ABSTRACT

Registration between 3-D volumes and 2-D fluoro images for Electrophysiology (EP) is a challenging task due to the lack of corresponding features between 2-D and 3-D data. This paper presents an automatic, accurate and workflow-friendly 2-D/3-D registration method specially designed for patient movement correction during EP procedures. Firstly, 2-D spines are enhanced by exploiting the temporal information in the EP fluoro sequence and by removing the moving structures (i.e. catheters) that can negatively affect the registration accuracy. Secondly, digitally reconstructed radiographs (DRRs) are computed efficiently from the segmentation of the 3-D spines using GPU-based method. Thirdly, special 2-D/3-D registration scheme is proposed, including an initial pose adjustment of the 3-D volume via automatic estimation of the spine orientation, a nonlinear histogram mapping between DRRs and X-ray images that maximizes their similarity in intensity distributions, and a gradient difference-based similarity measure that is tailored for comparing spine structures. Experiments on real clinical data demonstrate the efficacy of the proposed method.

Index Terms — X-ray angiography, C-Arm CT, 2-D/3-D registration, motion estimation, electrophysiology.

1. INTRODUCTION

Atrial fibrillation (AFIB) is a leading cause of stroke and one of the most common heart rhythm disorders. Radiofrequency catheter ablation (RFCA) has the potential of becoming the therapy of choice for treating AFIB. In order to guide the process of finding the site of origin where the excess firing of cells occur, 2-D X-ray fluoroscopy has been routinely used to provide real-time monitoring of EP procedures. However, fluoro images do not display the anatomic structures without the contrast agent, which on the other hand needs to be minimized for patients’ safety. To augment the doctor’s view of the heart anatomies and help navigate to the targeted area, high-resolution pre-operative computed tomography (CT) and/or magnetic resonance (MR) volumetric data can be fused with the intra-operative 2-D fluoroscopy using 2-D/3-D registration techniques.

Registration between 3-D volumes and 2-D fluoro images for EP is a challenging task due to the fact that there are few discernable features in a typical EP fluoro image. Rhode et. al. proposed to use multimodality fiducial skin markers to register MR volumes with X-ray images [1]. This method however requires specialized markers as well as a specialized workflow to guarantee that the markers are applied at the same position between the pre- and intra-operative data. Liao et. al. and Sra et. al. further proposed to utilize the coronary sinus (CS) and the catheter placed inside the CS as a location constraint to perform 2-D/3-D registration [2, 3]. This method is not applicable when CS catheter is not used during the procedure.

Recently cardiac C-Arm CT volume can be acquired on the same machine as fluoroscopy shortly before the intervention. Hence C-Arm CT is perfectly aligned with fluoro images if the patient does not move. However, patient movement is frequently observed during an EP procedure that typically takes a couple of hours. The goal of this paper is to design an accurate and workflow-friendly spine-based 2-D/3-D registration technique between C-Arm CT and fluoro images to compensate for patient movement in EP. Spines are a natural landmark with rich geometric variations that can be used for accurate registration. The proposed method can be applied when C-Arm CT volume is available for bringing the pre-operative CT/MR volume into alignment with the patient via 3-D/3-D registration, or when C-Arm CT is used in place of CT/MR for interventional guidance. General intensity-based 2-D/3-D registration techniques [4] will largely fail for EP applications due to the mismatch between the structures present in the 3-D and in the fluoro images. For example, the catheters are typically only shown in the fluoro images, while the chambers and/or vessels are only highlighted and visible in the 3-D volume. To our knowledge, this is the first fully-automatic intensity-based 2-D/3-D registration method that is designed and tested for EP, without the requirement for contrast agent injection.

2. 2-D & 3-D SPINE ENHANCEMENT

The spine and the EP catheters (CS, lasso, and ablation) are the most prominent structures in a typical EP fluoro image (Fig.1.c). We propose to enhance the spine from an EP fluoro sequence by taking advantage of the property that the spine is relatively static compared to the EP catheters that
are constantly moving due to cardiac and/or respiratory motion. For a fluoro sequence with N frames and a reference frame \( I_r \), the motion coefficient representing the change between \( I_2 \) and the \( k^{th} (k = 1, 2\ldots N) \) frame \( I_k \) at pixel \( p \) is calculated by Wronskian change detector [5]:

\[
mc_i(p) = \frac{1}{w} \sum_{j=1}^{w} \left( \frac{I_r(p_j)}{I_k(p_j)} \right)^2 - \frac{1}{w} \sum_{j=1}^{w} I_k(p_j)
\]

(1)

where \( w \) is the number of pixels in the neighborhood window of pixel \( p \) (e.g. \( w=9 \) for a 3x3 window), \( p_j \) is the \( j^{th} \) pixel in the neighborhood window of pixel \( p \), and \( I_r(p) \) and \( I_k(p) \) represent the image intensity at pixel \( p \) of \( I_r \) and \( I_k \) respectively. Wronskian change detector has shown to be superior to several state-of-the-art motion detectors such as shading model, derivative model, statistical change detection, and linear dependence change detector [5]. The amplitude of the motion coefficient is large for those pixels that contain appearance change, i.e. due to the moving catheters, and is small for the static background including the spine. In addition, the motion coefficient is negative (positive) for the pixels that are dark (bright) in \( I_r \) and bright (dark) in \( I_k \). Assuming \( H \) is the threshold on the motion coefficient for the change from a bright background pixel to a dark catheter pixel (set to 0.05 heuristically), a pixel \( p \) in the reference frame \( I_r \) is defined as a catheter pixel if \( mc_i(p) < -H \) for some \( k \ (1 \leq k \leq N) \). The intensity value of the background at pixel \( p \) is then calculated as:

\[
\frac{1}{N} \sum_{k=1}^{N} I_k(p)
\]

(2)

with \( mc_i(p) < -H \) for a catheter pixel

\( mc_i(p) < H \) for a non-catheter pixel

Since the choice of the reference frame can bias the background calculation in Eq. (2), multiple reference frames (set to 3 in our method) with the catheters at different locations are detected automatically by maximizing the average motion coefficients between given two frames. The final background image is obtained as the average of all the background images calculated using each of the reference frames, followed by a gaussian smoothing (Fig.1.d).

The volume of interest (VOI) that contains the spine without the contrast-filled heart is delineated via irregular volume punching. This operation can be performed before the intervention and is straightforward because the spine and the heart are well separated in 3-D space. More automatic method can be applied here [6], however, for our registration purpose an accurate segmentation of the spine from the 3-D volume is not necessary. DRRs are then generated from the VOI using the 2-D and/or 3-D texture-based volume rendering techniques on the graphics processing unit (GPU). GPU-based methods yield better computational efficiency than software-based techniques such as ray-casting. It takes about 20ms to generate a 256x256 DRR from a 256x256x253 volume with an NVidia Quadro FX Go1400.

**3. 2-D/3-D REGISTRATION OF SPINES**

The transformation relating points in the 3D volume to points on the projected X-ray image consists of six extrinsic rigid-body parameters \( T = \{ t_x, t_y, t_z, \theta_x, \theta_y, \theta_z \} \) that are estimated by the iterative registration algorithm, and four intrinsic perspective projection parameters determined by the X-ray imaging system based on a pinhole camera model (Fig.2). Since patient movements during the intervention are largely constrained by the table, the motion is presumed to be dominant in the table-plane. Hence a transformation of three parameters \( \{ t_x, t_z, \theta_z \} \) in the table plane is estimated first, followed by a fine tuning for the six parameters in \( T \).

**Fig.2. A schematic view of X-ray imaging and DRR generation.** X, Y and Z axes are from right to left, anterior to posterior, and feet to head of the patient on the table respectively.

It is well recognized that intensity-based 2-D/3-D registration has a relatively small capture range, e.g. with a convergence rate of 80% only up to 8mm displacement [7]. In order to increase the capture range, an initial orientation estimation of the spine is performed on both the fluoro and the DRR images (details are given later). Since an EP fluoro
image is commonly taken from the anterior-posterior (AP) view without table tilt (or this type of fluoro image can be taken specifically for our registration purpose when patient movements are observed), the angle between the orientations estimated from the fluoro and the DRR images can be used as the initial estimation for $\theta_j$. An initial estimation of $t_x$ and $t_z$ are further obtained by a global search in the area of [-20 20] mm around the current position with a step size of 3 mm given the above estimation on $\theta_j$.

Spine orientation estimation is performed by first applying histogram equalization on the 2-D image to enhance the spine structure. The detection of ridge points in the outer boundary of the vertebrae is then achieved by a morphological filtering called bottom-hat transform:

$$I_{\text{bottom\_hat}} = I \ast b - I$$

(3)

where $\ast$ denotes the closing operation and $b$ is the structural element of a disk of radius 5 for a 512x512 image. The non-vertebrae points are further reduced by retaining only the connected components that are larger than 30 pixels in size. Hough transform is performed to locate the strongest line in the bottom-hat transformed image. Since there are multiple strong lines corresponding to the edges of the multiple vertebrae, the mapping obtained by Hough transform is first thresholded and then summed across all the shifts. The orientation of the spine is estimated as the angle that has the highest number of votes summed across shifts. A similar method was proposed in [8] with two differences: top-hat transform was used for white structures on a dark background, and no summation across shifts was performed due to the difference in the type of spine images experimented.

Gradient-based similarity measure is observed to be suitable for comparing spines whose dominant features lie in the edge of the vertebrae. Since correction-based method is known to be sensitive to outliers, the sum of the absolute differences of the gradients (GD) is adopted in our method.

To further improve the robustness, only those pixels that have a gradient higher than a certain threshold (obtained as a given percentile) or are in the mask generated by the bottom-hat transform (described earlier) are counted in the similarity calculation. Furthermore, the absolute difference between the intensity profiles of the fluoro image $I_f$ and the DRR image $I_d$ needs to be minimized in order to reliably use the GD similarity measure. We suggest a monotonic nonlinear mapping method to align the intensity distribution of $I_f$ to that of $I_d$. Ideally the grey value $i$ in $I_f$ should be mapped to the grey value $i'$ where $C_f(i) = C_d(i')$. Here $C_f(.)$ and $C_d(.)$ denote the cumulative density function (CDF) of $I_f$ and $I_d$ respectively. For a discrete histogram distribution a perfect mapping typically can not be achieved and the following algorithm is proposed to achieve an optimal approximation:

$$i \rightarrow i' = \min_{t_i \leq i \leq t_2} \left| p_d(i) - p_f(k) \right|$$

(4)

with $t_1$ and $t_2$ defined by:

$$C_f(t_1 - 1) < C_d(t_1) \leq C_f(t_2)$$

$$C_f(t_1 - 1) < C_d(i) \leq C_f(t_2)$$

$$C_f(t_1) \leq C_d(i) \leq C_f(t_2)$$

Here $P_d(.)$ and $P_f(.)$ denote the histogram of $I_d$ and $I_f$ respectively. The proposed mapping is similar to histogram equalization (HE) with two major differences: 1) unlike HE whose targeted distribution is the uniform distribution, our targeted distribution is the histogram of the fluoro image which can be arbitrary; 2) In HE, the mapping is essentially $i \rightarrow i' = k$ where $C(k - 1) < C_f(i) \leq C(k)$ and $C(.)$ denotes the CDF of the uniform distribution. In contrast, there is a local histogram difference minimization step in our proposed mapping to optimally align the two histograms.

Fig. 3. Spine orientation estimation. (a) 2-D spine image; (b) ridge detection by bottom-hat transform; (c) Hough transform, with several local maxima corresponding to the edges of the vertebrae; (d) votes accumulated across shifts vs angle. The lower image is rotated by 20 degrees counter clock-wise from the upper image. The estimated orientations are at 9 and 29 degrees respectively.

Fig. 4. Nonlinear histogram mapping between fluoro and DRR images. (a) histogram of the fluoro image in (d); (b) histogram of the DRR image in (e); (c) the mapped histogram from (b) to (a); (f) the mapped DRR image which has better similarity to the fluoro image in terms of image intensity profile and variation.

4. RESULTS

The proposed method was experimented on three patient data acquired during EP procedures on an AXIOM Artis C-arm system (Siemens AG, Healthcare Sector, Forchheim,
Germany). The C-ARM CT volume was aligned perfectly with the fluoro images (512x512) without patient motion, and this alignment was used as the ground truth for quantifying the registration accuracy. Two hundred starting poses were generated by randomly generating a deviation from the ground truth position within the range of [-15 15] degrees of rotation and [-15 15] mm of translation in the table-plane, and the final results were averaged over all the 200 registrations. Several salient points on the spine and the pulmonary veins (PVs) were chosen to evaluate the accuracy. The performance of the proposed method was compared with that of the conventional method (Table 1). In the conventional method the original fluoro image was used as the reference, and the DRR was still generated from the VOI containing the spine. No initial estimation on spine orientation was performed, and GD over the whole image without histogram mapping was used as the similarity measure. A registration was considered to be successful if the average 3-D error on the salient points was smaller than 3 mm. Typical running time was 30s on 2.13 GHz Intel Pentium M for the proposed method.

![Registration examples.](image)

**Fig.5. Registration examples.** (a) spine-enhanced fluoro with the edge contours depicted; (b) DRRs before registration overlaid with the edge contours from the fluoro; (c) DRRs after registration overlaid with the edge contours from the fluoro; (d) 2-D/3-D overlay used for EP procedures. The 3-D mesh model is generated from the segmentation of the left atrium (LA) and PVs, and is overlaid with the fluoro by the proposed 2-D/3-D technique.

<table>
<thead>
<tr>
<th>Data</th>
<th>Success Rate (%)</th>
<th>Mean 2-D retro-projection error (pixel)</th>
<th>RMS 3-D error (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cov. Proposed</td>
<td>Cov. Proposed</td>
<td></td>
</tr>
<tr>
<td>Patient1</td>
<td>5 96</td>
<td>42.88 3.92</td>
<td>13.88 1.52</td>
</tr>
<tr>
<td>Patient2</td>
<td>3 93</td>
<td>46.14 4.47</td>
<td>14.78 1.97</td>
</tr>
<tr>
<td>Patient3</td>
<td>3 89</td>
<td>47.37 6.24</td>
<td>15.33 2.45</td>
</tr>
</tbody>
</table>

**5. DISCUSSIONS AND CONCLUSIONS**

We developed a novel 2-D/3-D registration method for patient movement correction during radio-frequency catheter ablation of atrial fibrillation. The spine is enhanced from the fluoro sequence using a computer vision-inspired motion detection algorithm, which can be reliably used as the reference image for 2-D/3-D registration. Similarly, DRRs are generated from the VOI containing the spine, resulting in a clean spine image without the requirement on accurate spine segmentation. Automatic spine orientation estimation is applied to increase the capture range of the algorithm, and a nonlinear histogram mapping is introduced for a reliable measurement of the similarity between the DRR and fluoro images using their absolute differences. Our experiments on clinical EP data show that compared to a simple application of the conventional intensity-based 2-D/3-D registration, the proposed method increased the success rate from below 5% to over 90%, and reduced the average 3-D error from ~15 mm to ~2 mm. To the best of our knowledge, this is the first spine-based 2-D/3-D registration method that is specifically designed for EP, a challenging task due to the lack of corresponding features in the 2-D and 3-D data. The proposed method is workflow-friendly (fully automatic during the intervention), and does not require fiducial markers or additional contrast agent.

From the results that is not shown here due to the space limit, we note that the registration error of the salient points on PVs was statistically larger (p<0.001) than that at the spine. Future work includes a systematic investigation of the influence of the registration error on the displacement at the LA and PVs (i.e. the target for EP), the capture range of the proposed algorithm for target offset, as well as the difference in the relative offsets of the heart and the spine between the pre-operative data and the C-Arm CT volume.

**6. ACKNOWLEDGEMENTS**

The authors are thankful to Mr. Andreas Meyer from Siemens AX Germany for providing the data in this study.

**7. REFERENCES**