Automatic Segmentation and Centroid Detection of Skin Sensors for Lung Interventions

Kongkuo Lu,1,* Sheng Xu,2 Zhong Xue,3 and Stephen T. Wong3

ABSTRACT
Electromagnetic (EM) tracking has been recognized as a valuable tool for locating the interventional devices in procedures such as lung and liver biopsy or ablation. The advantage of this technology is its real-time connection to the 3D volumetric roadmap, i.e. CT, of a patient’s anatomy while the intervention is performed. EM-based guidance requires tracking of the tip of the interventional device, transforming the location of the device onto pre-operative CT images, and superimposing the device in the 3D images to assist physicians to complete the procedure more effectively. A key requirement of this data integration is to find automatically the mapping between EM and CT coordinate systems. Thus, skin fiducial sensors are attached to patients before acquiring the pre-operative CTs. Then, those sensors can be recognized in both CT and EM coordinate systems and used to calculate the transformation matrix. In this paper, to enable the EM-based navigation workflow and reduce procedural preparation time, an automatic fiducial detection method is proposed to obtain the centroids of the sensors from the pre-operative CT. The approach has been applied to 13 rabbit datasets derived from an animal study and eight human images from an observation study. The numerical results show that it is a reliable and efficient method for use in EM-guided application.

Keywords: EM tracking, percutaneous intervention, lung cancer, CAD, 3D CT imaging, segmentation

1. INTRODUCTION
Electromagnetic (EM) tracking has been recognized as a valuable tool for many interventional and surgical procedures such as percutaneous lung and liver biopsy and ablation1–3 and image-guided surgeries. The advantage of this technology sits in its real-time connection to a 3D volumetric roadmap, e.g. CT, of a patient’s anatomy while the intervention is performed, so as to improve clinical outcomes while reducing adjustment of interventional device, procedure time, and complications4,5.

EM-based guidance requires tracking of the tip of the interventional device so that the location of the device can be superimposed with pre-operative CT images. A key process in CT-EM integration is to find the mapping between EM and CT coordinate systems2,6. For this purpose, skin fiducial sensors, either passive or active, are attached while a pre-procedural or intra-procedural scan is acquired for guidance. The sensors will be detected by EM field automatically, but the user may have to browse the guidance scan to identify the location of each skin sensor manually. Usually, three to six EM tracked sensors are adequate for guiding the procedure. Considering the small size of the sensors, around 20mm in diameter, and the large volume size of the scan, it is not convenient to go through the entire scan to locate the sensor during the intervention. In addition, the user may rely on 2D multi-planar reconstruction (MPR) sectional images: transverse, coronal, and/or sagittal, to define the centroid of each sensor, as a result, the process might be time consuming and user-biased.

Automatic approaches are desired, but the design is not straightforward. In CT image, the intensity value, defined based on Hounsfield Unit (HU) and used to represent different tissues, of sensors share similar distribution range with bony structures, wires external to patient bodies, and other devices involved in the procedure. Thus the goal of this paper is to develop an automatic fiducial detection method, so that the EM-tracked sensors can be recognized in both CT and EM coordinate systems to calculate the transformation matrix; EM-based navigation workflow can be enabled; and the burden for manual fiducial selection is reduced.

1 Philips Research North America, Briarcliff Manor, NY 10510, USA
2 NIH, Bethesda, MD 20892, USA
3 The Methodist Hospital Research Institute, Houston, TX 77030, USA
*Correspondence: Email: kongkuo.lu@philips.com; WWW: http://www.philips.com/; Telephone: 281-221-3787; Fax: 713-441-8696.
The method will be discussed in detail in Section 2. In Section 3, we present results derived from a series of human 3D chest CT scans, acquired from real EM-guided percutaneous liver biopsy and ablation procedures. A conclusion is given in Section 4.

2. METHODS

The proposed method is conducted based on observation of physical features of skin sensors and their appearance in 3D CT scans. In our guidance images with fiducial sensors attached, the sensors usually appear in similar or brighter intensities than bony structures and appear in regular shapes, such as spheres or cylinders, in MPR CT images with unknown position and orientation. The proposed approach includes three stages: 1) Detect and exclude body area; 2) Find fiducial candidates using thresholding, morphological operations, connected-component analysis, and conditional filtering; 3) Post-processing using self-assessed region growing to expand individual candidate areas to maximally cover the true sensor region, and calculate centroids based on segmentation. The detailed fiducial detection method is discussed as follows.

2.1. Body Detection

Since skin sensors and bony structures are represented in a similar intensity range in CT (Fig.1), it is not likely to distinguish these two types of structures based on simple thresholding. This is a major challenge in designing automatic fiducial segmentation methods, as shown in Fig.2. However, the bony structures are covered by skin and muscle. Also, voxel intensities of air and background are usually < 400 HU, bony and metal structures are > 1200, while fat and soft-tissue structures are in between. In this work, the 16-bit intensity value of the scans is in a range of [0, 4095]. A region-growing based method is conducted here to first include structures whose voxel intensities are in [400, 1200] external to the body area, so as to define the boundary of body areas to exclude all the structures covered by the boundary. In addition to method design, it is highly desired that the approach is implemented within several second to save procedure preparation time. Thus, this step is conducted as below:

1. Subsampling. Let \( r \) be the subsampling rate defined based on the number of slices, \( N_s \), of the 3D CT image. In this implementation, we subsample the scan in half, \( r = 1/2 \), when the 100 < \( N_s \) <= 200, resulting subvolume \( I_{1/2} \); a quarter, \( r = 1/4 \), when 200 < \( N_s \) < 400, resulting \( I_{1/4} \); and \( r = 1/8 \) when \( N_s \) >= 400, resulting \( I_{1/8} \). The last case is seldom an option. This definition of \( r \) is sufficient in this application. This process is not necessary when \( N_s \) < 100. The volume of \( I_r \) is \( V_r = r^3 V_I \).

2. Region growing. Four seeds are selected randomly from four corners of the subvolume \( I_r \), \( r = 1, 1/2, 1/4, 1/8 \). The region growing is implemented to include voxels whose intensities, \( I(x, y, z) \), are either < 400 or > 1200. The remain voxels are considered as body. Fig.3(a) gives an example result, where intensity values of body voxels were set to 1024 for better demonstration.

3. Image recovery. This step is to mapping the result derived from \( I_r \) to original image size. Then the neighbor regions of each body voxel is processed, and qualified voxels are further included. The size of the neighborhood area, \( S_N \), is determined by the subsample rate, where \( S_N = 1/r \). This process is sufficient to recover the body detection result in original CT scan, because the body region is a large enough object with respect to the sample rate used. Fig.3(b) shows an example result of body detection recovered in original CT image.

2.2. Sensor Candidate Identification

Process in 2.1 results in a 3D volume \( I_1 \), where only structures external to body remain. This step is to identify strong sensor candidates and includes following four processes.

1. Thresholding. Use \( I_1 \) as input, the output volume \( I_2 \) is generated as below:

\[
I_2(x, y, z) = \begin{cases} 
1 & \text{if} \ I_1(x, y, z) > T \\
0 & \text{otherwise}
\end{cases}
\]

where \( T \) is the intensity threshold higher than the intensity of bony structures in CT. The output include sensors, metal structures, wires, and objects represented with high intensity values. Some bony structures located in arms or shoulders may also be included due to the setup of CT acquisition.
Figure 1. Histogram of voxel intensity based on 10 human abdomen CT scan. The intensity values are shifted to start from 0 HU. And each bin includes a range of 100 HU. Air and background voxels are < 400HU; soft-tissue structure, fat, and other organs are in [400,1200] HU; and bony structure, fiducial sensors, and other metal objects are > 1200 HU. The unit of the horizontal axis is 100 HU.

Figure 2. Examples of skin fiducials shown in a CT image. (a) 2D axial section display. Fiducial sensors, pointed by red arrows, and bony structures share common intensity range in CT images. (b) 3D volume rendering. Fiducials have regular size and shape, and wires are attached when active EM tracked fiducial sensors are used.

2. Morphological opening operation. This step is to: a) remove small objects or noise while keeping volumes of the objects; and b) separate sensor candidates from attached wires. The size of the spherical structural element is denoted as $B$ and is selected based on sensor size and spacial resolution of the CT scan. Fig.4 gives an example of the fiducial sensor candidate resulting from the thresholding step.

3. Connected component analysis. Isolated objects are labeled in this step.

4. Filtering. The labeled objects resulted from the previous step will be evaluated resulting the major parts of the true candidate sensors. Features used are: volume/size, distance to other candidates, and volume similarity. The skin sensors used usually have similar size and shape, and may appear in a pair. Thus, these criteria can be used to further isolate true sensor candidates.
2.3. Self-Assessed Region Growing

The sensor candidates derived from the previous steps may only include the major regions and not cover the entire fiducial area. Thus, this step is to expand a segmented area to cover the true sensor. Unlike Mendoza’s self-assessed region growing approach, a conditional morphological dilation is conducted here since the skin fiducial sensors have regular size and shape. This process expands the candidate regions iteratively. In each iteration $i$, one layer in 3D is added to the surface of a sensor candidate. When terminated, only the voxels with desired intensities will be included to define the final sensor segmentation. The details are presented below. Given an area $A$, the mean and standard deviation of the voxel intensities in $A$ are calculated by:

$$\mu_A = \frac{1}{|A|} \sum_{(x,y,z) \in A} I(x, y, z)$$  \hspace{1cm} (2)
\[ \sigma_A = \sqrt{\frac{1}{|A|-1} \sum_{(x,y,z) \in A} (I(x,y,z) - \mu_A)^2} \]  

(3)

where \(|A|\) is the number of voxels included by area \(A\), and \((x, y, z)\) denotes a voxel in the CT volume. For object \(O_l\), we first calculate the follows:

- \(S_i\) be the surface layer added at iteration \(i > 0\), while \(S_0 = O_l\).
- For each voxel \(I(x, y, z)\) included by \(S_i\), calculate \(\mu_{N_j}\) and \(\sigma_{N_j}\) using equations 2 and 3, where \(j = 1 \ldots |S_i|\) and \(N_j\) is voxel \(I_j(x, y, z)\)'s 26-connection neighborhood region.
- Calculate
  \[
  \bar{\mu}_{S_i} = \frac{1}{|S_i|} \sum_{1 \leq j \leq |S_i|} \mu_{N_j}
  \]
  \[
  \bar{\sigma}_{S_i} = \frac{1}{|S_i|} \sum_{1 \leq j \leq |S_i|} \sigma_{N_j}
  \]

(4)
(5)

- A non-linear mapping might be applied to the original image \(I\) when it is noisy:
  \[
  I'(x, y, z) = \left(1 + \exp \left(\frac{-I(x, y, z) - \mu_{S_0}}{K \cdot \sigma_{S_0}/3}\right)\right)^{-1}
  \]
  (6)

Then, the iterative process is conducted as below:

1. For iteration \(i = 0\), calculate \(\bar{\mu}_{S_0}\) and \(\bar{\sigma}_{S_0}\).
2. For iteration \(i = 1, \ldots, M\), define \(S_i = S_{i-1} \oplus B - S_{i-1}\), where \(\oplus B\) is a morphological dilation operation using a unit 6-connection structuring element \(B\). So \(S_i\) is a layer attached to the object boundary resulted from iteration \(i - 1\).
3. Then calculate \(\bar{\mu}_{S_i}\) and \(\bar{\sigma}_{S_i}\).
4. Calculate decision measure:
   \[
   M_i = |\sigma'_{S_i} - \sigma'_{S_0}|
   \]
   (7)

If \(S_0 \cup S_1 \cup \ldots \cup S_i\) defines the object well, voxel intensities vary significantly between \(S_i\) and \(S_0\), resulting large \(M_i\). Otherwise, \(M_i\) is small.

5. The process stops when \(M_i \ll M_{i-1}\) and the output is \(O_l = S_0 \cup S_1 \cup \ldots \cup S_{i-1}\). Otherwise, start another iteration, \(i = i + 1\).

6. Apply same process on each isolated sensor candidate.

Similar to the 4th step in Section 2.2, the size, shape, and similarity measures of the fiducial sensors are applied to results derived from Section 2.3 to validate the true candidates. Then, centroids are calculated for detected sensors. Fig.5 gives an example of fiducial detection results from human CT scans, while Fig.6 shows example results derived from a rabbit scan.
3. RESULTS

The proposed approach was used to segment fiducial sensors and calculate their centroids for 13 rabbit scans derived from an animal study conducted to evaluate the performance of a minimally invasive multimodality image-guided system developed by Philips Research North America and the Methodist Hospital Research Institute, Houston, TX, for planning, guiding, diagnosing and/or treating lung cancer.9 The approach is also applied to 8 human scans, derived from 6 liver biopsy/ablation cases at the Methodist Hospital and guided by PercuNav system, Image Guided Intervention, Philips Healthcare. The rabbit images were produced by a Siemens SPECT/CT scanner, with spatial resolution: $\Delta X = \Delta Y = 0.23 \sim 0.3$ mm and slice thickness $\Delta Z = 1.25$ mm. Four fiducial sensors were placed on the chest skin for each rabbit to enable EM tracking. The human scans were produced by a Philips Brilliance 16-detector CT scanner, with spatial resolution: $\Delta X = \Delta Y = 0.7 \sim 0.97$ mm, and slice thickness $\Delta Z = 1 \sim 3$ mm.

Fig. 6 gives typical results from a rabbit scan, with spatial resolution: $\Delta X = \Delta Y = 0.29$ mm and $\Delta Z = 1.25$ mm. Four fiducial sensors were segmented and the corresponding centroids were calculated and highlighted by red crosses. Fig. 5 gives example results from a human scan, with spatial resolution: $\Delta X = \Delta Y = 0.7$ mm and $\Delta Z = 1$ mm. Six active fiducial sensors were detected and highlighted.

To evaluate the results, a live-wire-based method was used to segment the fiducial sensors in each scan and the centroids were calculated.10,11 The distance between the centroids, from the automatic and semi-automatic approaches, is calculated as an error measurement. Table 1 summarizes the performance of the automatic fiducial detection approach using the error measurement. The overall error is $0.3 \pm 0.27$ mm. Similar measurement was applied to evaluate the performance of the proposed approach when applied to rabbit images (four fiducial sensors were used), and the errors are $0.18 \pm 0.07$ mm.

4. CONCLUSION

Percutaneous intervention is an efficient and safe procedure for diagnosis and treatment of lung and liver cancer. Recently, EM tracking has been recognized as an useful tool for guiding this type of procedures. A preliminary but vital step in EM tracking is to detect the centroids of skin sensors to enable automatic transformation between EM and 3D CT. To simplify the workflow of EM-guided procedures, an approach is proposed in this work based on physical and imaging features of the sensors to automatically detect the sensor centroids. We have applied the method to 13 rabbit scans derived from an animal studies and eight human chest CT scans from an observation study. The numerical results showed its efficacy and robustness. This approach can be easily extended to other applications where skin sensors need to be isolated for EM tracking or other purposes.
(1-a) Fiducial 1 shown in a transverse slice. (1-b) Fiducial 1 was segmented and the boundary voxels are shown in red in the slice. The centroid was calculated and highlighted by the red cross. (2-a,b), (3-a,b), and (4-a,b) are fiducials 2-4 and the corresponding segmentation results shown in transverse slices. The centroids are highlighted by red crosses.

REFERENCES

Table 1. Summary of error measurements of the proposed automatic fiducial detection approach. Six skin fiducial sensors were used in each scan. The error was calculated based on the difference of the centroids of the same sensor but derived from the automatic method and the live-wire method. The overall error is $0.3 \pm 0.27$ mm.

<table>
<thead>
<tr>
<th>Fiducial</th>
<th>Scan1</th>
<th>Scan2</th>
<th>Scan3</th>
<th>Scan4</th>
<th>Scan5</th>
<th>Scan6</th>
<th>Scan7</th>
<th>Scan8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.13</td>
<td>0.22</td>
<td>0.27</td>
<td>0.16</td>
<td>0.45</td>
<td>0.36</td>
<td>0.34</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.11</td>
<td>1.89</td>
<td>0.13</td>
<td>0.09</td>
<td>0.36</td>
<td>0.90</td>
<td>0.28</td>
</tr>
<tr>
<td>3</td>
<td>0.14</td>
<td>0.13</td>
<td>0.26</td>
<td>0.20</td>
<td>0.27</td>
<td>0.31</td>
<td>0.45</td>
<td>0.29</td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td>0.25</td>
<td>0.15</td>
<td>0.17</td>
<td>0.14</td>
<td>0.44</td>
<td>0.56</td>
<td>0.31</td>
</tr>
<tr>
<td>5</td>
<td>0.18</td>
<td>0.17</td>
<td>0.21</td>
<td>0.21</td>
<td>0.33</td>
<td>0.44</td>
<td>0.55</td>
<td>0.31</td>
</tr>
<tr>
<td>6</td>
<td>0.13</td>
<td>0.10</td>
<td>0.15</td>
<td>0.10</td>
<td>0.26</td>
<td>0.40</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>Mean</td>
<td>0.16</td>
<td>0.16</td>
<td>0.49</td>
<td>0.16</td>
<td>0.26</td>
<td>0.39</td>
<td>0.51</td>
<td>0.28</td>
</tr>
<tr>
<td>Std</td>
<td>0.03</td>
<td>0.06</td>
<td>0.69</td>
<td>0.04</td>
<td>0.13</td>
<td>0.05</td>
<td>0.22</td>
<td>0.05</td>
</tr>
<tr>
<td>Min</td>
<td>0.13</td>
<td>0.10</td>
<td>0.15</td>
<td>0.10</td>
<td>0.09</td>
<td>0.31</td>
<td>0.28</td>
<td>0.19</td>
</tr>
<tr>
<td>Max</td>
<td>0.20</td>
<td>0.25</td>
<td>1.89</td>
<td>0.21</td>
<td>0.45</td>
<td>0.44</td>
<td>0.90</td>
<td>0.31</td>
</tr>
<tr>
<td>Median</td>
<td>0.16</td>
<td>0.15</td>
<td>0.23</td>
<td>0.17</td>
<td>0.26</td>
<td>0.38</td>
<td>0.50</td>
<td>0.30</td>
</tr>
</tbody>
</table>