Robust 2-D/3-D Registration of CT Volumes with Contrast-Enhanced X-ray Sequences in Electrophysiology Based on a Weighted Similarity Measure and Sequential Subspace Optimization

Xuan Zhao, Shun Miao, Liang Du, Rui Liao

Siemens Corporation, Corporate Technology, Princeton, NJ, USA

***Abstract*—2-D X-ray image navigation during an electrophysiology ablation procedure can be enhanced via the overlay of images derived from pre-operative 3-D data, to provide anatomical details that otherwise are not visible under X-ray. However, accurate registration of 3-D data and 2-D X-ray during electrophysiology ablation is a challenging problem, due to the relatively low image quality and the fact that the contrast medium typically fills only a small part of the left atrium in X-ray images. In this paper, we propose a robust 2-D/3-D registration method tailored for electrophysiology ablation procedures. In particular, a weighted similarity measure is utilized to handle the partial data problem, and a sequential subspace optimization is proposed to take advantage of the available bi-plane X-ray images, for a robust and efficient registration. Since the contrast medium washes out quickly and thus visible in only a small number of frames in the X-ray sequence, we furthermore streamline the workflow by automatically detecting the frame that is optimal for registration purpose. Experimental results on seven clinical data sets demonstrate the effectiveness of the proposed method.**

Keywords*—2-D/3-D Registration, Atrial fibrillation, Weighted Similarity Measure, Sequential Subspace Optimization, Optimal Frame Detection

1. INTRODUCTION

Atrial fibrillation (AF) has become a common disease that affects about two million people in the USA [1]. The fusion of pre-operative high-resolution data such as 3-D atrial CT and/or MR volumes and intra-operative fluoroscopic image sequences is introduced during radio-frequency catheter ablation (RFCA) therapy, an accepted option for treating AF [2,3]. Rhode et. al. proposed to use multimodality fiducial skin markers to register MR volumes with X-ray images [4]. However, this method requires a specialized workflow and it is difficult to guarantee that the specialized markers can be placed at the same position in the pre- and intra-operative data. Sra et. al. proposed to align the coronary sinus (CS) catheter with the CS and superior vena cava (SVC) in the CT images via user manual interaction [5]. Knecht et. al. compared three registration schemes [6], (i) by catheter, (ii) by manual alignment of the segmented heart in CT volume with the cardiac contour on the fluoroscopic image, and (iii) by spine from the segmented CT image. However, only 2-D in-plane registration is considered in their methods. Liao proposed a patient movement correction method where 2-D spines in the fluoroscopic image are used as the discernable features to be registered with the 3-D volume using a special 2-D/3-D registration scheme [7]. However, the motion estimated at the spines does not guarantee an accurate motion estimation of the left atrium, the target organ for RFCA, because the heart may have significant relative motion with regard to the spines between the pre- and intra-operative data.

In this paper, we propose a fully automatic bi-plane 2-D/3-D registration technique that directly registers the patient's left atrium in CT image to that in contrast- enhanced fluoroscopic images. The main contributions and novelties of the presented paper include:

1. A weighted similarity measure is proposed to make the algorithm robust to partial data. Since the left atrium is a relatively large chamber, it typically can be only partially filled by the limited dose of contrast medium. As a result the left atrium in fluoroscopic images is very likely to have only a partial shape. This problem is further complicated by the relatively low signal to noise ratio in the low dose EP fluroscopic images. Thus a properly defined similarity measure is crucial in order to handle the problem of partial and low quality data,
2. A sequential subspace optimization algorithm is proposed to take advantage of the bi-plane X-ray system. The proposed algorithm iteratively optimizes the registration parameters in the subspaces defined by the two image planes, to increase the capture range and the speed.
3. The workflow is furthermore streamlined by automatically detecting the optimal frame in the fluroroscopic sequence for registration.

2. METHOD

The proposed algorithm follows the general intensity-based 2-D/3-D registration framework, where Digitally Reconstructed Radiographs (DRRs) are simulated and the pose of the 3-D image is adjusted to minimize the difference between the generated DRR and the real 2-D X-ray image. Therefore our method consists of three major components: image processing, similarity measure and optimization, which will be discussed in section 2.1, 2.2 and 2.3 separately. In addition, the 2-D fluoroscopic sequence shows the perfusion of the injected contrast medium in the left atrium, and as a result the optimal frame needs to be selected from the sequence for registration purpose. In section 2.4, we propose an automatic contrast frame detection algorithm to streamline the registration workflow.

2.1. 2-D and 3-D Image Processing

Digitally Subtracted Angiograph (DSA) computed from the contrast-enhanced frame and a reference frame is widely used in contrast-enhanced fluoroscopy. The purpose of DSA is to highlight the contrast-filled organ and eliminate irrelevant steady structures, like bones. Denote the contrast-enhanced frame as I_f and the reference frame as I_r, the DSA is calculated as:

$$DSA = I_f - I_r \quad (1)$$

DRRs rendered for the entire volume tend to contain irrelevant structures that may lead to a lot of local maximums, and as a result make the registration prone to local optimum. To eliminate irrelevant structures, a segmented left atrium is used for DRR rendering. DRRs are generated using the 3-D texture-based volume rendering technique on the graphics processing unit (GPU), which yields better computational efficiency than software-based technique such as ray-casting. It takes about 15ms to generate a 256 × 256 DRR from a 256 × 256 × 256 volume using an NVidia Quadro FX 360M. Examples of the fluoroscopic, DSA and DRR images are shown in Fig. 1. Note that, DSA image shows only a small part of contrast-enhanced region while DRR image shows a complete left atrium with a clear boundary.

2.2. Similarity Measure

The similarity between DSA and DRR is computed and maximized iteratively during the registration process. Conventional image similarity measures like Mutual Information (MI) and Normalized Cross Correlation (NCC) give all the pixels the same weight. However, because left atrium is a relatively large chamber, and is usually not filled completely with the contrast medium, the intensity of the left atrium in the contrast-enhanced fluoroscopy is inhomogeneous and the shape of the left atrium might be incomplete. In addition, to minimize the dose of contrast medium for patients' safety, the left atrium shown in the DSA is typically relatively faint (Fig. 1 (b)). To make the registration robust to a partial and faint shape, we propose a spatially weighted similarity measure. It is confirmed with collaborating physicians that the contrast delivery catheter is typically put close to the roof of the left atrium when contrast medium is injected. This means the roof of the left atrium usually has a noticeably higher contrast than other regions, especially during the initial stage of the contrast agent injection. Moreover, the inner region of the left atrium has less discernible patterns than the boundary. Therefore, we use a spatially weighted Gradient Correlation (GC) to give the roof of the left atrium more weight. Given the DSA and the DRR, the weighted GC is calculated as:

$$SM(DSA, DRR) = \sum_i \alpha_i \, GC(DSA \cdot M_i, DRR \cdot M_i) \quad (2)$$

Where

$$M_i = \begin{cases} 1 & x \in \Omega_i \\ 0 & otherwise \end{cases} \quad (3)$$

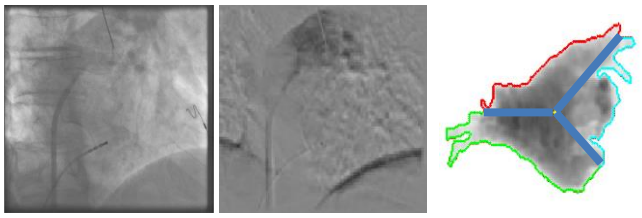
is the mask and $\bigcup_i \Omega_i = \Omega$. α_i is the weight for the *i*-th region. In addition:

$$GC(I,J) = \frac{1}{2}\left(NCC(\nabla_x I, \nabla_x J) + NCC(\nabla_y I, \nabla_y J)\right) \quad (4)$$

where $NCC(\cdot)$ is the Normalized Cross Correlation:

$$NCC(I_1, I_2) = \frac{\sum_{x,y}[I_1(x,y) - \bar{I}_1(x,y)][I_2(x,y) - \bar{I}_2(x,y)]}{\sqrt{\sum_{x,y}[I_1(x,y) - \bar{I}_1(x,y)]^2 [I_2(x,y) - \bar{I}_2(x,y)]^2}} \quad (5)$$

Other similarity measures such as MI or NCC were also tested, and as expected, gave inferior performance compared to GC, because the boundary area provides the most reliable information to quantify the similarity for this type of low quality data. To properly define the mask *M,* we take advantage of the prior knowledge about the fluoroscopy imaging geometry typically used during EP, which puts the roof of the left atrium to the top left in the 2-D fluoroscopic image. Therefore, the region can be defined based on the angular coordinate of the pixel in a polar coordinate system centered at the center of mass of the left atrium in the DRR. Three angular thresholds are empirically chosen as $\{\pi/4, \ \pi, \ 7\pi/4\}$. An example of the partitions is shown in Fig. 1 (c). We assign weight $\alpha_1 = 0.6$ to the red region which is the roof of the left atrium, and $\alpha_2 = 0.4$ to the green region which is the bottom of the left atrium in our experiments, respectively. The weight α_3 for the cyan region is set as zero, as there is no prominent boundary within this region for most data.



(a) (b) (c)

Fig. 1. An example of the (a) fluoroscopic image, (b) DSA and (c) DRR overlaid with region partitions.

2.3. Optimization

In the proposed method, the registration parameters are estimated in two steps: 1. three translation parameters are first estimated by a sequential subspace optimization method. 2. Starting from the position estimated in step 1, six rigid-body parameters are fine tuned by a Best Neighbor optimization method.

We focus on the three translation parameters in the first step because after the calibration-based alignment using the machine parameters there is only limited rotation mismatch between the acquisition of the cone-beam CT and the fluoroscopy for most data. Altough the dimensionality of the parameter space is reduced to three in the first step, a brute force global search of three dimensions simultaneiously is still too computationally expensive for clinical use. We propose the following subspace optimization algorithm to reduce the computational complexity, compared to the brute force global search. On the other hand, compared to local op-

timizer-based optimization, our proposed sequential subspace optimization has the advantage of an extremely large capture range and robust convergence.

We denote the camera reference coordinate systems of the two image planes as $[\boldsymbol{i_1}, \boldsymbol{j_1}, \boldsymbol{k_1}]$ and $[\boldsymbol{i_2}, \boldsymbol{j_2}, \boldsymbol{k_2}]$, where $\boldsymbol{k_i}$ is the direction along the camera axis. For simplicity of explanation, we use the example case that the two image planes are perpendicular to each other. Therefore we can choose the two coordinate systems so that

$$\begin{aligned} \boldsymbol{i_1} &= \boldsymbol{i_2}, \\ \boldsymbol{j_1} &= \boldsymbol{k_2}, \\ \boldsymbol{k_1} &= -\boldsymbol{j_2}. \end{aligned} \tag{6}$$

We alternatively search in the subspace spanned by $\boldsymbol{i_1}, \boldsymbol{j_1}$ and the subspace spanned by $\boldsymbol{i_1}, \boldsymbol{k_1}$ to maximizes the similarity measure in Equation (2). Denote the 3-D translation as $\boldsymbol{T} = x \cdot \boldsymbol{i_1} + y \cdot \boldsymbol{j_1} + z \cdot \boldsymbol{k_1}$, the search in the first image plane would be:

$$\{x_{k+1}, y_{k+1}\} = \max_{\substack{x \in [x_k - r_k, x_k + r_k] \\ y \in [y_k - r_k, y_k + r_k]}} SM\big(DSA, DRR_1(\boldsymbol{T_k})\big) \tag{7a}$$

Similarly, the search in the second image plane is:

$$\{x_{k+1}, z_{k+1}\} = \max_{\substack{x \in [x_k - r_k, x_k + r_k] \\ z \in [z_k - r_k, z_k + r_k]}} SM\big(DSA, DRR_2(\boldsymbol{T_k})\big) \tag{7b}$$

We call the above two consecutive searches an iteration. After each iteration, the search range r_k is halved. Usually after two to three iterations the registration is fairly accurate. From this point, a six-dimensional rigid-body registration fine tuning can be launched to further improve the accuracy.

This algorithm takes advantage of the fact that 2-D projection image is not sensitive to the object's position along the camera axis direction. Even if the error in z is relatively large, the search of x and y using Equation (7a) is still fairly accurate. Based on this property, we can decouple the parameter space and focus on only two dimensions in each search to dramatically reduce the computational complexity. In particular, the complexity is reduced from $O(N^3)$ to $K \cdot O(N^2)$, where N is the searching range and K is the number of iterations, which is usually 2 or 3 and thus much smaller than N (typically on the order of hundreds).

Note that the proposed sequential subspace optimization is not limited to the setup of two orthogonal image planes, but can be applied to general bi-plane image systems with an arbitrary angle between the two image planes. The advantage of using an orthogonal setup, as typically done in an EP procedure, is to achieve the optimal accuracy in 3-D. This is because for a given image plane, in-plane (depth) estimation is the most (least) accurate direction. In an orthogonal setup, the 3-D space is optimally covered because the depth direction in one image plane is the in-plane direction for the other image plane. We observed that in practice, our sequential subspace optimization can be applied robustly to achieve an accurate registration in 3-D when the two image planes are at least 30 degrees apart.

2.4. Optimal frame detection

In previous sections, we have been assuming that the contrast-enhanced frame is already selected. However, manual selection of the contrast-enhanced frame is not desirable in clinical practice because it interferes with the workflow. In addition, a manually-selected frame via visual inspection may not be the optimal one to be used as the input for a given registration algorithm. Therefore, we propose an optimal frame detection method to streamline the workflow. The basic idea is to perform 2-D/3-D registration on all frames with contrast, and select the one with the highest similarity measure. To detect the first frame after contrast injection, we compute the difference between neighboring fluoroscopic frames:

$$DDSA_j = DSA_j - DSA_{j-1}. \tag{8}$$

Because the intensity of contrast region decreases very fast after contrast injection, we apply a heuristically chosen threshold to $DDSA_j$ to detect the region with contrast medium presence. Denote n_j as the number of pixels whose $DDSA_j$ is lower than the threshold. The first frame with $n_j > N_{pixel}$, where N_{pixel} is empirically set as 1000 in a 256 x 256 image, is selected as the first contrast-enhanced frame. Starting from this frame, registration is performed on each contrast-enhanced frame, using the registration result of the previous frame as the starting position. As the registration results of two neighboring frames should be close, small search range can be used for the following frame for computational efficiency. The frame with the highest similarity measure is then selected and the corresponding registration result is used as the final registration result.

3. EXPERIMENTAL RESULTS

We evaluated the proposed algorithm on 7 clinical datasets. The number of frames of the datasets varies within 16 ~ 36. We compared the proposed method with conventional 2-D/3-D registration methods using a Best Neighbor optimizer with Mutual Information (MI) and Gradient Correlation (GC) as the similarity measure. To evaluate the registration results, we computed the 3-D Target Registration Error (TRE) with respect to the ground truth position, which was chosen manually based on the discernible boundaries in the DSA and DRR bi-plane images in the selected optimal frame and validated by an expert. The results are listed in Table 1. It shows that the conventional method using MI achieves much higher registration accuracy than using GC. This is mainly because GC has very sharp similarity profile and very small capture range. The proposed method achieves much smaller TRE compared with both conventional methods on most cases except case #6, where the proposed method has slightly larger error than the MI based conventional method. By looking into case #6 (Fig. 2(c)), the registered DRR is slightly higher than the ground truth position (see the roof part in the fluoroscopic image). This is likely to be caused by the catheter near the bottom boundary of the atrium, which has large gradient magnitude in the DSA image and contributes more significantly to the GC score than the MI score. Overlays of the 3-D model and the 2-D X-ray image of case #1, #3 are also

shown in Fig. 2 for qualitative evaluation. Case #1 has a significantly larger TRE (7.1mm) compared to other 6 datasets, mainly due to the noticeably larger difference in the shape of the left atrium between the 3-D and 2-D images (for unclear reason). However, Fig. 2(a) shows that with a TRE of 7.1mm the overlay of case #1 is still reasonable and clinically useful.

To demonstrate the necessity of the optimal frame detection, Fig. 3 shows the registration result using two different frames in the same sequence. Both frames contain contrast medium. The strong roof boundary and a faint but noticeable bottom boundary help the registration in Fig. 3(a), while the diffused contrast in Fig. 3(b) makes the registration less accurate.

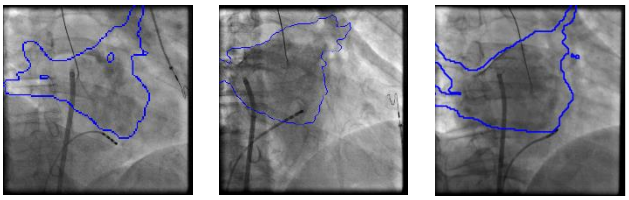
Table 1. TRE (mm) of proposed method and conventional methods

Data #	Conv. (MI)	Conv. (GC)	Proposed
1	13.5	48.2	7.1
2	2.9	43.9	2.7
3	11.0	50.8	1.4
4	6.9	66.1	3.9
5	6.4	9.2	1.0
6	3.4	48.2	4.3
7	9.8	34.8	2.6

The translation-only registration result is fine-tuned by a rigid-body registration using a Best Neighbor optimizer with Gradient correlation (GC) as the similarity measure. Table 2 shows the adjusted parameter in addition to the translation-only results. It shows that the additional rotation after machine calibration is relatively small for most cases. More noticeable rotations are observed for case #1 and #4. Unfortunately, it is hard to even determine the ground truth rotations due to the missing structures in the contrast-enhanced fluoroscopic images. Note that in the current clinical practice, the physicians only translate the 3-D volume manually to align it with the fluoroscopy, due to the difficulty in mentally predicting the right direction for out-of-plane rotation. Nevertheless, we observed that the rigid-body registration result is more reasonable than the translation-only registration result for these cases based on visual inspection. One of the examples (case #4) is shown in Fig. 4, where the rigid-body fine-tuned registration (Fig. 4. (c) and (d)) fits the roof of the left atrium better than the translation-only result (Fig. 4. (a) and (b)).

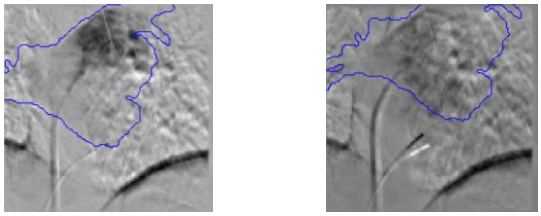
Table 2. Adjusted parameters by rigid-body fine-tuning registration. Translations are in *mm*, while rotations are in *degree*.

Data #	Tx	Ty	Tz	Rx	Ry	Rz
1	0	-0.45	0.1	-10.5	-5.5	0.5
2	0.1	0	-0.1	7.8	-1.0	0.9
3	0.4	-0.1	-1.1	0.1	-2.0	0
4	-0.6	-0.3	-0.9	8.7	2.3	4.5
5	-0.4	-0.1	-0.4	0.6	1	0.3
6	-1	-1.1	-3.1	1.7	-3.2	0.8
7	-0.6	-0.6	4	0.6	8.7	-0.1



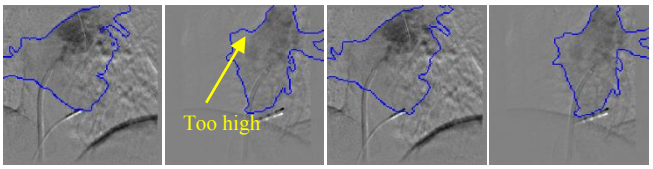
(a) (b) (c)

Fig. 2. Registration results using the proposed method (blue contour represents the boundary of the DRR image). (a) case #1, (b) case #3, (c) case #6.



(a) (b)

Fig. 3. Registration results of two different frames of the same data. (a) 6th frame after contrast injection (b) 9th frame after contrast injection. Blue contour represents the boundary of the DRR image.



(a) (b) (c) (d)

Fig. 4. Registration result w/ and w/o rigid fine-tuning. (a) and (b) are translation-only registration result in two image planes. (c) and (d) are rigid-body fine-tuned registration result.

4. CONCLUSION

In this paper, we proposed a fully automatic 2-D/3-D registration scheme for fusing the 2-D fluoroscopic image with pre-operative 3-D volume in EP procedures. As a partial shape of the left atrium is common in X-ray images due to the limited dose of contrast agent, we utilize a weighted similarity measure to deal with the partial shape problem. By taking advantage of the two imaging planes, we define two 2-D subspaces in the 3-D parameter space and optimize the parameters in these two subspaces alternatively. To make the whole workflow fully automatic, we furthermore streamline the workflow by automatically detecting the frame that is optimal for registration purpose. The proposed algorithm is evaluated on 7 clinical datasets, demonstrating its efficiency and robustness.

ACKNOWLEDGEMENT

Authors would like to thank Dr. Atilla Kiraly and Dr. Norbert Strobel for providing the clinical data, and Dr. Christophe Chefd'hotel for his support to this work.

REFERENCES

[1] S. Stewart, C. L. Hart, D. J. Hole, J. McMurray, *et al.*, "A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the renfrew/paisley study.," *The American journal of medicine*, vol. 113, no. 5, p. 359, 2002.

[2] R. Cappato, H. Calkins, S.-A. Chen, W. Davies, Y. Iesaka, J. Kalman, Y.-H. Kim, G. Klein, D. Packer, and A. Skanes, "Worldwide survey on the methods, efficacy, and safety of catheter ablation for human atrial fibrillation," *Circulation*, vol. 111, no. 9, pp. 1100–1105, 2005.

[3] O. M. Wazni, N. F. Marrouche, D. O. Martin, A. Verma, M. Bhargava, W. Saliba, D. Bash, R. Schweikert, J. Brachmann, J. Gunther, *et al.*, "Radiofrequency ablation vs antiarrhythmic drugs as first-line treatment of symptomatic atrial fibrillation," *JAMA: the journal of the American Medical Association*, vol. 293, no. 21, pp. 2634–2640, 2005.

[4] K. Rhode, Y. Ma, A. Chandrasena, A. King, G. Gao, P. Chinchapatnam, M. Sermesant, D. Hawkes, T. Schaeffter, J. Gill, *et al.*, "Evaluation of the use of multimodality skin markers for the registration of pre-procedure cardiac mr images and intra-procedure x-ray fluoroscopy images for image guided cardiac electrophysiology procedures," in *Proc. SPIE*, vol. 6918, 2008.

[5] J. Sra, G. Narayan, D. Krum, A. Malloy, R. Cooley, A. Bhatia, A. Dhala, Z. Blanck, V. Nangia, and M. Akhtar, "Computed tomography-fluoroscopy image integration-guided catheter ablation of atrial fibrillation," *Journal of cardiovascular electrophysiology*, vol. 18, no. 4, pp. 409–414, 2007.

[6] S. Knecht, H. Skali, M. D. O'Neill, M. Wright, S. Matsuo, G. M. Chaudhry, C. I. Haffajee, I. Nault, G. H. Gijsbers, F. Sacher, *et al.*, "Computed tomography–fluoroscopy overlay evaluation during catheter ablation of left atrial arrhythmia," *Europace*, vol. 10, no. 8, pp. 931–938, 2008.

[7] R. Liao, "2-d/3-d registration of c-arm ct volumes with fluoroscopic images by spines for patient movement correction during electrophysiology," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pp. 1213–1216, IEEE, 2010.