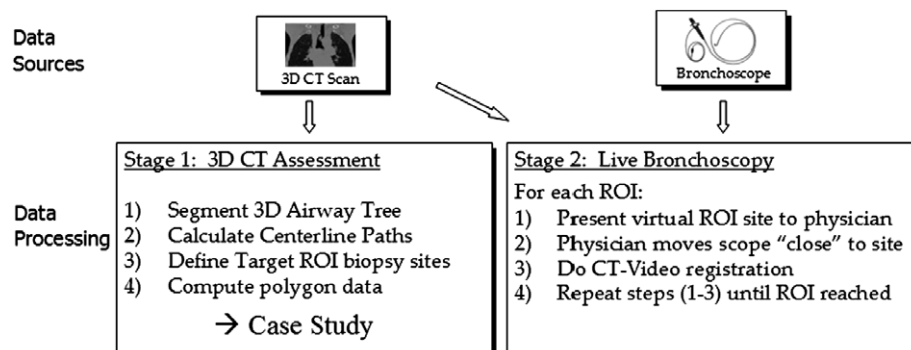
Fig. 1. Two-stage image-guided lung-cancer assessment system.

Fig. 2. System usage during a live procedure. The standard video monitor attached the bronchoscope suite (upper left) depicts video during a procedure, while the computer display (lower center) provides extra visual feedback.

in Visual C++ 6.0 and employs a few graphics and visualization utilities available in OpenGL and vtk [23,24]. This inexpensive computer set-up provides real-time computation of high-quality endoluminal renderings and real-time presentation of the bronchoscopic video. Fig. 2 depicts the system in the surgical suite. The physician observes both the computer display and standard fail-safe bronchoscope video monitor during the procedure.

The discussion below overviews the operation of the system. The operations for Stage 1 only appear as a summary below, since most of the data-processing methods of this stage have been described in previous publications. Section 3 gives details on many of the other data-processing steps. As many choices had to made during the construction of this large system, no individual method can be construed as being "optimal."

2.1. Stage 1: 3D CT-based planning

Given a patient's 3D CT image scan, the airway tree is first segmented using a 3D technique based on region growing and mathematical morphology [25]. Next, using the segmented airway tree as input, the major central axes of the airways are computed using a method combining techniques from 3D skeletonization, branch pruning, and cubic-spline analysis [26,27]. The techniques for segmentation and central-axes analysis were previously devised in our laboratory and heavily validated on many 3D human CT scans [25–27].

Next, two sets of triangles representing the interior (endoluminal) and exterior surfaces of the airway tree are generated. These data, necessary for generating 3D renderings of the airway tree, are computed as follows. First, a gray-scale voxel-based mask of the airway-tree surfaces is constructed by combining a simple $5 \times 5 \times 5$ dilation of the airway-tree segmentation with the original 3D gray-scale image data. Next, we apply the standard Marching Cubes algorithm, employed in computer graphics, to this masked gray-scale image to produce the requisite sets of triangles [24]. Note that the exterior surface is not merely a dilation of the interior surface. The dilation enables the creation of a liberally defined masked gray-scale image encompassing both the interior and exterior surfaces of the airway tree. By using this masked image, the Marching Cubes algorithm can finely compute exterior and interior surface triangle-mesh boundaries to the sub-voxel level.

Finally, the physician interacts with the system's computer display to define the target biopsy sites. The physician does this by either manually drawing regions of interest (ROIs) on 2D slice views of the 3D CT data or by employing a semi-automatic image-segmentation method. Once a site is defined, a triangle representation suitable for later rendering is derived for it and a guidance path is selected by locating the closest precomputed central axis to the site. When all target biopsy sites have been defined, a guidance plan, consisting of the original 3D CT scan, polygonal representations of the airway tree and defined 3D biopsy sites, and associated guidance paths, is saved in a data structure referred to as the *case study* and available for subsequent bronchoscopy.

2.2. Stage 2: image-guided bronchoscopy

During a guided bronchoscopy procedure, our system simultaneously draws upon both the bronchoscope's video stream and the previously built case study. In the bronchoscopy laboratory, the bronchoscope's video feed is interfaced to the computer, giving a live video stream.

During the procedure, the following steps are performed for each preplanned biopsy site. First, the computer display presents the physician with an initial CT rendering along the guidance path. Next, the physician moves the scope "near" the presented site. An automatic registration step is then performed to adjust the virtual CT world to the real video world, bringing the two worlds into registration. When registration is complete, a rendition of the target biopsy site is fused onto the video view and distance information related to the scope's current position and biopsy-site position is presented.

With the aid of the computer, the physician continues along the guidance path to the biopsy site, iterating the steps above. This continues until the physician reaches the biopsy site and performs the biopsy. Section 5 further illustrates the features and use of the system.

The key step during guided bronchoscopy is the registration of the 3D CT to the bronchoscopic video. The bronchoscopic video—the *Real World*—is a live real manifestation of the patient's chest during the procedure. The 3D CT image—the *Virtual World*—acts as a high-resolution copy of the patient's chest.

The registration problem can be looked upon as one of matching the viewpoints of two cameras. The first camera—the bronchoscope—gives 2D endoluminal airway video images $I_V(x, y)$ inside the Real World of the human chest. The second camera provides 2D rendered endolumi-
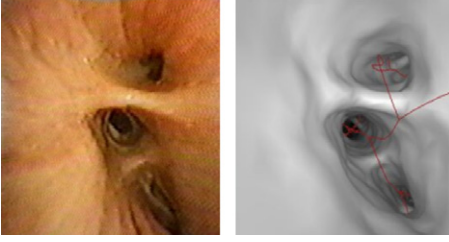
Fig. 3. Matching sample views $I_V(x,y)$ (left) and $I_{CT}(x,y)$ (right) for a typical interior airway location.

nal airway images $I_{CT}(x,y)$ inside the Virtual World of the 3D CT image. Both cameras provide information, albeit in slightly different forms, on the same physical 3D structure: the interior of the 3D airway tree. See Fig. 3 for examples of $I_V(x,y)$ and $I_{CT}(x,y)$. The goal of registration is to align the viewpoints of the two cameras so that they are situated at the same point in space and simultaneously give images of the same region.

The registration process is initialized by assuming that the bronchoscope (Real World camera) is at a fixed viewpoint, giving a fixed reference video image $I_V^F(x,y)$, while the Virtual World camera begins at an initial viewpoint χ_i that is "within a reasonable vicinity" of the bronchoscope's viewpoint, giving view $I_{CT}^{\chi_i}(x,y)$. During registration, an optimization process searches for the optimal viewpoint χ_o via

$$\chi_o = \arg \left\{ \max_{\chi \in N_{\chi_i}} [S_{NMI}(I_{CT}^{\chi}(x,y), I_V^F(x,y))] \right\} \quad (1)$$

to give the virtual image $I_{CT}^{\chi_o}(x,y)$ best matching the fixed video target $I_V^F(x,y)$; in (1) N_{χ_i} represents a search neighborhood about the starting viewpoint χ_i and S_{NMI} represents the normalized mutual information (NMI) between views of the two cameras [28]. Section 3 fully describes the registration problem (1).

3. Mathematical methods

Section 3.1 describes the shared camera geometry assumed for both the virtual CT world and the real bronchoscopic video world. Section 3.2 discusses image modelling considerations specific to the bronchoscopic video images $I_V(x,y)$, while Section 3.3 describes details related to computing the virtual-world endoluminal views $I_{CT}(x,y)$. Finally, mathematical considerations pertaining to the NMI-based registration problem (1) are given in Section 3.4.

3.1. Camera geometry

Each data source, $I_V(x,y)$ and $I_{CT}(x,y)$, acts as a camera that provides a 2D image of an observed 3D scene. As discussed below, our system sets up both cameras to abide by the same imaging geometry.

What a camera sees is determined by its viewpoint, specified by the six-parameter quantity $\chi = (X,Y,Z,\alpha,\beta,\gamma)$.

(X,Y,Z) represents the camera's 3D global spatial position in World coordinates, while (α,β,γ) are Euler angles describing the camera orientation about the focal point. A local coordinate system (x,y,z) can be set up about World point (X,Y,Z). For the local system, the positive z axis is in front of the camera, the positive x axis points to the right, and the positive y axis points up. World point (X,Y,Z), which is point $(0,0,0)$ in local coordinates, coincides with the camera's focal point; α, β, and γ are the rotation angles about the x, y, and z axes, respectively. The camera's viewing screen, which captures the resultant 2D image, is perpendicular to the camera's z axis and is situated a distance f from the focal point, where f is the focal length.

The observed 3D scene is projected onto the camera's viewing screen through a standard 3D-to-2D perspective projection. For a given observable World point $\mathbf{p} = (X_p, Y_p, Z_p)$, we first transform it into the camera's local coordinate system:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{R}(\alpha,\beta,\gamma) \begin{bmatrix} X_p - X \\ Y_p - Y \\ Z_p - Z \end{bmatrix}, \quad (2)$$

where (X_c, Y_c, Z_c) is the transformed point and $\mathbf{R}(\alpha,\beta,\gamma)$ is the rotation matrix [29]. Finally, the point is converted to a 2D viewing screen location (x,y) through the perspective transformation

$$x = \frac{fX_c}{Z_c}, \quad y = \frac{fY_c}{Z_c}. \quad (3)$$

The viewing screen's focal length f and physical dimensions determine a camera's field of view (FOV). To facilitate straightforward registration, we make both the bronchoscope and virtual-world cameras have the same FOV. Thus, if the two cameras are perfectly registered, then pixel (x,y) in bronchoscope image $I_V(x,y)$ and virtual-world image $I_{CT}(x,y)$ arises from the same physical 3D scene point.

To match the FOVs of the two cameras, we do two things. First, prior to bronchoscopy, we calculate the bronchoscopic camera's focal length f (Section 3.2) and use f for the virtual-world camera's geometry (Section 3.3). Second, we make the World coordinate system coincide with the 3D CT image's voxel coordinates. Let the intensity value of voxel (i,j,k) in the 3D CT image be given by $I(i,j,k)$, where i, j, and k are the column, row, and slice indices of the 3D CT image. Then, the World coordinate position of CT voxel (i,j,k) is given by

$$(X_i, Y_j, Z_k) = (i\Delta x, j\Delta y, k\Delta z), \quad (4)$$

where Δx, Δy, and Δz are the sampling intervals.

3.2. Bronchoscopic image modeling

A bronchoscope uses a built-in illumination source and a CCD camera to produce a continuous 2D video stream

of the observed airway-tree interior. The aperture of the bronchoscope is situated at its tip. The tip denotes the 3D World position of the bronchoscope inside the airway tree. Most modern CCD-based bronchoscopes produce pseudo-color images. For our work, we only need the luminance (gray-scale) component. Optically, the bronchoscope tip can be modeled as a point light source that coincides with the device's CCD camera viewpoint [30]. Within this model, the illuminated endoluminal airway surface is Lambertian (diffuse), and the image brightness (irradiance) of an illuminated airway surface point $\mathbf{p} = (X_p, Y_p, Z_p)$ is

$$I(\mathbf{p}) = \sigma \frac{L \cos \theta_s}{\pi R^2}, \tag{5}$$

where L is the intensity of the bronchoscope's light source, θ_s is the angle between the light source (same as the camera's z axis) and $\mathbf{p}$'s surface normal, R is the distance from the light source to $\mathbf{p}$, and σ is a proportionality factor that takes into account airway-surface albedo and the bronchoscope's device characteristics. The value $I(\mathbf{p})$ then passes through the bronchoscope camera's imaging optics, per (2, 3), to give the final value $I_V(x, y)$. Bronchoscopic video does tend to have some specular component as well, but this tends to be in small, wet areas of the airway's interior surface and has little impact on the overall scene illumination model.

In reality, a bronchoscope's camera employs a barrel distortion to give a wide angle ("fish eye") FOV. This feature gives the physician more detail near the center of the image. Since the barrel distortion literally stretches the observed scene nonlinearly, it has become common to correct for this distortion [31]. For example, Stefansik et al., in their efforts to build an image-guided liver-surgery system, used video-distortion correction to help register 3D CT-based liver-surface renderings and corrected video from a rigid laparascope [12]. For our system, prior to bronchoscopy, we perform a simple off-line calibration that gives a corrective transformation for undoing this barrel distortion. More importantly, this transformation also enables real-time matching of the bronchoscope's FOV to the virtual-world camera's FOV [32].

To do this prior off-line computation, the bronchoscope is first mounted in a calibration device at a known distance from a predefined calibration dot pattern, and a bronchoscope image of the dot pattern is captured. Next, a series of calculations are run on the captured (distorted) dot-pattern image. These calculations provide a set of polynomial coefficients that define the distortion-correction transformation. Ref. [32] gives complete detail for these calculations.

These calculations also give the focal length f of the bronchoscope camera as follows. Let X_r and X_1 denote the horizontal positions of the right-most and left-most dots in the calibration dot pattern, and let x_r and x_1 denote analogous quantities for the distortion-corrected image of the pattern. Then, from the perspective Eqs. (2, 3),

$$x_r = \frac{f(X_r - X_m)}{Z_m} + x_m, \quad x_l = \frac{f(X_1 - X_m)}{Z_m} + x_m, \tag{6}$$

where x_m is the horizontal coordinate of the viewing-screen center, X_m is an analogous coordinate on the original dot pattern, and Z_m is the known distance of the calibration pattern from the bronchoscope. From (6), the viewing-screen width of this image is $x_r - x_l = \frac{f(X_r - X_1)}{Z_m}$. Thus, the focal length is given by

$$f = \frac{(x_r - x_l) Z_m}{(X_r - X_1)}. \tag{7}$$

All quantities X_r, X_1, X_m, Z_m, x_r, x_l, and x_m are known or easily computed using the calibration calculations. Also, the bronchoscope's field-of-view angle is readily computed:

$$\theta_{FOV} = 2 \tan^{-1} \left(\frac{x_r - x_l}{2f} \right). \tag{8}$$

Let x_r' and x_l' denote the horizontal coordinates of the right-most and left-most dots in the distorted (uncorrected) calibration pattern image. The correction-polynomial coefficients are scaled by the factor

$$\frac{x_r - x_l}{x_r' - x_l'} \tag{9}$$

so that the corrected image has the same width as the uncorrected image. Now, the distortion-corrected image will fit into the same window size as the incoming distorted bronchoscopic image. Also, the FOV angle is known.

The bronchoscope's focal length (7) and FOV angle (8) will be used for the Virtual-World camera, as discussed in Section 3.3. This results in two images sources, I_V and I_{CT}, arising from cameras having matching FOVs. In addition, during live bronchoscopy, the calculated parameters are used to produce distortion-corrected video in real-time [32].

3.3. Endoluminal rendering

The 3D CT image defines the Virtual-World representation of the chest. During image-guided bronchoscopy, the Virtual-World camera is maneuvered through the major airways depicted in the 3D CT image. At each viewpoint $\chi = (X, Y, Z, \alpha, \beta, \gamma)$, an endoluminal rendering $I_{CT}^\chi(x, y)$ is produced. These endoluminal renderings act as simulated endoscopic views of the airway-tree interior.

The intensity value $I_{CT}^\chi(x, y)$ of a particular screen point (x, y) abides by the endoscope Lambertian shading model (5). For a particular viewpoint χ, a range map and a set of angles θ_s for all needed 3D scene points $\mathbf{p}$ within the required FOV are computed. Calculation of the angles θ_s use the triangle surface normals. The viewing-screen dimensions are determined by the FOV angle θ_{FOV} from (8) and World coordinate-system dimensions from (4). The computation of the viewing-screen image uses the imaging geometry of (2, 3) and the focal length (7). A small ambient lighting term is added to the rendering calculation to keep all values >0. In addition a small constant factor is

added to the denominator of (5) to avoid unstable calculations. All endoluminal rendering calculations are performed in hardware using standard OpenGL commands [23].

3.4. Registration algorithm

The two data sources do not provide physically identical images, but they do have much in common. For the image-guided bronchoscopy problem, Fig. 3 shows examples of a video image $I_V^\chi(x, y)$ and endoluminal rendering $I_{CT}^\chi(x, y)$ observing the 3D World from the same viewpoint χ. Looking at the two images, it is obvious that they are aligned to view the same 3D structure. As described earlier, both image sources abide by the same imaging geometry and have the same FOV. Also, their intensity characteristics are similar: both sources depict surface-shape information in the form of a depth-shaded Lambertian-surface model [29,30]. The bronchoscopic video image, however, deviates somewhat from this simple intensity model in that it is able to depict airway-wall mucosal detail and small specular reflections near wet portions of the airway wall. The endoluminal rendering, on the other hand, does not deviate from the model and only depicts surface-shape information, albeit with high quality. Also, a small ambient intensity bias tends to exist between the two images, as a result of rendering options and bronchoscope gain characteristics. The similarities between the two image sources, however, make them well-suited for image registration.

Mutual information, which arises in information theory to measure the statistical dependence between two random variables, is commonly used for registering images from different modalities [9,28,33,34]. It can be measured with the Kullback-Leibler metric [34], which for our problem is given by

$$S_{MI}(I_V, I_{CT}) = \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} p_{V,CT}(k, l) \log \frac{p_{V,CT}(k, l)}{p_V(k) \cdot p_{CT}(l)}, \quad (10)$$

where images I_V and I_{CT} are the two "random variables" being compared, $p_V(k)$ and $p_{CT}(l)$ are the respective marginal probability density functions of the images (normalized image histograms), $p_{V,CT}(k, l)$ is the joint density function between the two images (normalized joint histogram between the two images), and $M = 256$ is the number of gray-levels used. S_{MI} can also be written in terms of entropy:

$$S_{MI}(I_V, I_{CT}) = h(V) + h(CT) - h(V, CT), \quad (11)$$

where $h(V)$ and $h(CT)$ are the marginal image entropies and $h(V, CT)$ is the joint entropy between the two images. Studholme et al., however, performed a detailed study demonstrating that the basic S_{MI} measure can fail to properly align two images if the amount of image overlap is large [28]. Also, in bland overlapping image regions, where $h(CT) \approx h(V) \approx h(V, CT)$, S_{MI} is sensitive to the lack of statistical information contributed by these overlapping regions. This reduces the measure's ability to recover from

larger initial misalignments and to register images that have much overlap. For these reasons, we use the idea of normalized mutual information (NMI) by Studholme et al. [28], which involves normalizing S_{MI} in (11) by $h(V, CT)$:

$$\begin{aligned} S_{NMI}(I_V, I_{CT}) &= \frac{h(V) + h(CT) - h(V, CT)}{h(V, CT)} \\ &= \frac{h(V) + h(CT)}{h(V, CT)} - 1. \end{aligned} \quad (12)$$

For S_{NMI}, any increase in marginal entropies is counterbalanced by a change in joint entropy, making the measure less dependent on the amount of overlap [28]. Note that $S_{NMI}(I_V, I_{CT})$ as defined in (12) tends to have a magnitude on the order of 10^{-2}. Studholme et al. actually excluded the "-1" factor in their work, but we have found that this makes the optimization less sensitive. Hence, we have found it advantageous to keep this term.

Bricault et al. have pointed out the following for endoluminal airway images [8]: (1) the significant image information resides primarily near airway bifurcations and the corresponding dark "holes" leading to upcoming airways and near large (darker) deviations in the airway walls; and (2) the brighter bland wall regions tend to have little useful information. Also, the specular reflections that appear in small regions of the video images correspond to saturated intensity points and do not appear in the endoluminal renderings. Drawing upon these observations, we modify the entropy calculations by varying the weight assigned to darker and brighter image points. Noting that $p_V(k) = \sum_{l=0}^{M-1} p_{V,CT}(k, l)$ and $p_{CT}(l) = \sum_{k=0}^{M-1} p_{V,CT}(k, l)$, we can write

$$h(V) = -\sum_{k=0}^{M-1} \sum_{l=0}^{M-1} p_{V,CT}(k, l) \log p_V(k),$$

$$h(CT) = -\sum_{k=0}^{M-1} \sum_{l=0}^{M-1} p_{V,CT}(k, l) \log p_{CT}(l),$$

$$h(V, CT) = -\sum_{k=0}^{M-1} \sum_{l=0}^{M-1} p_{V,CT}(k, l) \log p_{V,CT}(k, l).$$

These are modified by adding weighting factors:

$$h(V) = -\frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} w_k w_l p_{V,CT}(k, l) \log p_V(k),$$

$$h(CT) = -\frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} w_k w_l p_{V,CT}(k, l) \log p_{CT}(l),$$

$$h(V, CT) = -\frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} w_k w_l p_{V,CT}(k, l) \log p_{V,CT}(k, l),$$

$$\quad (13)$$

where the weights w_k are given by

$$w_k = (M - k)/M, \quad k = 0, 1, \ldots, M - 1. \quad (14)$$

(The w_l are also given by (14) with l replacing k.) While other weighting schemes are possible, this scheme empha-

sizes darker pixels while attenuating brighter pixels. Thus, greater emphasis is placed on the important darker image structures. Our earlier efforts showed that (13) can be used successfully for registration with no weights (i.e., $w_k = 1$, $k = 0, 1, \ldots, M - 1$) [35,36]. But the unweighted method proved to be less robust for situations depicting little visible upcoming airway structure, which generally appears dark as an "entrance to a cave." Hence, we followed the observation of Bricault et al. stating that darker areas contain more information when inside the airway tree [8].

For our circumstance, S_{NMI} is maximized when the joint density $p_{\mathrm{V,CT}}(k,l)$ is approximately a diagonal matrix. This behavior can be explained as follows. If the two image sources are perfectly aligned and have identical intensity characteristics, then $\forall k$, $p_{\mathrm{V,CT}}(k,k) = p_{\mathrm{V}}(k) = p_{\mathrm{CT}}(k)$. But, as stated earlier, the two image sources differ somewhat in intensity characteristics. If the images become properly aligned, then $P_{\mathrm{V,CT}}(\cdot,\cdot)$ will still be concentrated along a diagonal in the k-l space, with some deviation about the diagonal to account for small local deviations in source differences and with a possible shift to account for ambient intensity bias.

Finally, to solve the registration problem (1), the optimization starts with initial viewpoint χ_i for the endoluminal renderings and with fixed target video frame $I_{\mathrm{V}}^{\mathrm{F}}(x,y)$. During optimization, viewpoint $\chi = \{X, Y, Z, \alpha, \beta, \gamma\}$ is varied in a neighborhood N_{χ_i} about χ_i. An endoluminal rendering $I_{\mathrm{CT}}^{\chi}(x,y)$ is computed for each candidate viewpoint χ and compared to $I_{\mathrm{V}}^{\mathrm{F}}(x,y)$ through the S_{NMI} measure (12), weighted entropies (13), and weights (14). This is a six-parameter optimization problem, per the six parameters constituting χ. χ is varied until an optimal viewpoint χ_o is found that maximizes S_{NMI}. Three standard optimization algorithms were investigated: (1) steepest-ascent [37], (2) Nelder-Mead simplex [38], and (3) simulated annealing [39]. In our experiments, all algorithms converged well in the large majority of circumstances. Section 4.1 studies the performance of these algorithms. We used the simplex algorithm for the majority of our tests.

The following section gives detailed validation results for the registration method. Section 5 then gives complete system results.

4. Validation of registration method

A major system-feasibility issue is the robustness and practicality of the CT-video registration method, critical to the guidance system. This section provides validation results for the method. First, Section 4.1 defines the error criteria used to measure performance. Next, Sections 4.2–4.5 present four sets of tests for the CT-video registration method.

4.1. Optimization set-up and error measures

All tests were performed on a dual-CPU Dell Precision 620 workstation, as discussed earlier. We did not optimize the computer code. The 3D MDCT images used in Sections 4.2–4.4 had typical resolutions of $\Delta x \approx \Delta y \approx \Delta z \approx 0.6$ mm and generally consisted of 400–500 2D slices. The bronchoscope video source provided analog video at 30 frames/s. In real time video frames were digitized to 264×264 arrays and underwent distortion correction.

The parameters varied during optimization were the increments for the six parameters constituting the viewpoint χ: ΔX, ΔY, ΔZ, $\Delta\alpha$ $\Delta\beta$, and $\Delta\gamma$. For the simplex and annealing methods, the parameter increments define the six side lengths of the initial simplex. The steepest ascent and simplex methods were run for 200 iterations and simulated annealing was run in two steps of 100 iterations. In the first 100 iterations of simulated annealing, a random variable was added to each S_{NMI} calculation to keep the algorithm from stalling in local capture regions. The second 100 iterations were run like the simplex method without random changes. With 200 iterations, all algorithms ran in a tolerable amount of time and showed consistently successful results. These parameter values are realistic and flexible for our airway imaging scenario involving MDCT images and videobronchoscopy data. They were selected after an extensive ad hoc study done over various animal and human data sets [40]—see Section 5.

All tests were run by setting an initial start position for I_{CT}^{χ} and target frame I_{V}. In all tests the I_{CT}^{χ} were virtual-world renderings. But, in some tests (Sections 4.2 and 4.3), the "video image" I_{V} was actually another virtual image. This was done to eliminate certain variables introduced by the video and to provide more control over a given test. All tests were set up so that the true viewpoint for the target "video image" I_{V} was known. The final position $I_{\mathrm{CT}}^{\chi_o}$ after optimization was compared to the true position of the fixed target video image I_{V} by calculating three different error measures that consider position, angle, and biopsy needle distance. Let the final registration's viewpoint χ_o and true viewpoint χ_t be defined as:

$$\chi_o = (X_o, Y_o, Z_o, \alpha_o, \beta_o, \gamma_o),$$
$$\chi_t = (X_t, Y_t, Z_t, \alpha_t, \beta_t, \gamma_t).$$

Denote the final registration's viewpoint position as $p_o = (X_o, Y_o, Z_o)$ and the true viewpoint's position as $p_t = (X_t, Y_t, Z_t)$. Also, the unit vector for the viewing direction at final registration is $\mathbf{v}_o = \mathbf{v}(\alpha_o, \beta_o, \gamma_o)$, while the unit vector for the viewing direction of the true position is $\mathbf{v}_t = \mathbf{v}(\alpha_t, \beta_t, \gamma_t)$. The *position error* is defined as the difference between p_o and p_t:

$$e_{\mathrm{p}} = \|p_o - p_t\|, \tag{15}$$

where $\|\cdot\|$ is vector magnitude calculated as $\|(x,y,z)\| = \sqrt{x^2 + y^2 + z^2}$. The *angle error* is defined as

$$e_{\mathrm{a}} = \cos^{-1}(\mathbf{v}_o \cdot \mathbf{v}_t), \tag{16}$$

where $\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$ and θ is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$. Finally, the *needle error* is the distance between points situated a fixed length in front of the two viewpoints. This
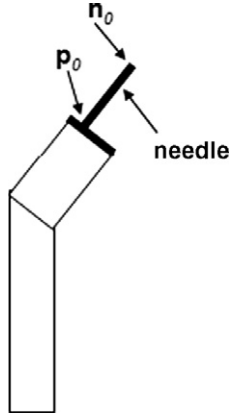
Fig. 4. Schematic figure of bronchoscope tip. The figure illustrates a needle protruding from the end of the bent tip. The location of the scope tip is $\mathbf{p}_0$ and the needle end is $\mathbf{n}_0$. The length of the needle protruding beyond the scope tip is d_n.

represents the resulting error if the virtual view is used in a needle biopsy. Suppose during bronchoscopy that a biopsy needle extends a distance d_n away from the tip of the bronchoscope; see Fig. 4. The needle tip's location is based on the final registered position $\mathbf{n}_o$ and on the true bronchoscope position $\mathbf{n}_t$:

$$\mathbf{n}_o = \mathbf{p}_o + d_n \mathbf{v}_o, \quad \mathbf{n}_t = \mathbf{p}_t + d_n \mathbf{v}_t. \tag{17}$$

Thus, the *needle error* is the distance between these two needle-tip locations:

$$e_\mathrm{n} = \|\mathbf{n}_o - \mathbf{n}_t\|. \tag{18}$$

4.2. Performance of optimization algorithms

The first set of tests focused on the relative performance between the three optimization algorithms. One airway-tree ROI was used. A virtual image at this ROI was used as the target video image I_V during registration. By doing this, we eliminate the source differences between the video and virtual CT, and we can measure performance without being influenced by the qualitative differences between the video and virtual CT images. We also can do precise matching between the I_CT^{χ} and the target view I_V, since we know the precise location of the target site.

For each of the three optimization algorithms, we varied each of the six viewpoint parameters separately. For a test varying ΔX, the initial viewpoint started within the range -10 mm to $+10$ mm from the true viewpoint's X location, with all other viewpoint parameters Y, Z, α, β, and γ starting at the true values per I_V. The -10 mm to $+10$ mm range was also used for tests varying the Y and Z viewpoint starting positions. Tests varying the starting points for the initial roll, pitch, or yaw angles ranged from $-20°$ to $20°$, with the other five viewpoint parameters again starting at the correct values. For the needle error e_n, we considered a point 10 mm in front of a view's actual viewpoint to arrive at n_o and n_t in (17).

For a given optimization algorithm, all three error measures, e_p, e_a, and e_n, were computed. Therefore, 18 total error plots (6 viewpoint parameters, 3 error measures) were generated for each optimization method. Fig. 5 gives a sample of the error plots generated. In some instances a starting viewpoint may result in a view situated outside the airways. Since such views are not valid starting points (and the bronchoscope could not be positioned at these locations!), they are not included in the plotted interval. Fig. 5a gives an example of this phenomenon, where the interval for ΔY variation could only be considered from -2 mm to 6 mm. The error plots provide a numerical summary of an optimization algorithm's sensitivity to starting point. For a given viewpoint parameter, acceptable registration performance occurs if the measured error is low for a wide range of initial values. Such a range can be inferred from an error plot by noting the range of values that give a low error. Based on these tests and on the exhaustive results presented elsewhere, we used the following benchmarks to signify good final registrations [40]:

$$\text{final } X, \; Y, \; Z \text{ position/needle errors} < 5 \text{ mm} \tag{19}$$

$$\text{final } \alpha, \; \beta, \; \gamma \text{ angle errors} < 5° \tag{20}$$

To get a sense of size within the airways, note, for example, that the typical diameter of the left main bronchus is on the order of 20 mm. The values above give acceptable deviations of the final computed optimal viewpoint of $I_\mathrm{CT}^{\chi_o}$ relative to I_V's actual viewpoint. Based on these benchmarks, we can infer from Fig. 5b, for example, that the simplex algorithm (solid lines in the plots) gave acceptable registrations in terms of e_a for the starting Z value of χ_i deviating in the range -9 mm $\leqslant \Delta Z \leqslant 10$ mm.

Average computation times for the optimization algorithms were as follows:

Steepest ascent—8.91 s.
Simplex—23.16 s.
Simulated annealing—24.04 s.

As expected, steepest ascent runs the quickest, but it in general converges the poorest. In reality, however, this algorithm usually gives acceptable registrations, but it is more sensitive to starting point than the other algorithms. Overall, the simplex algorithm was found to be most practical: it is both robust and runs sufficiently quickly. This is an interesting result, since simulated annealing might have been expected to perform better, particularly at more distant starting points (where undesirable local maxima are more likely to occur). This was not the case. A possible reason for this was that the search step-size was relatively large. The simplex algorithm was able to search a relatively large area at the beginning of registration. This allowed it to find the best solution overall and avoid local maxima. We emphasize that our computer implementations could be improved and that it was not our intention to construct the "best" optimization technique. Our results do show the
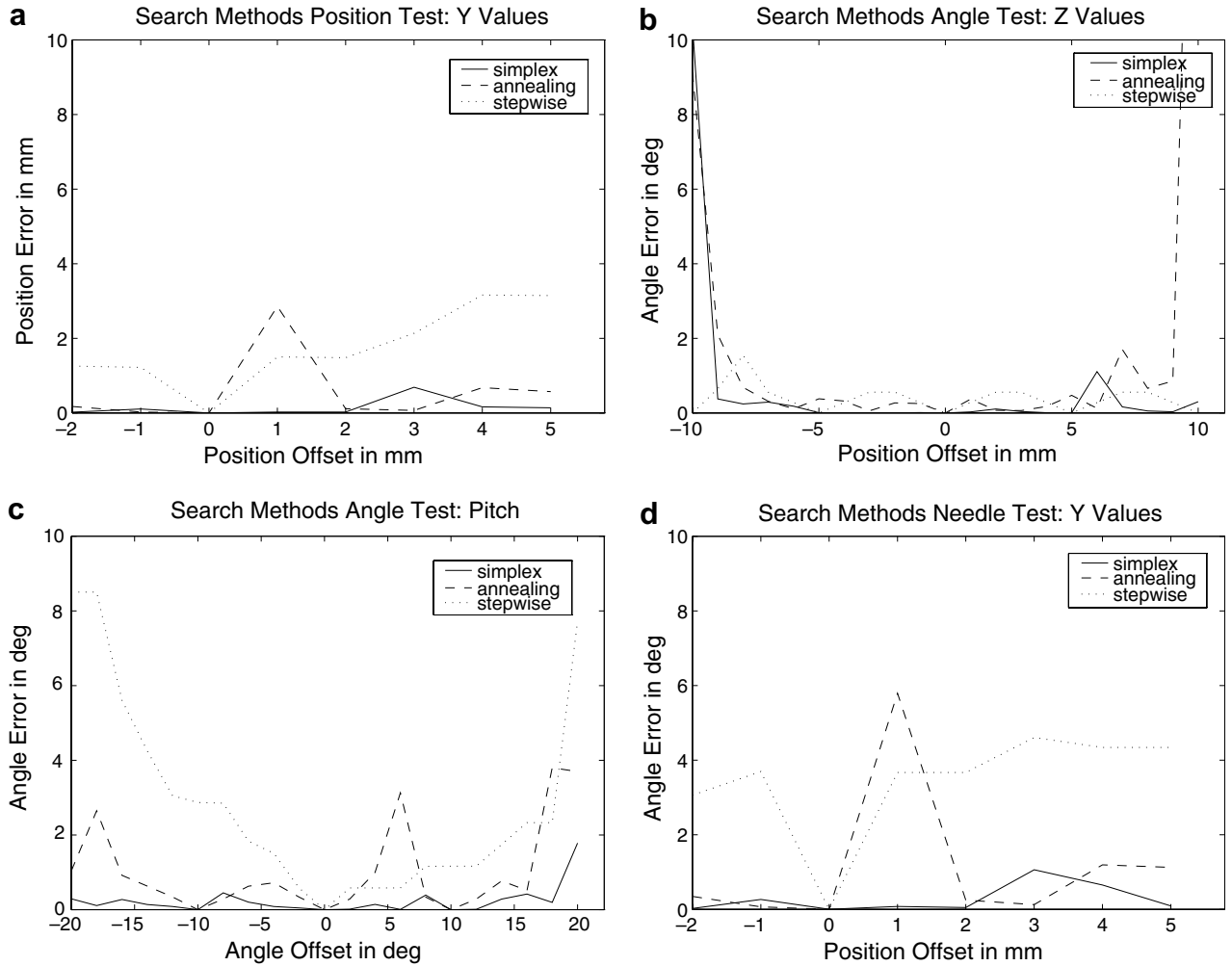
Fig. 5. Sample position, angle, and needle errors for simplex, simulated annealing and steepest ascent (stepwise) search methods. For I_{CT}^{χ}, the starting viewpoint before registration would have all parameters of χ_i equal to those of I_V's viewpoint except one. Registration would then be done to find the optimal view I_{CT}^{χ} matching I_V. Each plot shows the effect of varying one of the parameters in χ_i's initial starting viewpoint from the final solution. The horizontal axis gives the value of the initial offset, and the vertical axis gives the final error after optimization. Figure parts are as follows: (a) position error e_p: variation of initial Y position (ΔY); (b) angle error e_a: variation of initial Z position (ΔZ); (c) angle error e_a: variation in pitch angle; (d) needle error e_n: variation in initial Y position (ΔY).

robustness of the registration method for our real airway-analysis scenario. We used the simplex optimization algorithm for the remainder of our tests [40].

4.3. Sensitivity to variations in airway morphology

The airways can differ significantly in their shape and size within the airway tree. It is important to test the robustness of the CT-video registration method to various locations within the airway tree. We again fix the target "video" frame I_V as a virtual-world endoluminal airway rendering and vary the initial position of I_{CT}^{χ}. We used the same parameters as before in varying the initial viewpoint and ran the simplex optimization algorithm. For I_V, we used the six test ROIs depicted in Fig. 6. These ROIs gave a good mix of airway size and geometry.

Table 1 gives tabular results for e_p. (Exhaustive results for this test and all others to follow appear elsewhere [40,37].)

The table gives the range of initial viewpoint values that result in final acceptable registrations, per (19, 20). As an example, for ROI 2, the method gives acceptable final position error ($e_p < 5$ mm per (19)) if the initial viewpoint χ_i of $I_{CT}^{\chi_i}$ satisfies the following: -10 mm $\leqslant \Delta X \leqslant 7.3$ mm, -6.4 mm $\leqslant \Delta Y \leqslant 5.4$ mm, -5.1 mm $\leqslant \Delta Z \leqslant 10$ mm, etc. The averages over all six ROIs appear at the bottom of the table. Table 2 summarizes these averages for e_a and e_n. As is clear from these results, the performance is robust for ROIs located over widely varied regions within the airway tree. All of the results of Tables 1 and 2 point to the considerable robustness of the method. To gain perspective on these numbers, note that a typical airway diameter is under 10 mm, with the trachea being by far the largest at around 20–30 mm. In addition a rotation angle of $\pm 20°$ represents a gross positional difference. This robustness has been borne out in our live studies, where five different technicians successfully employed the system.
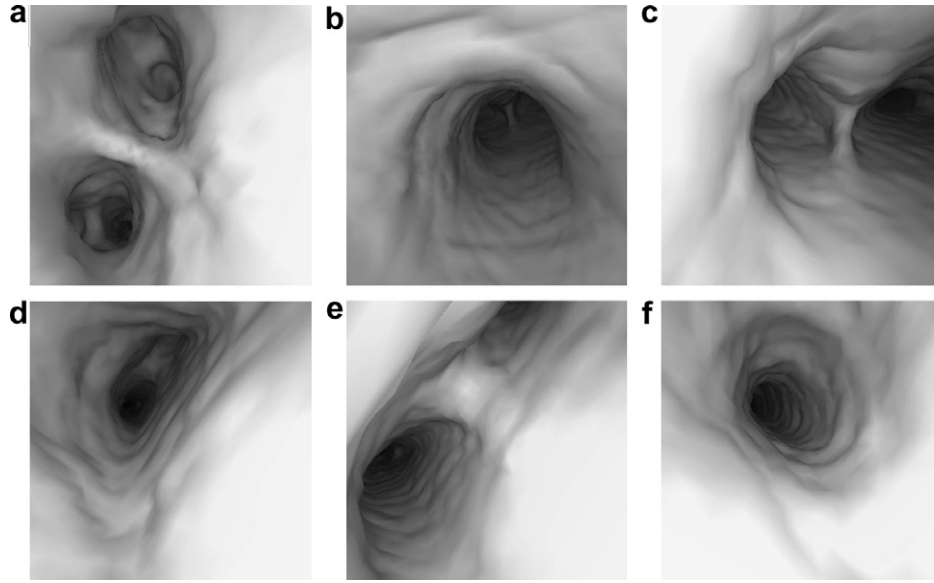
Fig. 6. ROIs used in testing the registration method's sensitivity to airway morphology. Each ROI served as a target "video" frame I_V: (a) ROI 0—middle of trachea; (b) ROI 1—trachea near main carina; (c) ROI 2—proximal end of right main bronchus; (d) ROI 3—distal end of right main bronchus; (e) ROI 4—proximal end of left main bronchus; (f) ROI 5—distal end of left main bronchus.

Table 1
Sensitivity to airway morphology: bounds on initial viewpoint variations that result in acceptable position error e_p, per (19)

ROI Number	X bound min, max	Y bound min, max	Z bound min, max	Roll bound min, max	Yaw bound min, max	Pitch bound min, max
ROI 0	−10.0, 4.9	−10.0, 10.0	−10.0, 8.8	−20.0, 20.0	−20.0, 17.6	−20.0, 20.0
ROI 1	−7.6, 10.0	−10.0, 10.0	−10.0, 10.0	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
ROI 2	−10.0, 7.3	−6.4, 5.4	−5.1, 10.0	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
ROI 3	−10.0, 10.0	−10.0, 6.5	−10.0, 5.7	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
ROI 4	−10.0, 10.0	−7.1, 10.0	−10.0, 4.6	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
ROI 5	−10.0, 10.0	−10.0, 10.0	−10.0, 10.0	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
Average	−9.6, 8.7	−8.9, 8.6	−9.2, 8.2	−20.0, 20.0	−20.0, 19.6	−20.0, 20.0

Each row gives the performance for one of the ROIs of Fig. 6. Each column gives the acceptable deviation in initial value for one of the viewpoint parameters defining χ_i for $I_{CT}^{\chi_i}$.

Table 2
Sensitivity to airway morphology: ranges of acceptable registration performance for e_a and e_n averaged over six ROIs (Fig. 6), per (19, 20)

Error Measure	X bound min, max	Y bound min, max	Z bound min, max	Roll bound min, max	Yaw bound min, max	Pitch bound min, max
e_a	−9.6, 8.5	−7.7, 8.6	−9.1, 8.0	−20.0, 20.0	−20.0, 19.5	−20.0, 20.0
e_n	−9.7, 9.1	−7.2, 8.4	−9.1, 6.3	−20.0, 20.0	−19.3, 19.5	−20.0, 19.6

As in Table 1, each column gives the acceptable deviation in initial value for one of the viewpoint parameters for $I_{CT}^{\chi_i}$.

4.4. Registration of bronchoscopic video to an MDCT-based endoluminal rendering

Registration was next tested using a true human bronchoscopic video frame for I_V. To be able to measure accuracy, we first located a reference virtual view I_{CT} deemed to match I_V's viewpoint "perfectly." We found this reference view I_{CT} by running many tests about a given video site and noting the result deemed the best. This then gave a pair of matching views for a given ROI. This test used six pairs of ROIs, exhibiting varying airway morphology. We then varied the initial starting position of $I_{CT}^{\chi_i}$, per the simplex registration parameters used in the earlier tests, and performed the optimization using the video frame I_V as the fixed target I_V^F. After registration terminated, the reference I_{CT} view, corresponding to I_V^F, was compared to the final optimization result $I_{CT}^{\chi_o}$ to compute registration error. Since the image sources are now different, this test evaluated the S_{NMI} criterion as well as the search method.

Fig. 7 gives summary error plots for one ROI pair, and Table 3 gives performance ranges over the six ROI pairs. These ranges indicate the values for which an acceptable
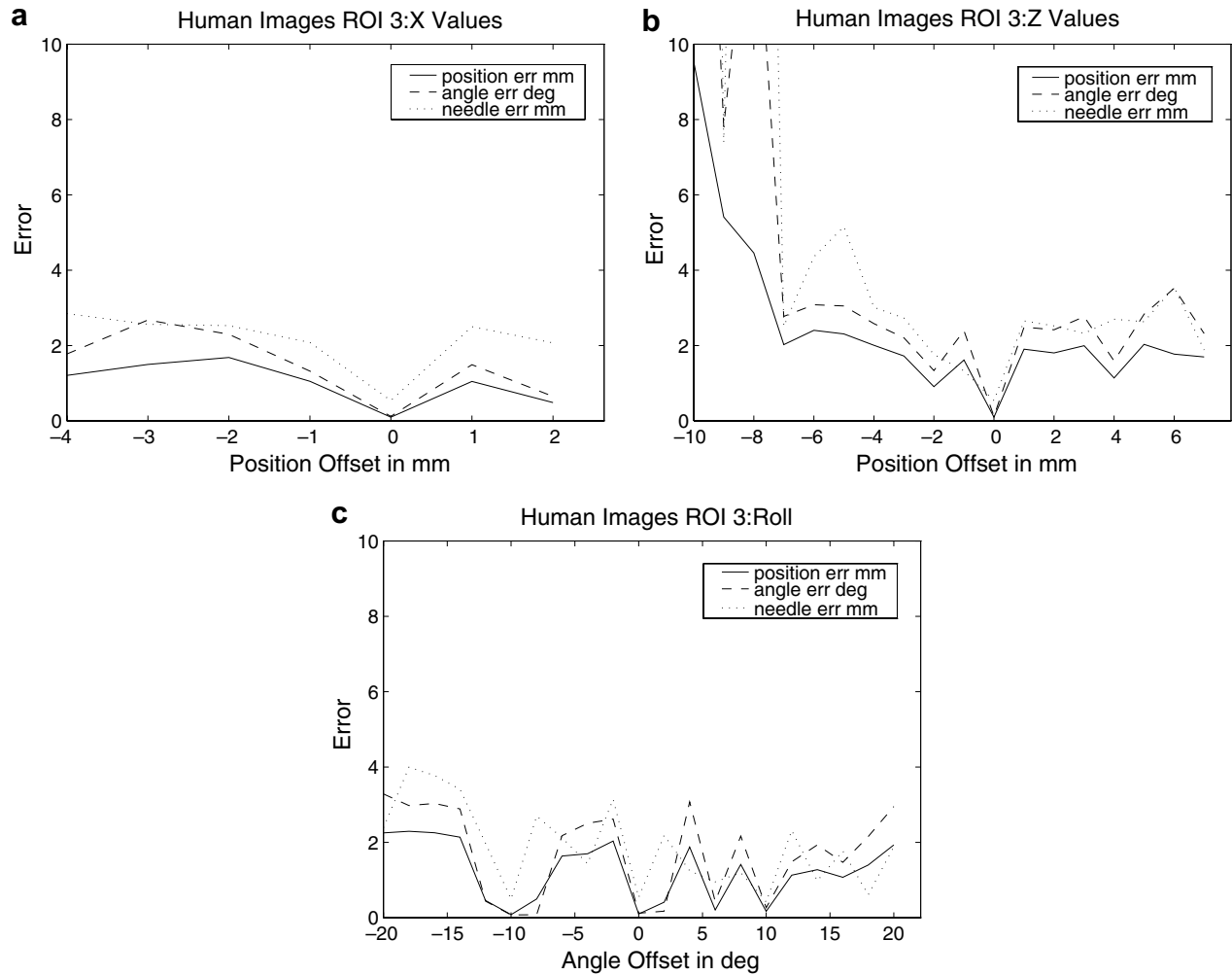
Fig. 7. Registration of video and virtual views: error plots for ROI pair 3. (a) variation of initial ΔX value; (b) variation of initial ΔZ value; (c) variation in initial roll angle.

registration of $I_{CT}^{\chi_i}$ could be made to I_V (again, $I_{CT}^{\chi_i}$ was actually compared to the reference I_{CT} view to arrive at an error value). Comparing these ranges to the ranges given in the earlier tables based on virtual-to-virtual registration tests, we observe that the robustness drops, but only slightly. This indicates that the normalized mutual information criteria works well for registering a virtual view to a video view.

4.5. Sensitivity to different lung capacities

When a patient undergoes a 3D MDCT scan during the initial 3D CT Assessment stage, he/she is asked to fully inflate their lungs to total lung capacity (TLC). Later, during bronchoscopy, when the patient lies on the bronchoscopy suite table, the patient's lungs are typically near functional residual capacity (FRC), the lung volume when their lungs are nearly completed deflated; the patient only performs very shallow tidal breathing during the procedure. Thus, for our guidance system, we have a TLC MDCT chest volume, while, during bronchoscopy, the chest is at the lower FRC volume. It is important to see if this change in lung volume affects registration accuracy. It is possible that the airways could differ significantly in size for different lung volumes, and, hence, adversely influence registration performance. This final series of tests considered this issue in a controlled circumstance.

The tests considered the scenario when the target "video" frame I_V is a virtual view of an airway site during FRC and the varied virtual-world MDCT view I_{CT}^{χ} is computed from an MDCT scan at TLC. The images used for this test were from a pig (Section 5.2). A lung volume controller maintained consistent air capacity during two separate scans of the pig's chest: (1) at FRC; and (2) at 20 cm H_2O for the inspired volume (near TLC) [41]. For each of the two volumes, we located three corresponding ROI pairs for the test. The positions of these corresponding ROIs were carefully located prior to the test. By using only CT data, we precisely measured performance. Table 4 gives a summary of the error measures over the six ROI pairs. While the acceptable ranges of initial viewpoints are smaller than in the earlier tests, wide ranges of start point

Table 3
Registration of video and virtual views: ranges of acceptable performance averaged over the six ROI pairs when various parameters of χ were varied, per (19,20)

Error Measure	X bound min, max	Y bound min, max	Z bound min, max	Roll bound min, max	Yaw bound min, max	Pitch bound min, max
e_p	−9.7, 8.9	−8.9, 8.9	−7.3, 9.4	−20.0, 20.0	−20.0, 20.0	−20.0, 20.0
e_a	−9.4, 8.0	−7.8, 7.9	−6.6, 7.7	−18.6, 20.0	−20.0, 20.0	−20.0, 20.0
e_n	−10.0, 7.9	−7.7, 6.0	−6.1, 7.1	−15.2, 14.0	−11.0, 17.4	−11.1, 13.8

Table 4
Registration for different lung capacities: ranges of acceptable registration performance averaged over six ROI pairs when various parameters of χ were varied, per (19,20)

Error Measure	X bound min, max	Y bound min, max	Z bound min, max	Roll bound min, max	Yaw bound min, max	Pitch bound min, max
Position	−6.6, 10.0	−10.0, 8.3	−8.7, 8.5	−20.0, 20.0	−20.0, 20.0	−20.0, 18.9
Angle	−6.5, 10.0	−8.0, 6.5	−8.1, 8.2	−20.0, 20.0	−19.5, 20.0	−19.9, 18.5
Needle	−6.2, 7.6	−7.9, 6.3	−9.4, 7.0	−19.0, 20.0	−17.2, 20.0	−19.8, 17.3

deviation are permitted. We note that the ROIs constituting each pair appear to be nearly identical, despite the differences in lung volumes [37]. This property, which arises because the airways are locally rigid, enable us to perform robust registration even under differing lung capacities.

5. System results

Fig. 2 illustrates the use of the system during bronchoscopy. A technician performs all set up tasks, freeing the physician to perform more essential tasks. The system fits smoothly into the real work flow of the procedure. In this section we present three sets of results showing the complete system's performance: (a) a phantom study, which involves a controlled test, free of subject motion; (b) animal studies, which permit controlled tests in an in vivo circumstance; and (c) human studies, involving real lung-cancer patients in the standard clinical work flow.

5.1. Phantom study

We first performed a controlled study involving no motion. Six physicians, ranging in experience from new clinical fellows in training to clinical faculty, performed bronchoscopic "biopsy" on a rubber model of the airway tree augmented with five 1.4 mm platinum beads (the biopsy sites). The phantom model was made of rigid rubber and serves as a surgical training device. It is a molded replica of the human airway tree over four generations; see Fig. 8a. A 3D MDCT scan was done of the phantom, using a Marconi Mx8000 four-detector MDCT scanner. This 3D CT scan was of size $453 \times 155 \times 160$, with sampling intervals $\Delta x = \Delta y = 0.35$ mm and $\Delta z = 1.5$ mm. Five film sheets were printed for this scan with twelve 5 mm-thick transverse-plane $(x - y)$ slices printed on each film sheet; three coronal views $(x - z)$ were also included on the film for reference; see Fig. 8b. Fig. 8c is a coronal projection image of the phantom, which shows the locations of the desired biopsy sites.

Each physician then performed two separate biopsy tests on the phantom. In the first test, they performed the standard procedure, where they had the CT film available in the bronchoscopy lab, but no other guidance aid. In the second test, they used our proposed guidance system to performed the biopsies. For each biopsy site, the physician stuck a needle into the rubber wall of the original air-

Fig. 8. Set-up for phantom study: (a) rubber airway-tree model for phantom study; (b) three of the CT film sheets used for the phantom study; (c) coronal weighted-sum projection of the 3D phantom CT scan—the squares indicate the positions of the five platinum beads serving as biopsy sites for the test.

way-tree model and then called for a measurement. A technician made the measurement using a caliper, accurate to within 0.01 mm, as the distance between the metal bead (the biopsy site) and the line perpendicular to the needle direction.

The results for the individual physicians are give in Table 5, while Table 6 summarizes their performance. Notice that using the standard procedure, the physicians varied greatly in their performance. With the guidance system, the physicians all improved in performance. *In addition, physician performance almost appeared to become independent of experience, and they all performed nearly the same!* (Note that typical biopsy needles have length 20 mm and diameter 0.5–1.0 mm.) Considering that target sites, such as lymph nodes and suspect cancer lesions, typically have a diameter >1 cm, these results are excellent [3].

Table 5
Phantom study results

Physician	Standard (mm)	Guided (mm)
1	5.80	1.38
2	2.73	1.33
3	4.00	1.49
4	8.87	1.60
5	8.62	2.45
6	3.19	1.24

Each row signifies the individual average performance of each physician for the 5 target biopsy sites. The standard approach involves using the CT film as the only guidance aid. The "Guided" column gives results using the proposed guidance system.

Table 6
Phantom study results averaged over the six physicians

Measure	Average biopsy error (mm)	Standard deviation (mm)
Standard	5.53	4.36
Guided	1.58	1.57

"Standard" refers to the standard film approach, while "Guided" gives results using the proposed guidance system. The average biopsy error is an average of all biopsy errors for the six physicians, while standard deviation gives a standard deviation of these errors.

5.2. Animal studies

We next performed five animal studies to determine the system's effectiveness during a live procedure in controlled circumstances. The first three studies were done to test the overall efficacy, safety, and feasibility of the system. Besides showing that the system can function safely and effectively during a live surgical situation, these three studies also demonstrated the fundamental condition that the system and associated CT-video registration method function effectively despite the difference in lung volumes during CT scanning and live bronchoscopy; these data were used for the studies of Section 4.5.

The latter two animal studies were done to test the accuracy of the system during a live controlled situation. We first performed a 3D CT scan of the animal. The CT scans were generated on an Imatron electron-beam CT scanner. The scans were of size $512 \times 512 \times 140$ and had resolution $\Delta x = \Delta y = 0.412$ mm and slice thickness $\Delta z = 1.5$ mm. Given the scan, we then generated a series of "virtual" ROI sites by drawing 3D regions having diameters on the order of 0.5cm in various locations around the airway tree. These ROIs were all situated outside the airways, however, just as in the real scenario where suspect cancer nodules or lymph nodes are the regions of interest. Fig. 9a depicts a 3D rendering of these ROIs for one of the animal studies.

We then used the system to guide the physician to each of the preplanned biopsy sites. Upon reaching a site, the physician would deposit a metallic dart at the site (Fig. 9b). When all sites had been "biopsied," we then rescanned the animal and compared the before and after CT scans—Fig. 9c depicts a sample "after" scan, while Fig. 10 depicts before and after views focused on one ROI. For this test, we achieved the following results averaged over all ROIs considered:

4.95 mm—distance from ROI centroid,
2.37 mm—distance from ROI.

"Distance from ROI centroid" is defined as the distance from the centroid of the preplanned virtual biopsy site to the closest dart point appearing in the second CT scan.
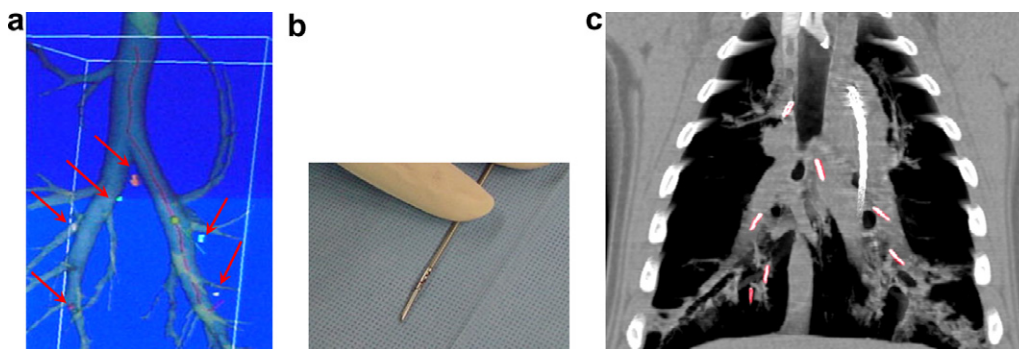

Fig. 9. Results for an animal test. (a) 3D surface rendering showing segmented airway tree, central axes (thin lines), and six planned biopsy sites (pointed to by arrows). (b) Example of a dart used in the animal tests. (c) Coronal thin-slab rendering showing metallic darts (appear as bright flashes) deposited at preplanned biopsy sites [42].
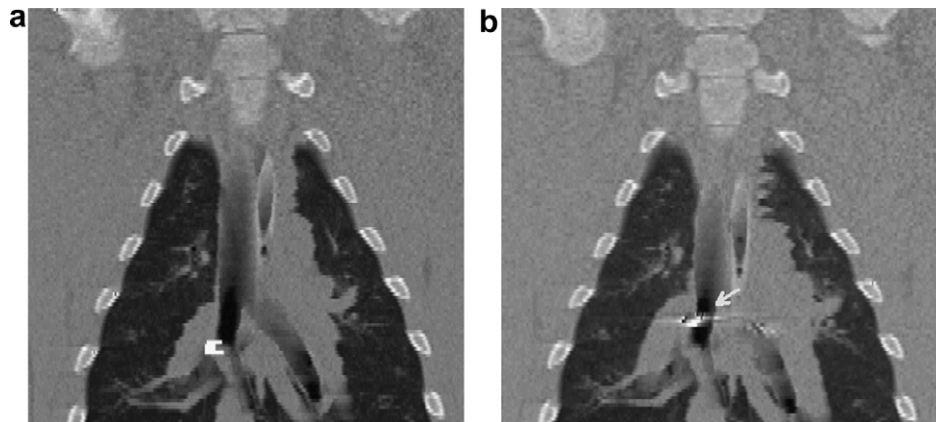
Fig. 10. Coronal slab views (based on depth-weighted maximum slab over 15 slices [42]) comparing the CT scan *before* the bronchoscopy procedure and a second CT scan *after* the platinum darts were in place. (a) First target ROI in "before" scan. (b) View of dart (bright flash indicated by arrow) deposited at site of first ROI in "after" scan.

"Distance from ROI" is defined as the distance from the closest virtual-site voxel to the closest dart point. Despite the motion artifacts in these tests (from breathing and the beating heart), the performance is excellent, especially considering the typical sizes of ROIs of interest ($\approx$1 cm in diameter).

5.3. Human studies

This section reports on the efficacy of the system for human lung-cancer assessment. For an initial pilot series of 11 subjects, the following procedures were done, per Fig. 1 and Section 2. After giving informed consent, the patient underwent a 3D MDCT scan using a Philips Mx8000 scanner (patient holds breath for 20 s during the scan). The image data were downloaded to the computer and 3D CT Assessment was done. During this stage, a technician ran automated procedures to segment the airway tree and define the central axes of the major airways. Next, the physician assisted in locating and defining target ROIs. For this pilot study, the ROIs were suspect mediastinal (central chest) lymph nodes. The technician then completed the case study by running an automated procedure for computing airway-tree and ROI polygon data.

After 3D CT Assessment, the computer was brought to the bronchoscopy laboratory and interfaced to the bronchoscope's video feed through a standard video output. During the procedure, a technician would interact with the computer to provide visual feedback to the physician. These interactions consisted of loading the appropriate guidance path for a given ROI and presenting sites along the path toward the ROI to the physician. Using the image guidance provided by the system, the physician maneuvered the bronchoscope to the ROI (a lymph node) and performed the biopsy. The biopsy samples were analyzed during the procedure by an on-site cytopathologist.

Fig. 11 shows a computer screenshot during a procedure. The original 3D MDCT scan was done on a 4-detector Philips Mx8000 MDCT scanner, with sample spacings

$\Delta x = \Delta y = 0.59$ mm and $\Delta z = 0.60$ mm; the scan consists of 479 slices, each 512×512. Extensive display capability is provided to the physician by the various graphical tools to give a much fuller vision of the procedural circumstances. In the figure, the upper right view shows a weighted-sum projection of 3D CT image data, plus the projected red central axes. The current chest location of the guidance system is marked with a blue ball. The upper center and right views show a transverse 2D CT slice (mediastinal viewing window: window level = 40, window width = 400) and a coronal front-to-back thin-slab rendering (focus = 30, vision = 40) [42]; the cross-hairs and red ball in these views indicate current 3D chest location. The lower left 3D surface rendering depicts the airway tree, extracted central axes (red lines), and current green biopsy site; a needle in this view indicates current 3D position and viewing direction.

The lower right Video Match view is the crux of the guidance and image fusion. It shows the following. Left: live video at current 3D location. Center: registered CT-based endoluminal rendering with green biopsy site in view and red-line guidance path—note that the green site is *outside* the airway. Right: video view with CT-based biopsy site fused onto it. Distance information is also provided by this tool. For example, "airway to ROI surface" indicates the system's current closest distance from the airway-wall surface to the ROI's surface. In the figure, this distance is 3.3 mm, indicating that a biopsy needle need only travel 3.3 mm through the airway wall to puncture the ROI surface. The measure "Dist to ROI" indicates the current total closest distance of the system to the ROI centroid and states how far the bronchoscope currently is from the biopsy site's centroid. Distances from any visible point in the rendered view can be found by moving the computer mouse over the rendered view. As can be seen in the figure, the previously unobservable biopsy ROI now appears as a large green object—the biopsy site is hard to miss in this view! When this view is coupled with the distance information, it gives the physician considerable added confidence in performing a previously blind biopsy.
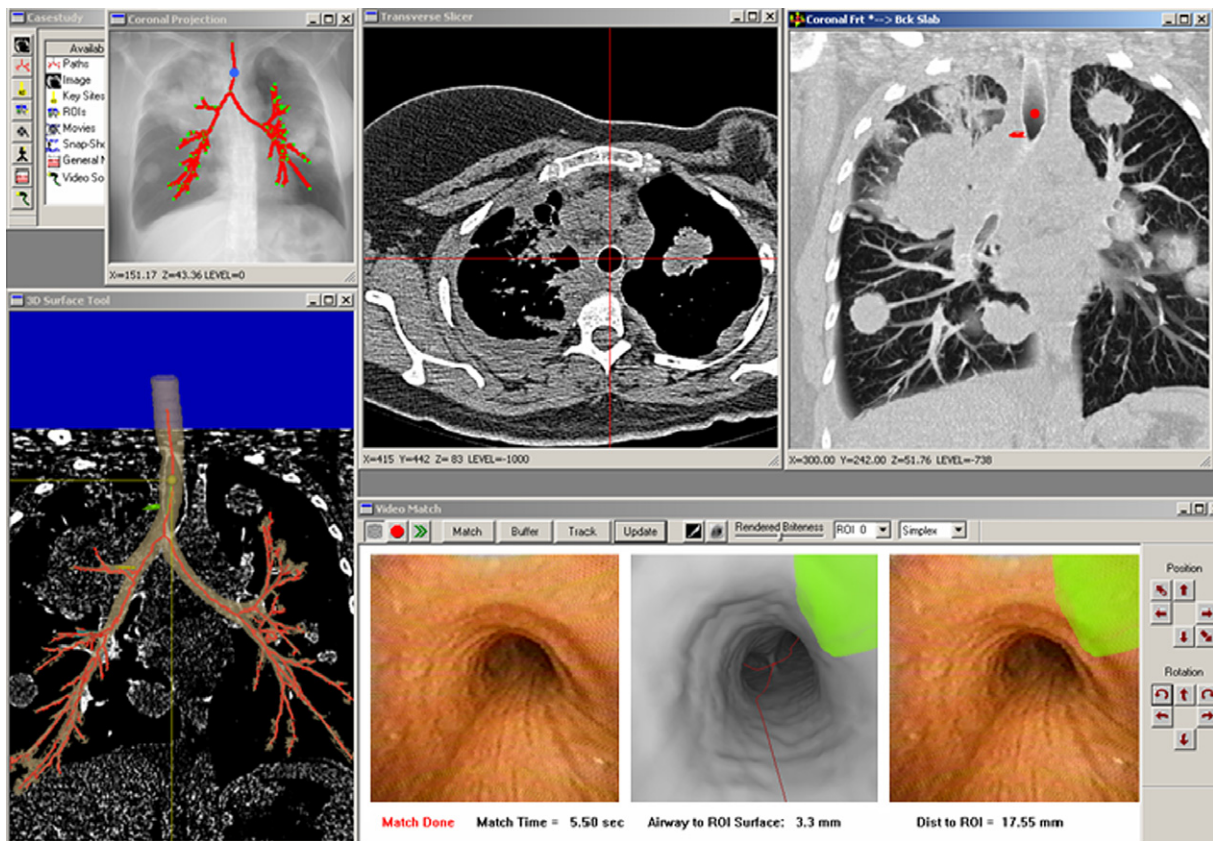
Fig. 11. System view during Stage-2 Image-Guided Bronchoscopy for a human lung-cancer case.

A major feature of the system involves the constant real-time updating of the complete system during navigation: as the physician is led to a biopsy site along a navigation path, all active viewing tools automatically update and follow in synchrony the same registered 3D position and viewing direction.

Fig. 12 focuses on the Video Match tool for another human lung-cancer case. In the figure, the physician has introduced a biopsy needle into one of the bronchoscope's working channels. The needle is shown as the physician is about to make a biopsy of a target site. Since the needle is bright, it has minimal impact on the registration (the registration emphasizes dark areas).

The physicians involved in the studies noted that the system's considerable extra visual feedback greatly helped in making biopsy decisions. This helped to lower the stress
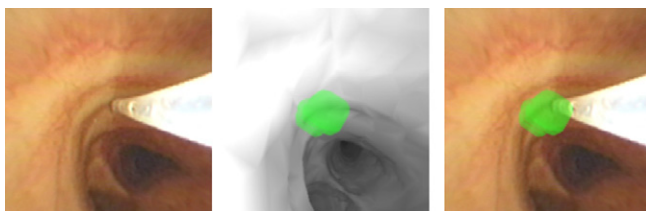

Fig. 12. Video Match tool view during actual biopsy for human case DC. Note the needle in the field of view piercing the virtual biopsy site in the far-right bronchoscopic video frame.

associated with the procedure, over that of the standard procedure. This stress arises, because the physician could conceivably puncture major blood vessels, such as the aorta; this could result in a serious life-threatening event. Blood vessels are also situated outside the airway interior and are not visible in the video. These vessels can be viewed in a computer-vision sense as occlusions to avoid. Note again that with the standard procedure, the physician must essentially make biopsy-location decisions blindly to avoid blood vessels.

During our tests, there were no complications or life-threatening events. An average of 4.3 biopsy attempts were made per ROI. In general >1 biopsy attempt is required per suspect ROI, because a given suspect region may not consist entirely of malignant material; the multiple biopsies enable multiple "approaches" to the region. The rate of return of diagnostic material from the biopsies was 60%, a far greater rate than the 20–30% range typically noted for the standard procedure [3]. Further clinical studies are ongoing.

6. Discussion

A vital assumption made is that the "real world" 3D space captured by the bronchoscopic video is in synchrony with the "virtual world" 3D space depicted by the CT renderings. This implies that the physician "cooperates" when

he moves the scope close to a biopsy site. Through the detailed studies presented here, we have found that the registration technique behaves robustly over a wide range of translations and rotations, giving the physician much leeway in the level of cooperation.

The system's registration procedure is only run at discretely selected sites along a path toward a target biopsy site. The bulk of the computing time is taken up by the large number of endoluminal renderings that must be computed during a typical optimization. More efficient implementation can speed up this procedure. Note that the physician spends far more time during a procedure deciding on when precisely to perform a needle biopsy and on replacing needles for subsequent biopsies. We have observed an occasional misregistration during a few of the tests. But misregistration is very easily corrected by slightly adjusting the starting point of the CT-based rendering and redoing the registration. Since the bronchoscope readily stays in a stationary position—the device is generally well anchored by the nose, trachea, and any other airways it has been guided through—the act of registering again is straightforward.

We have also found that the time difference between when the CT scan is done and when the procedure is performed also does not have an impact. In addition, the difference in inspiration level (how full the lungs are with air) between when the CT scan is done and when bronchoscopy is performed also does not have a significant effect. This appears to be true, because the airways are relatively rigid and bony, and their local relative shape and size do not change during inspiration.

In addition to the phantom, animal, and human studies presented here, the system has been successfully used on nearly 30 other human lung-cancer patients to date. The system increases the physician's vision of the procedure circumstances, greatly eases decision making and reduces stress, and appears to increase biopsy success rate. Most notably, the system appears to nullify the skill-level difference between different physicians, while also improving accuracy. The system has also been successfully applied to the problem of examining human airway obstructions [27].

The computer system interface requires little effort to use and greatly augments the physician's vision during a procedure. A technician performs all tasks except biopsy-site selection, which requires the physician's expertise. Thus, the system adds no new work burden, as biopsy-site planning must be done in any event. In fact, the system fits seamlessly into the current work flow of the patient's lung-cancer management team.

Further work can be done in a number of areas. While we have devised a large number of visual and quantitative tools for the system, we have by no means optimized their use or studied their efficacy. Such studies could benefit radiologists who analyze now common high-resolution 3D CT chest scans. Specific analysis and visualization protocols could be designed for planning airway stent design and insertion, guiding laser ablation, and performing treatment. Finally, work is needed in segmenting the diagnostic sites of interest, such as the hilar and mediastinal lymph nodes and suspect cancer nodules, and then integrating them with the system.

References

[1] R. Greenlee, M. Harmon-Hill, T. Murray, M. Thun, Cancer statistics, 2001, CA Cancer J. Clin. 51 (1) (2001) 15–36.

[2] A.D. Sihoe, A.P. Lim, Lung cancer staging, J. Surg. Res. 117 (1) (2004) 92–106.

[3] H.P. McAdams, P.C. Goodman, P. Kussin, Virtual bronchoscopy for directing transbronchial needle aspiration of hilar and mediastinal lymph nodes, Am. J. Roentgen. (1998) 1364–1381.

[4] K. Hopper, T. Lucas, K. Gleeson, J. Stauffer, R. Bascom, D. Mauger, R. Mahraj, Transbronchial biopsy with virtual CT bronchoscopy and nodal highlighting, Radiology 221 (2) (2001) 531–536.

[5] P.M. Boiselle, K.F. Reynolds, A. Ernst, Multiplanar and three-dimensional imaging of the central airways with multidetector CT, Am. J. Roentgen. 179 (2002) 301–308.

[6] C.S. White, E.A. Weiner, P. Patel, E.J. Britt, Transbronchial needle aspiration: guidance with CT fluoroscopy, Chest 118 (6) (2000) 1630–1638.

[7] H. Minami, Y. Ando, F. Nomura, S. Sakai, K. Shimokata, Interbronchoscopist variability in the diagnosis of lung cancer by flexible bronchoscopy, Chest 105 (2) (1994) 1658–1662.

[8] I. Bricault, G. Ferretti, P. Cinquin, Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy, IEEE Trans. Med. Imaging 17 (5) (1998) 703–714.

[9] W.E. Grimson, G.J. Ettinger, S.J. White, T. Lozano-Perez, W.E. Wells III, R. Kikinis, An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization, IEEE Trans. Med. Imaging 15 (2) (1996) 129–140.

[10] C.R. Maurer Jr., J.M. Fitzpatrick, M.Y. Wang, R.L. Galloway Jr., R.J. Maciunas, G.S. Allen, Registration of head volume images using implantable fiducial markers, IEEE Trans. Med. Imaging 16 (4) (1997) 447–462.

[11] Y. Sato, M. Nakamoto, Y. Tamaki, T. Sasama, I. Sakita, Y. Nakajima, M. Monden, S. Tamura, Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization, IEEE Trans. on Med. Imaging 17 (5) (1998) 681–693.

[12] J.D. Stefansic, A.J. Herline, Y. Shyr, W.C. Chapman, J.M. Fitzpatrick, B.M. Dawant, R.L. Galloway Jr., Registration of physical space to laparoscopic image space for use in minimally invasive hepatic surgery, IEEE Trans. Med. Imaging 19 (10) (2000) 1012–1023.

[13] S.B. Solomon, P. White Jr., C.M. Wiener, J.B. Orens, K.P. Wang, Three-dimensionsal CT-guided bronchoscopy with a real-time electromagnetic position sensor: a comparison of two image registration methods, Chest 118 (6) (2000) 1783–1787.

[14] Y. Schwarz, A.C. Mehta, A. Ernst, F. Herth, A. Engel, D. Besser, H.D. Becker, Elecromagnetic navigation during flexible bronchoscopy, Respiration 70 (5) (2003) 515–522.

[15] W.E. Lorensen, F.A. Jolesz, R. Kikinis, The exploration of cross-sectional data with a virtual endoscope, Interactive Technology and the New Health Paradigm (1995) 221–230.

[16] D.J. Vining, K. Liu, R.H. Choplin, E.F. Haponik, Virtual bronchoscopy: relationships of virtual reality endobronchial simulations to actual bronchoscopic findings, Chest 109 (2) (1996) 549–553.

[17] P. Rogalla, J. Van Scheltinga, B. Hamm, Virtual Endoscopy and Related 3D Techniques, Springer-Verlag, Berlin, 2002.

[18] W.E. Higgins, K. Ramaswamy, R. D Swift, G. McLennan, E.A. Hoffman, Virtual bronchoscopy for 3D pulmonary image assessment: State of the art and future needs, Radiographics 18 (3) (1998) 761–778.

[19] E.F. Haponik, S.L. Aquino, D.J. Vining, Virtual bronchoscopy, Clin. Chest Med. 20 (1) (1999) 201–217.

[20] R.M. Summers, N.R. Aggarwal, M.C. Sneller, M.J. Cowan, B.J. Wood, C.A. Langford, J.H. Shelhamer, CT virtual bronchoscopy of the central airways in patients with Wegener's granulomatosis, Chest 121 (1) (2002) 242–250.

[21] P. Turcza, M. Duplaga, Navigation systems based on registration of endoscopic and CT-derived virtual images for bronchofiberscopic prcoedures, in: M. Duploga et al. (Eds.), Transformation of Health Care with Information Technologies, IOS Press, 2004, pp. 253–263.

[22] K. Mori, D. Deguchi, J. Sugiyama, Y. Suenaga, J. Toriwaki, C.R. Maurer, H. Takabatake, H. Natori, Tracking of bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images, Med. Image Anal. 6 (2002) 321–336.

[23] R.S. Wright Jr., B. Lipchak, OpenGL Super Bible, 3rd ed., SAMS Publishing, 2005.

[24] W. Schroeder, K. Martin, B. Lorensen, The Visualization Toolkit: An Object-Oriented Approach To 3D Graphics, Prentice Hall, Upper Saddle River, NJ, 1997.

[25] A.P. Kiraly, W.E. Higgins, E.A. Hoffman, G. McLennan, J.M. Reinhardt, 3D human airway segmentation methods for virtual bronchoscopy, Acad. Radiol. 9 (10) (2002) 1153–1168.

[26] R.D. Swift, A.P. Kiraly, A.J. Sherbondy, A.L. Austin, E.A. Hoffman, G. McLennan, W.E. Higgins, Automatic axes-generation for virtual bronchoscopic assessment of major airway obstructions, Comput. Med. Imaging Graph. 26 (2) (2002) 103–118.

[27] A.P. Kiraly, J.P. Helferty, E.A. Hoffman, G. McLennan, W.E. Higgins, 3D path planning for virtual bronchoscopy, IEEE Trans. Med. Imaging 23 (11) (2004) 1365–1379.

[28] C. Studholme, D.L.G. Hill, D.J. Hawkes, An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition 32 (1) (1999) 71–86.

[29] E. Trucco, A. Verri, Introductory Techniques for 3-D Computer Vision, Prentice-Hall, Upper Saddle River, NJ, 1998.

[30] T. Okatani, K. Deguchi, Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center, Computer Vision and Image Understanding 66 (2) (1997) 119–131.

[31] K.V. Asari, S. Kumar, D. Radhakrishnan, A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation, IEEE Trans. Med. Imaging 18 (4) (1999) 345–354.

[32] J.P. Helferty, C. Zhang, G. McLennan, W.E. Higgins, Videoendoscopic distortion correction and its application to virtual guidance of endoscopy, IEEE Trans. Med. Imaging 20 (7) (2001) 605–617.

[33] P. Viola, W.M. Wells III, Alignment by maximization of mutual information, Int. J. Comp. Vis. 24 (2) (1997) 137–154.

[34] F. Maes, A. Colligeon, D. Vandermeulen, G. Marchal, P. Suetens, Multimodality image registration by maximization of mutual information, IEEE Trans. Med. Imaging 16 (2) (1997) 187–198.

[35] J.P. Helferty, A.J. Sherbondy, E.A. Hoffman, G. McLennan, 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.

[36] J.P. Helferty, W.E. Higgins, Technique for registering 3D virtual CT images to endoscopic video, IEEE Inter. Conf. Image Process. (2001) 893–896, Oct. 7-10 2001.

[37] J.P. Helferty, E.A. Hoffman, G. McLennan, and W.E. Higgins, CT-video registration accuracy for virtual guidance of bronchoscopy, in: SPIE Medical Imaging 2004: Physiology, Function, and Structure from Medical Images, vol. 5369, A. Amini, and A. Manduca, eds., pp. 150–164, 2004.

[38] J.A. Nelder, R. Mead, A simplex method for function optimization, Comput. J. 7 (1965) 308–313.

[39] W.H. Press, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, 1994.

[40] J.P. Helferty, Image-Guided Endoscopy and its Application to Pulmonary Medicine, Ph.D. thesis, Penn State University, May 2002.

[41] B.Q. Tran, J.K. Tajik, R.A. Chiplunkar, E.A. Hoffman, Lung volume control for quantitative x-ray CT, Ann. Biomed. Engin. 24 (Supplement 1) (1996) S66.

[42] J.Z. Turlington, W.E. Higgins, New techniques for efficient sliding thin-slab volume visualization, IEEE Trans. Med. Imaging 20 (8) (2001) 823–835.