3D Image Fusion and Guidance for Computer-Assisted Bronchoscopy

William E. Higgins, Lav Rai, Scott A. Merritt, Kongkuo Lu, Nicholas T. Linger, and Kun-Chang Yu
Penn State University
Depts. of Electrical Engineering and Computer Science and Engineering
University Park, PA 16802, USA

Lung Cancer

- Lung Cancer: #1 cancer killer, 30% of all cancer deaths, 1.5 million deaths world-wide, < 15% 5-year survival rate (nearly the worst of cancer types)

- To diagnose and treat lung cancer,
  1) 3D CT-image preplanning – noninvasive
  2) Bronchoscopy – invasive

- 500,000 bronchoscopies done each year in U.S. alone

  → Procedure is LITTLE HELP if diagnosis/treatment are poor

- A test for CT Image-based Lung-Cancer Screening in progress!

  10-30 million patient population in U.S. alone!

  → Screening is WORTHLESS if diagnosis/treatment are poor
3D CT Chest Images

Typical chest scan $V(x,y,z)$:
1. 500 512X512 slices $V(x,y,z)$
2. 0.5mm sampling interval
3D Mental Reconstruction

➔ How physicians assess CT scans now
Visualization Techniques — see “inside” 3D Images

- multi-planar reconstruction
- projection imaging
- volume/surface rendering
- virtual endoscopic rendering
- curved-section reformatting

STSMIP: sliding-thin-slab maximum intensity projection

Bronchoscopy

For “live” procedures

I\textsubscript{V}(x,y)

Figure 19.4, Wang/Mehta ‘95
Difficulties with Bronchoscopy

1. Physician skill varies greatly!
2. Low biopsy yield. Many “missed” cancers.
3. Biopsy sites are beyond airway walls – biopsies are done blindly!

“Let’s just start cutting and see what happens.”
Virtual Endoscopy (Bronchoscopy)

- Input a high-resolution 3D CT chest image
  - virtual copy of chest anatomy
- Use computer to explore virtual anatomy
  - permits unlimited “exploration”
  - no risk to patient

Endoluminal Rendering

$I_{CT}(x,y)$ (inside airways)
Image-Guided Bronchoscopy Systems

Show potential, but recently proposed systems have limitations:

- CT-Image-based
  - McAdams et al. (AJR 1998) and Hopper et al. (Radiology 2001)
  - Bricault et al. (IEEE-TMI 1998)
- Electromagnetic Device attached to scope
  - Schwarz et al. (Respiration 2003)

→ Our system: reduce skill variation, easy to use, reduce “blindness”
Our System: Hardware

Endoscope
Scope Monitor
Scope Processor
Light Source
Computer display

PC Enclosure
Matrox PCI card
Video Capture
Main Thread
Video Tracking
OpenGL Rendering
Worker Thread
Mutual Information
Dual CPU System

Video AGP card
Rendered Image
Polygons, Viewpoint Image

AVI File
Video Stream

Software written in Visual C++.
Our System: Work Flow

Stage 1: 3D CT Assessment
1) Segment 3D Airway Tree
2) Calculate Centerline Paths
3) Define Target ROI biopsy sites
4) Compute polygon data

→ Case Study

Stage 2: Live Bronchoscopy
For each ROI:
1) Present virtual ROI site to physician
2) Physician moves scope “close” to site
3) Do CT-Video registration and fusion
4) Repeat steps (1-3) until ROI reached
Stage 1:  3D CT Assessment (Briefly)

1. Segment Airway tree  
   (Kiraly et al., Acad. Rad. 10/02)

2. Extract centerlines  
   (Kiraly et al., IEEE-TMI 11/04)

3. Define ROIs  
   (e.g., suspect cancer)

4. Compute tree-surface polygon data (Marching Cubes – vtk)

→ CASE STUDY to help guide bronchoscopy
Stage 2: Bronchoscopy - Key Step: CT-Video Registration

Register

Virtual 3D CT World

\[ I^i_{CT}(x, y) \] (Image Source 1)

Maximize normalized mutual information to get

\[ I_{CT}^o(x, y) \]

To the

Real Endoscopic Video World

\[ I^F_V(x, y) \] (Image Source 2)
CT-Video Registration: 1) Match viewpoints of two cameras

Both image sources, $I_V$ and $I_{CT}$, are cameras.

6-parameter vector representing camera viewpoint $\chi = (X, Y, Z, \alpha, \beta, \gamma)$

3D point $p = (X_p, Y_p, Z_p)$ mapped to camera point $(X_c, Y_c)$ through the standard transformation

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

The final camera screen point is given by $(x, y)$ where

$$
x = \frac{fX_c}{Z_c}, \quad y = \frac{fY_c}{Z_c}
$$
Bronchoscope Video Camera Model

Following Okatani and Deguchi (CVIU 5/97), assume video frame \( I(p) \) abides by a Lambertian surface model; i.e.,

\[
I(p) = \sigma \frac{L \cos \theta_s}{\pi R^2}
\]

where

\[ p = (X_p, Y_p, Z_p) \]

\( \theta_s = \) light source-to-surface angle

\( R = \) distance from camera to surface point \( p \)
Make FOVs of both Cameras equal

To facilitate registration, make both cameras $I_V$ and $I_{CT}$ have the same FOV.

To do this, use an endoscope calibration technique (Helferty et al., IEEE-TMI 7/01).

Measure the bronchoscope’s focal length (done off-line):

$$f = \frac{(x_r - x_l) Z_m}{(X_r - X_l)}$$

Then, the angle subtended by the scope’s FOV is

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

Use same value for endoluminal renderings, $I_{CT}$. 
Bronchoscope Calibration Device

- Capture known dot pattern.
- Compute Calibration parameters from frame.

<table>
<thead>
<tr>
<th>Endoscope</th>
<th>Olympus BF P200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern size</td>
<td>4.5” x 3”</td>
</tr>
<tr>
<td>Dot Diameter</td>
<td>.1”</td>
</tr>
<tr>
<td>Grid size</td>
<td>19 x 13</td>
</tr>
<tr>
<td>Distance</td>
<td>3”</td>
</tr>
</tbody>
</table>

CT (Endoluminal Rendering) Camera Model

Related to video frame model $I(p)$:

$$I_{CT}(p) = \frac{(\cos \phi_p)^{1/2} \cos \theta_s}{(1 + 0.0025R^2)} + L_a$$

where

$p = (X_p, Y_p, Z_p)$

$\theta_s =$ light source-to-surface angle

$\theta_p =$ angle between camera axis and vector pointing toward $p$

$R =$ distance from camera to surface point $p$ (from range map)

$L_a =$ ambient light (indirect lighting)

Use OpenGL
Normalized Mutual Information

*Mutual Information* (MI) – used for registering two different image sources:

a) Grimson *et al.* (*IEEE-TMI* 4/96)

b) Studholme *et al.* (*Patt. Recog.* 1/99) → normalized MI (NMI)
Normalized Mutual Information

Normalized mutual information (NMI):

\[ S_{NMI}(I_V, I_{CT}) = \frac{h(V) + h(CT)}{h(V, CT)} \]

where

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

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

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

and

\[ p_V(k) = \sum_{l=0}^{M-1} p_{V,CT}(k, l) \]

is a histogram (marginal density)
Given a fixed video frame $I^F_V(x, y)$ and starting CT view $I^i_{CT}(x, y)$

Search for the optimal CT rendering $I^o_{CT}(x, y)$ subject to

$$\chi_o = \arg \left\{ \max_{\chi \in N_{\chi_i}} \left[ S_{NMJ} (I^\chi_{CT}(x, y), I^F_V(x, y)) \right] \right\}$$

where viewpoint $\chi = (X, Y, Z, \alpha, \beta, \gamma)$ is varied over neighborhood $N_{\chi_i}$ about $\chi_i$

Optimization algorithms used: step-wise, simplex, and simulated annealing
A simplex is an N dimensional figure with N+1 vertices

Simplex Optimization

Reflection

\[ V^* = \frac{2}{N} \sum_{i=0; i \neq i_{HI}}^{N} V^i - V^i_{HI} \]

Expansion

\[ V^{**} = \frac{3}{N} \sum_{i=0; i \neq i_{HI}}^{N} V^i - 2V^i_{HI} \]

Collapse

\[ V^{***} = \frac{1}{2N} \sum_{i=0; i \neq i_{HI}}^{N} V^i - \frac{1}{2} V^i_{HI} \]

Collapse Simplex

\[ V^i = \frac{1}{2} V^i + \frac{1}{2} V^i_{LOW} \]
CT-Video Registration: Simplex Optimization

Calculate vertices of initial simplex (N=6):
\[ \chi_0 = \{X, Y, Z, \alpha, \beta, \gamma\} \] (initial viewpoint), \[ \chi_1 = \{X + \Delta X, Y, Z, \alpha, \beta, \gamma\}, \ldots, \chi_N = \{X, Y, Z, \alpha, \beta, \gamma + \Delta \gamma\} \]
iteration = 0

Do

Compute \( S_{NMI}(\chi_i), \quad i = 0, 1, \ldots, N \)
Note the best and worst vertices and associated NMI measures: \( \chi_{\text{min}}, \chi_{\text{max}}, S_{\text{min}}, S_{\text{max}} \)

Get reflection of \( \chi_{\text{min}} \) across face of simplex: \( \chi^* = (\frac{2}{N} \sum_{\chi_i \neq \chi_{\text{min}}} \chi_i) - \chi_{\text{min}} \)

If \( S_{NMI}(\chi^*) > S_{\text{max}} \)
    Get 2× reflection of \( \chi_{\text{min}} \) across face of simplex: \( \chi^{**} = (\frac{3}{N} \sum_{\chi_i \neq \chi_{\text{min}}} \chi_i) - 2\chi_{\text{min}} \)
    If \( S_{NMI}(\chi^{**}) > S_{\text{max}} \) \( \chi_{\text{max}} = \chi^{**} \)
    Else \( \chi_{\text{max}} = \chi^* \)

Else

Contract \( \chi_{\text{min}} \) \( \frac{1}{2} \) distance toward face of simplex: \( \chi^{***} = (\frac{1}{2N} \sum_{\chi_i \neq \chi_{\text{min}}} \chi_i) + \frac{1}{2}\chi_{\text{min}} \)

If \( S_{NMI}(\chi^{***}) > S_{\text{min}} \) \( \chi_{\text{min}} = \chi^{***} \)

Else contract all vertices \( \chi_i \neq \chi_{\text{max}} \) \( \frac{1}{2} \) distance toward \( \chi_{\text{max}} \): \( \chi_i = \frac{1}{2}(\chi_i + \chi_{\text{max}}) \)

iteration = iteration + 1

While \( (S_{\text{max}} - S_{\text{min}}) > \text{TOLERANCE} \) and iteration < \text{MAX_ITER} \)
CT-Video Optimization Example

Fixed Video Frame

Optimal CT Rendering

initial $p_{v,CT}$

optimal $p_{v,CT}$
System Results

Three sets of results are presented:

A. Phantom Test
   controlled test, free of subject motion

B. Animal Studies
   controlled in vivo (live) tests

C. Human Lung-Cancer Patients
   real clinical circumstances
A. Phantom Test

Goal: Compare biopsy accuracy under controlled stationary circumstances using (1) the standard CT-film approach versus (2) image-guided bronchoscopy.

Experimental Set-up:

- **Rubber phantom** - human airway tree model used for training new physicians.
- **CT Film** - standard form of CT data.
Computer Set-up during Image-Guided Phantom “Biopsy”
Phantom Accuracy Results (6 physicians tested)

Film biopsy accuracy: \(5.53\text{mm}\)  
Std Dev: \(4.36\text{mm}\)

Guided biopsy accuracy: \(1.58\text{mm}\)  
Std Dev: \(1.57\text{mm}\)

<table>
<thead>
<tr>
<th>Physician</th>
<th>film accuracy (mm)</th>
<th>guided accuracy (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.80</td>
<td>1.38</td>
</tr>
<tr>
<td>2</td>
<td>2.73</td>
<td>1.33</td>
</tr>
<tr>
<td>3</td>
<td>4.00</td>
<td>1.49</td>
</tr>
<tr>
<td>4</td>
<td>8.87</td>
<td>1.60</td>
</tr>
<tr>
<td>5</td>
<td>8.62</td>
<td>2.45</td>
</tr>
<tr>
<td>6</td>
<td>3.19</td>
<td>1.24</td>
</tr>
</tbody>
</table>

→ ALL physicians improved greatly with guidance

→ ALL performed nearly the SAME with guidance!
B. Animal Studies

**Goals:** Test the performance of the image-guided system under controlled in vivo circumstances (breathing and heart motion present).

**Experimental Set-up:**

![Biopsy dart](image1)

![Computer system during animal test (done in EBCT scanner suite)](image2)
Animal Study - Stage 2: Image-Guided Bronchoscopy

Coronal Weighted-Sum Projection of 3D CT Scan with centerline to biopsy site.

Position of dart after biopsy made.

Positions of rendered preplanned site (from guidance system) and dart after biopsy made.

Rendered airway tree. Green ball and virtual needle indicate 3D position. Light-blue rendered object is target biopsy site.
Composite View after All Real Biopsies Performed

Thin-slab DWmax depth-view of 3D CT data AFTER all darts deposited at predefined sites. Bright “flashes” are the darts.
C. Human Studies
Stage 2: Image-Guided Bronchoscopy

Real-World target video $I_V$

Virtual-World CT rendering $I_{CT}$

Registered Virtual ROI on Video

(case h005 [UF], mediastinal lymph-node biopsy, in-plane res. = 0.59mm, slice spacing = 0.60mm)
Animal Study - Stage 1: 3D CT Assessment
Stage 2: Image-Guided Bronchoscopy
Case DC: performing a biopsy

Left view: Real-time bronchoscopic video view; biopsy needle in view
Center: Matching virtual-bronchoscopic view showing preplanned region (green)
Right: Preplanned region mapped onto bronchoscopic view, with biopsy needle in view.

Distance to ROI = scope’s current distance from preplanned biopsy site (ROI).

- All nodal-region samples showed normal appearing lymphocytes.
- Subsequent open-lung biopsy showed a suspect mass to be inflammatory tissue.

⇒ 40 lung-cancer patients done to date
Case UF: approaching a biopsy site

Left view: Real-time bronchoscopic video view
Center: Matching virtual-bronchoscopic view showing preplanned region (blue); red line is preplanned guidance path
Right: Preplanned region mapped onto real bronchoscopic view.

Distance Measures:
Min Airway to ROI Surface = distance of closest airway surface point to ROI (target lymph node)
Distance to ROI Center = cursor’s current distance from center of preplanned biopsy site (ROI).
Airway Surface = distance of cursor to airway surface
Airway to ROI Surface = distance of airway surface point that cursor is on to ROI.
Guidance to Peripheral Lung-Cancer Nodules – In Progress
Real-Time Image-Guided Bronchoscopy – In Progress
Real-Time Image-Guided Bronchoscopy – In Progress

Error = 1.479236

Error = 10414.725680
Comments on System

- Effective, easy to use
  - A technician – instead of $$ physician – performs nearly all operations
- Gives a considerable “augmented reality” view of patient anatomy
  - less physician stress
- Fits seamlessly into the clinical lung-cancer management process.
- Appears to greatly reduce the variation in physician skill level.

This work was partially supported by:
  NIH Grants #CA74325, CA91534, HL64368, and RR11800
  Whitaker Foundation, Olympus Corporation
Thank You!
Lung Cancer

- Lung Cancer: #1 cancer killer, 30% of all cancer deaths, 1.5 million deaths world-wide, < 15% 5-year survival rate (nearly the worst of cancer types)

- To diagnose and treat lung cancer,
  1) 3D CT-image assessment – preplanning, noninvasive
  2) Bronchoscopy – invasive

→ Procedure is LITTLE HELP if diagnosis/treatment are poor
Normalized Mutual Information

*Mutual Information* (MI) – used for registering two different image sources:

a) Grimson *et al.* *(IEEE-TMI 4/96)*

b) Studholme *et al.* *(Patt. Recog. 1/99) → normalized MI (NMI)*

We use normalized mutual information (NMI) for registration:

\[ S_{NMI}(I_V, I_{CT}) = \frac{h(V) + h(CT)}{h(V, CT)} \]