SINGLE-VIEW 2D/3D REGISTRATION FOR X-RAY GUIDED BRONCHOSCOPY

Di Xu1, Sheng Xu3, Daniel A. Herzka1, Rex C. Yung2, Martin Bergholdt4, Luis F. Gutiérrez5, Elliot R. McVeigh1

1Department of Biomedical Engineering, Johns Hopkins University, USA
2Pulmonary Oncology, Johns Hopkins Medical Institutions, USA
3Department of Image Guided Technology, Philips Research North America, USA
4Department of Digital Imaging, Philips Research Europe, Germany
5Image Guided Interventions, Philips Healthcare, USA

ABSTRACT
X-ray guided bronchoscopy is commonly used for targeting peripheral lesions in the lungs which cannot be visualized directly by the bronchoscope. The airways and lesions are normally not visible in X-ray images, and as a result, transbronchial biopsy of peripheral lesions is often carried out blindly, lowering the diagnostic yield of bronchoscopy. In response to this problem, we propose to superimpose the lesions and airways segmented from preoperative 3D CT images onto 2D fluoroscopic images. A feature-based 2D/3D registration method is used for image fusion between the two datasets. The algorithm extracts features of the bony structures from both CT and X-ray images to compute the registration. Phantom and clinical studies were carried out to validate the algorithm's performance, showing an accuracy of 3.48±1.38mm. The convergence range and speed of the algorithm were also evaluated to investigate the feasibility of using the algorithm clinically. The results are presented.

Index Terms— fluoroscopic navigation, bronchoscopy, feature-based 2D/3D image registration, surface mesh render, distance map

1. INTRODUCTION
Lung cancer is the leading cause of cancer death in the world, with over 160,000 deaths every year in the US alone, and an additional 215,000 new cases each year [1]. Minimally invasive biopsies can be performed through the airways using a bronchoscope inserted into the patient’s mouth. However, the airways are small in the mid and lower lung, making it difficult to use a bronchoscope to reach peripheral lesions. Physicians often have to turn to X-ray fluoroscopy to see biopsy devices which are extended beyond the field of view of the bronchoscope camera. Due to the low image contrast of fluoroscopic images, often neither the lesions nor the airways are visible. As a result, bronchoscopic biopsies of peripheral lesions are obtained in a blind way, leading to a significant false negative rate [2].

CT is often used in lung cancer diagnosis. Both the airways and lesions of the lungs can be segmented under high resolution CT. Though people have used CT to directly guide interventions, concerns about procedure cost and increased radiation exposure make it less desirable. One possible solution in image guidance of bronchoscopy is to fuse the 3D CT volume with 2D fluoroscopic images and overlay the lesion and airways on the fluoroscopic images so that both the anatomic and diagnostic information from CT can be brought to the interventional procedure. The alignment between the CT volume and fluoroscopic images can be achieved by a 2D/3D registration. Image fusion of CT and fluoroscopic images can provide intra-operative image guidance for the bronchoscopic biopsy procedure, with the aim of precisely targeting peripheral lung lesions.

The registration algorithms can be grouped into two categories: intensity based and feature based [3].

Intensity based registration is frequently used for registration of bony structures between CT and fluoroscopic images [4]. This class of algorithms projects the CT volume onto a virtual 2D space to generate a digitally reconstructed radiograph (DRR) which is compared with real fluoroscopic images for registration. The main limitation for applying intensity-based registration algorithms in real-time intervention is the algorithm’s speed, especially the speed it takes to generate DRR [5]. Many methods have been reported to accelerate DRR generation, including a splatting method rather than ray-casting for fast DRR, and DRR generation with a graphics processing unit (GPU) [6,7], but speed is still a bottleneck for using intensity-based algorithms for interventional guidance.

The feature-based approaches match corresponding points of the two modalities [8]. The performance of feature-based registration highly depends on the robustness of feature point detection [5]. Currently, feature-based registration algorithms are commonly used for neural and cardiac applications [9,10], where the center lines of vascular structures from both CT/MR and digital subtraction angiography (DSA) image can be robustly segmented [11]. Some applications use distance map technique to accelerate...
registration speed. However, no contrast agent is used during X-ray guided bronchoscopic procedures so DSA is not an option. Therefore, the bronchial tree cannot be directly segmented in X-ray images, making it difficult to use the center lines of the airways for feature-based registration. To address this problem, we propose to extract the edges of bony structures for the registration between the CT and X-ray images, allowing the segmented airways to be overlaid on the X-ray images during the intervention.

2. METHODS

Mobile C-arms are commonly used in X-ray guided bronchoscopy. Information about the angle of the fluoroscope and the table position is typically not available during the intervention. Therefore, we propose a single-view 2D/3D registration algorithm, as described in the following steps.

2.1. Bone segmentation in CT and surface mesh generation

Bony structures are often used in 2D/3D registration since they have good contrast in X-ray images and can be robustly segmented. As shown in Figure 3.a, the ribs (sometimes also spine) are typically included in the X-ray images of bronchoscopy. Therefore, it is desirable to segment the bones in both CT and X-ray images which can be used as registration features. The bones are first segmented in the CT image by applying an intensity threshold. Triangle surface meshes of the bones are then generated from the segmentation results [12]. Since the CT image is acquired before the bronchoscopy procedure, this step is done offline.

2.2. Projection of surface mesh to virtual image plane

After the ribs and spine are segmented in the CT images, each triangle of the surface mesh is projected onto a virtual 2D image plane [13]. The position and orientation of the virtual image in the X-ray coordinate system is the estimated pose of the fluoroscope’s intensifier and the starting point of the 2D/3D registration, which can be obtained using prior knowledge of the patient’s pose with limited manual adjustment. The result of the mesh projection is a binary image, with pixels set to 1 if they are inside the projections of one or more triangles. The boundaries of the surface mesh projections are then extracted from the virtual 2D image and will be matched to the bone segmentation of the X-ray image.

2.3. Canny edge detection and distance transform

Canny edge detection is then conducted on the fluoroscopic image [14], yielding a binary image with a set of feature points representing the rib and spine boundaries. A distance transform was then applied to the binary image. For every pixel of the fluoroscopic image, the distance from the pixel to the closest edge point detected by the Canny filter is calculated. A distance map, \( f(x) \), is generated to store the distance for every pixel, where \( x \) is the coordinates of the pixel in the fluoroscopic image.

2.4. Cost function and optimization

The CT and X-ray images can be registered by matching the boundaries of the projected CT bone segmentations to the edges detected in the fluoroscopic image. After the bone segmentations of the CT image are projected to the X-ray image plane, the boundaries of the projections can be represented by a set of points in the X-ray image coordinate system, \( \mathbf{x} = \{x_1, x_2, ..., x_n\} \). A Euclidean distance-based similarity measure can be obtained for a given transformation between the CT image and the X-ray system by calculating the distance from each projected boundary point to the closest Canny edge of the fluoroscopic image. Since the distance of each pixel has already been computed in the distance transform of section 2.3, the cost function is given by

\[
D = \frac{1}{n} \sum_{i=1}^{n} f(x_i)
\]

where \( x_i \) is a point from the point set \( \mathbf{x} = \{x_1, ..., x_n\} \), extracted from the boundaries of the surface projections; and \( f(x) \) represents the distance map computed from the fluoroscopic image.

Numerical optimization using the Powell algorithm [15, 16] is used to minimize the cost function to find the best rigid-body transformation that aligns the CT and X-ray images. Since the optimization is an iterative procedure, the cost function is evaluated many times during the optimization. The algorithm is very efficient because the distance map only needs to be calculated once at the beginning of the optimization. It is not necessary to compute the distance between the projected boundary and its closest Canny edge in every iteration of the optimization.

3. EXPERIMENTS AND RESULTS

3.1. Phantom evaluation

The phantom study was conducted at the Johns Hopkins University on an Allura Xper FD20 fluoroscope system (Philips Healthcare, Best, the Netherlands) as shown in Figure 1. Four fiducial markers were attached to the phantom (Elastrat, Geneva, Switzerland) for quantitative validation. The fiducial markers were manually identified in the CT volume coordinates and in two X-ray images.
acquired at different view angles. The 3D X-ray coordinates were reconstructed from the biplane pair using epipolar geometry. The ground truth transformation between the CT image and the X-ray system was then calculated using point based rigid-body registration. (Note that the C-arm used in phantom validations is a high-end fluoroscopic system that is typically not available for bronchoscopy procedures, but it was used to evaluate the registration algorithm because it provided calibrated view geometries that were used to establish the ground truth for the 2D/3D registration.)

Table 1. Convergence range (manually explored) and average accuracy of the registration algorithm. The registration parameters were defined relative to the X-ray image coordinates, so x- and y-parameters are in-plane and z-parameters are through-plane.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Upper Convergence Range</th>
<th>Lower Convergence Range</th>
<th>Average Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx</td>
<td>5.00°</td>
<td>-5.00°</td>
<td>4.06 mm</td>
</tr>
<tr>
<td>Ry</td>
<td>5.00°</td>
<td>-9.99°</td>
<td>4.52 mm</td>
</tr>
<tr>
<td>Rz</td>
<td>14.02°</td>
<td>-21.00°</td>
<td>3.89 mm</td>
</tr>
<tr>
<td>Tx</td>
<td>100.00 mm</td>
<td>-75.02 mm</td>
<td>4.02 mm</td>
</tr>
<tr>
<td>Ty</td>
<td>45.00 mm</td>
<td>-50.00 mm</td>
<td>1.48 mm</td>
</tr>
<tr>
<td>Tz</td>
<td>4.63 mm</td>
<td>-5.37 mm</td>
<td>4.71 mm</td>
</tr>
</tbody>
</table>

The accuracy of the algorithm was evaluated by comparing the distance of the four fiducial markers’ positions on the fluoroscopic image to their positions on the projected virtual image. Down-sampled 1.19/1.19/3 mm CT scan images and 1.19/1.19 mm X-ray images were used, 30 experiments conducted, resulting in an error of 3.48 ±1.38 mm (Where in-plane 3.21±1.54 mm, through-plane 4.34±1.30 mm). Figures 2.c and 2.d show a typical phantom result.

Table 1 shows the convergence range of the registration algorithm. The starting point of the registration was shifted away from the ground truth location for each of the six transformation parameters. The offset from starting point to ground truth starts from 5 mm (deg) then has 2-fold increase for every trial. The registration algorithm was then carried out trying to recover the ground truth. A registration was considered to be successful if the registration error was smaller than 5 mm.

Figure 2. Algorithm validation on phantom data. (a) is fluoroscopic image; (b) shows the result of Canny edge detection; The white edges are the projections of CT bone segmentations that are superimposed on the X-ray images. (c) is the initial position before the 2D/3D registration; (d) is the result after registration.

Figure 3. Algorithm validation on patient data. (a) is fluoroscopic image; (b) shows the result of Canny edge detection; The white edges are the projections of CT bone segmentations that are superimposed on the X-ray images. (c) is the initial position before the 2D/3D registration; (d) is the result after registration.
The speed of the algorithm was compared with the intensity-based algorithms that rely on DRR generation. Using the same phantom dataset and the same computer, the DRR generation time (in seconds) was 0.082, 0.441 and 3.241 for surface rendering, splatting and ray-casting algorithms respectively. The mean registration time for the proposed algorithm was 73 seconds as opposed to 189 and 510 seconds for the intensity-based algorithms by splatting and ray-casting DRR generation respectively. The intensity-based registration method used in our experiment is described in [5].

3.2. Clinical evaluation

The algorithm was also tested on patient data as shown in Figure 3. Since no ground truth was available in patient studies, visual inspection was used to evaluate the algorithm’s performance. Figure 3.c and 3.d show an example of the alignment before and after registration, respectively. The white edges are the projections of CT bone segmentations that are superimposed on the X-ray image.

4. CONCLUSIONS AND DISCUSSION

The single-view 2D/3D registration algorithm is feasible for X-ray guided bronchoscopy. The proposed algorithm shows good robustness for in-plane parameters (Rx, Ry and Tz). The performance of the algorithm is inferior for the three through-plane parameters (Rz, Tx and Ty). This result is expected because only a single view is used in the registration.

Another source of errors might be the different intrinsic patterns that virtual binary images and real X-ray images have. Therefore, detected feature points might have difference. Hence some errors might be involved.

Since the registration is designed for interventional use, it needs to be sufficiently fast so that procedure time will not be significantly prolonged. Our experiments show that the algorithm is much faster than both of intensity-based algorithms by splatting and ray-casting DRR generation. The proposed algorithm is faster and has a large capture range, although it sacrifices some through-plane accuracy in the process. This trade-off was motivated by the observation that in X-ray guided bronchoscopy, a rough alignment between the CT can X-ray images may already provide sufficient information for the pulmonologist to choose the correct respiratory airway branch for inserting the surgical devices. This is in contrast to today’s standard practice, in which no CT overlay is used, such that the physician has no visual feedback of the airways (see Figure 3.a). Therefore, we believe that our work can improve the clinical outcome of X-ray guided bronchoscopy by showing the patient’s airways corresponding to the X-ray image.

5. REFERENCES

[1] American Cancer Society, "What Are the Key Statistics About Lung Cancer?"