3D-2D Registration Based on Mesh-Derived Image Bisection

David Thivierge-Gaulin[1], Chen-Rui Chou[2], Atilla P. Kiraly[3], Christophe Chefd'Hotel[3], Norbert Strobel[4], and Farida Cheriet[1]

[1] École Polytechnique de Montréal, Montreal, QC, Canada
david.thivierge-gaulin@polymtl.ca
[2] University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[3] Siemens Corporation, Corporate Research and Technology, Princeton, NJ, USA
[4] Siemens AG, Healthcare Sector, Forchheim, Germany

Abstract. Electrophysiology procedures such as catheter ablation for atrial fibrillation are non-invasive approaches for treating heart arrhythmia. These operations necessitate contrast liquid injections for the left atrium and pulmonary veins to be visible under fluoroscopy. However, injections have to be minimized because of their toxicity. To provide visual guidance after the contrast liquid has washed away, it is possible to overlay a mesh of the left atrium obtained from a pre-operative 3D volume over the intra-operative 2D fluoroscopic images. This paper presents a novel mesh-based registration algorithm providing such an overlay by registering the left atrium mesh to fluoroscopic images showing contrast liquid injection. The registration is based on image bisections generated by mesh projections, which bypasses the original volumetric data and digitally reconstructed radiographs generation. The algorithm was validated on 7 clinical datasets and registers with a mean target registration error of 6.56 ± 2.67mm.

Keywords: Mesh Registration, 3D/2D Registration, 2D/3D Registration, Model-Based Registration, Hybrid Registration, Image-Guided, Atrial Fibrillation, Catheter Ablation, Electrophysiology.

1 Introduction

Image guidance during electrophysiology (EP) procedures such as catheter ablation (CA) for atrial fibrillation (AF) has been shown to decrease procedure duration and likelihood of AF recurrence [1]. Since the left atrium (LA) and pulmonary veins (PV)s are not visible under fluoroscopy without the injection of contrast liquid [1], the operation is facilitated by overlaying a 3D mesh extracted from pre-operative 3D volume (CT or MRI) over the intra-operative fluoroscopic images. To provide a correctly aligned overlay, the 3D mesh of the LA with attached PVs is *registered* at the time on contrast liquid injection to serve as a visual reference after the contrast liquid has washed away.

It is possible to register the mesh manually, but a quick and automatic registration algorithm is desirable because it allows reducing the number of operative workflow steps, higher reproducibility of results and does not require

B.M. Dawant et al. (Eds.): WBIR 2012, LNCS 7359, pp. 70–78, 2012.

a trained professional to be present. Typical automatic 3D-2D registration algorithms transform a 3D pre-operative volumetric image into a 2D digitally reconstructed radiograph (DRR), which is in turn compared to intra-operative fluoroscopic 2D images [2]. The assumption is that since the DRRs simulate fluoroscopic images, they will be resemblant enough so that a well chosen similarity measure between the two will have its global minimum coincide with the registered position of the two (3D-2D) modalities.

An alternative to DRR-based methods is to directly register the mesh to the fluoroscopic images, thus allowing exploitation of the information contained in the mesh's manual segmentation from pre-operative 3D volume. This paper describes such an algorithm that relies on analysis of the image bisection generated by the projection of the mesh. It allows bypassing the production of DRRs as well as not requiring the use of volumetric data. This is useful in EP procedures that do not use the volumetric data as part of their workflow as well as having the potential to be faster than DRR-based registration. The DRR production is usually the main bottleneck to achieve fast registration because it has to be iteratively evaluated in the optimizer inner loop – replacing the DRR by a faster process would greatly enhance registration speed.

The algorithm is described in section 2 and validated on 7 clinical cases (section 3). The potential use, advantages and drawbacks of the solution are outlined in section 4.

2 Methods

This section describes the steps of the mesh-derived registration algorithm:

1. Pre-process the 2D and 3D data (section 2.1).
2. Bisect the fluoroscopic images using mesh-to-mask projection (section 2.2).
3. Compute a cost from the image bisection (section 2.3).
4. Find the registered position using an optimizer (section 2.4).

The four steps are illustrated in figure 1.

2.1 Data Pre-processing

Generating the 2D Subtracted Images. Our source data consists of a biplane DICOM sequence of between 15 to 40 fluoroscopic frames of 1024x1024 pixels (2D) showing the injection of contrast liquid in the LA. As can be seen in figure 2 (a), the region of interest (LA and PVs) is not visible under fluoroscopy unless injected with contrast liquid [1]. Contrast liquid cannot be constantly injected during the operation because it is harmful to the patient. It is therefore crucial to process the images taken during the injection in order to get the best approximation of the 2D LA topology. In order to obtain a good delineation of the LA from the background and reduce interference from other image components, a frame that contains contrast liquid is subtracted to a frame that does not. No motion compensation is applied to account for movement between the two time points. The images are downsampled to a resolution of 256x256 pixels in order to speed-up the registration process.

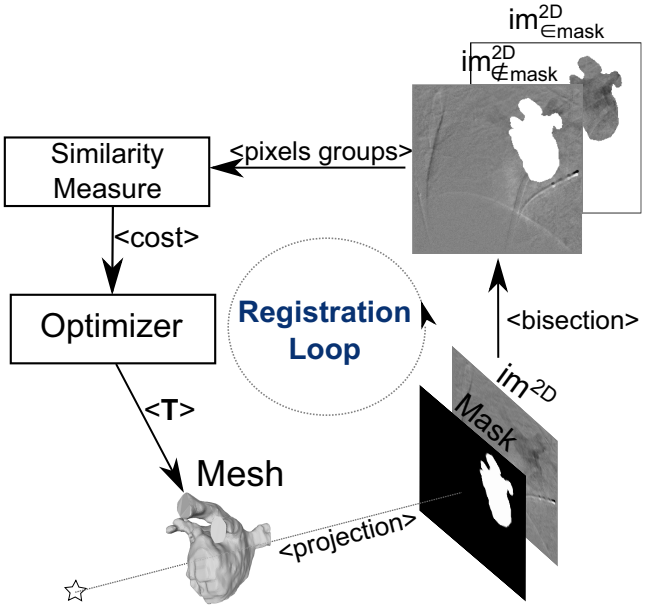

Fig. 1. Overview of the registration algorithm. The mesh is first projected into a mask, which creates a bisection of the 2D image. The cost of the bisection is evaluated by the similarity measure which is fed to the optimizer. The optimizer then iteratively modifies the parameters of the rigid transformation **T** to find the minimum cost.

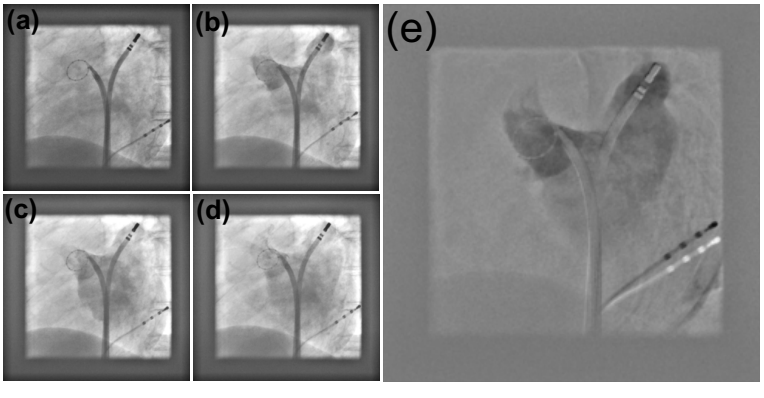

Fig. 2. (a) to (d): Sequence of fluoroscopic images showing the injection of contrast liquid in the LA (frames 0, 10, 19 and 35). (e): Subtracted image (frame 10 - frame 0). Note that the surgical instruments used for EP procedures are present in the images.

Segmentation of the Volumetric Data. In the clinical cases used for this paper, the MRI data was manually segmented into a mesh by a health-care professional.

Mesh Pre-processing. 3D meshes of the LA with attached PVs were used in our experiments. Since the extremities of the small PVs are not visible even during the injection of the contrast liquid, they are manually cut off the mesh before the operation in order to have a better match between the 2D fluoroscopic images and the projected mesh (figure 3).

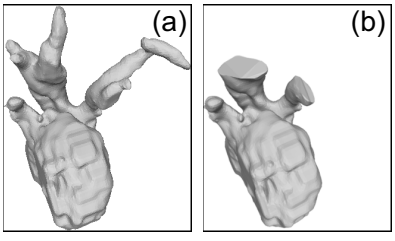

Fig. 3. The LA with PVs mesh, without pre-processing (a) and with shortening of the PVs (b)

2.2 Bisection Using Mesh to Mask Projection

The mesh-derived image bisection method directly uses a mesh extracted from volumetric data to create different *groups* of pixels once projected over a fluoroscopic image. The two groups formed are the pixels that fall under the projection of the mesh ($\in mask$) and the ones that do not ($\notin mask$) (see figure 4). The main insight is that when the mesh is properly *registered*, the grouped pixels will share *common characteristics* because they belong to the same entity (e.g. an organ or a zone that contains contrast liquid).

A projection system is setup in order to transform the mesh into a *mask* that aggregates the pixels in two groups.

$$mask := \text{MaskProjection}\big(\mathbf{T},\ \mathbf{P},\ mesh\big) \tag{1}$$

where $\mathbf{T} = \{T_x, T_y, T_z, \theta_x, \theta_y, \theta_z\}$ are the extrinsic rigid-body transformation parameter and $\mathbf{P} = \{f, o_x, o_y, s_x, s_y, \theta_y^{biplane}\}$ the intrinsic perspective projection parameters. $\theta_y^{biplane}$ is a rotation parameter centered on the middle of the mesh used to create a second view in cases of biplane registration. The intrinsic parameters $\mathbf{P}$ are determined from the fluoroscopic imaging system and the extrinsic parameters $\mathbf{T}$ are estimated by the registration algorithm. The projection system and the parameters are illustrated in figure 5.

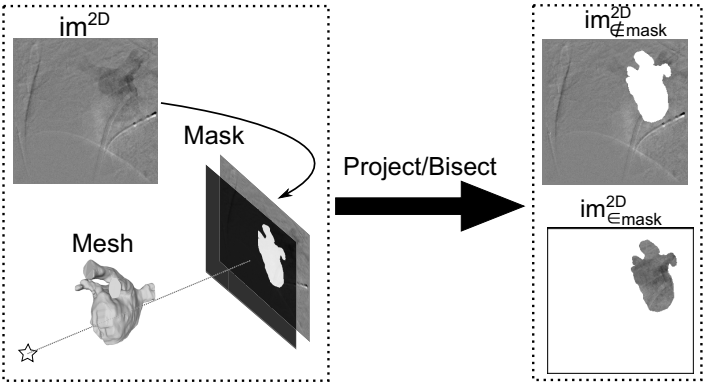

Fig. 4. The projections of the mesh creates a bisection of the image (im^{2D}) into two pixel groups: $im^{2D}_{\in mask}$ and $im^{2D}_{\notin mask}$

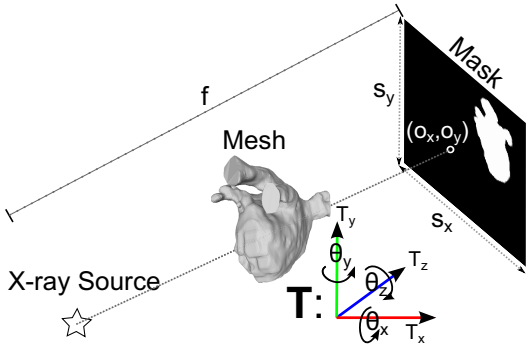

Fig. 5. Projection system used to create the mask from the mesh with extrinsic rigid-body parameters $\mathbf{T} = \{T_x, T_y, T_z, \theta_x, \theta_y, \theta_z\}$ and intrinsic perspective projection parameters $\mathbf{P} = \{f, o_x, o_y, s_x, s_y, \theta_y^{biplane}\}$

2.3 Similarity Measure Driven by Image Bisection

In order to evaluate if the groups of pixels formed by the current mesh pose correspond to a registered mesh, it is necessary to derive a similarity measure that is minimum when the mask is overlaid over the 2D image's target structure and high when over other image regions. The idea to register using pixel groups is inspired by snake methods, where a segmentation is found by iteratively evolving a curve via the minimization of an energy function. The difference in our approach is that the rigid-body parameters $\mathbf{T}$ are iteratively modified instead of the curve's control points, thus *indirectly* changing the contour of the segmentation curve according to the mesh's topology. Another way to see our solution is that it constrains the possible curves to the subset of curves that can be obtained by projecting the mesh.

If one assumes that the target 2D region is relatively homogeneous and markedly different from the other zones of the 2D image, a simple comparison of the average pixel values that fall inside and outside of the mask with the pixels in and out of these groups can be a good indication of the fitness of the position. This is inspired by the cost function of a level-set segmentation approach introduced in [3], which leads to the definition of the following similarity measure:

$$\begin{aligned}\mathrm{CostFn}(im^{2D},\ mask) := &\sum_{\forall(x,y)\in mask}\left(im^{2D}_{(x,y)} - avg(im^{2D}_{\in mask})\right)^2 \\ &+ \sum_{\forall(x,y)\notin mask}\left(im^{2D}_{(x,y)} - avg(im^{2D}_{\notin mask})\right)^2 \end{aligned} \tag{2}$$

where $im^{2D}_{(x,y)}$ is the intensity value of the fluoroscopic image at position (x, y) and $avg(im^{2D}_{\in mask})$, $avg(im^{2D}_{\notin mask})$ are the average intensity values for the *group* of pixels inside and outside the mask respectively.

2.4 Finding the Registered Position Using an Optimizer

The complete registration algorithm, illustrated in figure 1, solves the following equation:

$$\hat{T_n} = \arg\min_{T_n}\ \mathrm{CostFn}\left(im^{2D},\ mask_{T_n}\right) \tag{3}$$

where $mask_{T_n}$ is a mask created by the projection of the mesh under transformation T_n (equation 1). The 'arg min' is approximated by a *chain* of two Powell optimizers. The first operates over translation only, followed by an optimization over translation and rotation. The solution of the registration is the rigid transform $\hat{T_n}$ applied to the atrial mesh, generating the grouping of pixels on the 2D image that minimizes equation 2.

3 Results

3.1 Experiment Description

Our dataset contains 7 cases (labeled as 'C#', e.g. C200) of CA for AF, each of which has an atrial mesh that was manually segmented from MRI data along with intra-operative biplane fluoroscopic sequences showing the injection of contrast liquid. The biplane intrinsic perspective projection parameters and the ground truth extrinsic rigid-body transformations that register the meshes to the biplane images are found by careful interactive visual examination of the mesh

and subtracted fluoroscopic images. The cost function (equation 2) is adapted for biplane cases by summing the cost for each plane. In order to evaluate the accuracy of the registration algorithm, a deviation of the rigid transformation $\mathbf{T}$ is applied to the ground truth before registration. The deviation is in millimeters/degrees and contained in the interval: $\Delta\mathbf{T}_{deviation} = \{\Delta T_x, \Delta T_y, \Delta T_z, \Delta\theta_x, \Delta\theta_y, \Delta\theta_z\} = \{-15..15, -15..15, -15..15, -10..10, -10..10, -10..10\}$ where 'A..B' signifies a random number between A and B following a uniform distribution.

Both the mean target registration error (mTRE) and mean projection distance (mPD) [4] are used to assess the accuracy of the algorithm. The mTRE is the mean distance between the registered and ground truth points in 3D space and mPD is similar but *after 3D-2D projection*:

$$\text{mTRE}(P, T_{regist}, T_{truth}) = \frac{1}{k}\sum_{i=1}^{k} \|T_{regist}\mathbf{p}_i - T_{truth}\mathbf{p}_i\| \tag{4}$$

$$\text{mPD}(P, M_{regist}, M_{truth}) = \frac{1}{k}\sum_{i=1}^{k} \|M_{regist}\mathbf{p}_i - M_{truth}\mathbf{p}_i\| \tag{5}$$

where $P = \{\mathbf{p}_1, \ \ldots, \ \mathbf{p}_k\}$ are the mesh's vertices (typically $k \approx 15,000$). T_{regist} and T_{truth} are the rigid body transformation found by the registration algorithm and the ground truth; M_{regist} and M_{truth} the *perspective projection* matrixes (the mPD is understood to be calculated after division by the homogeneous coordinate).

3.2 Experiment Results

Using the experiment parameters described in section 3.1, 100 starting positions randomly deviated according to $\Delta\mathbf{T}_{deviation}$ are generated for every case (total 700 starting positions). After registration, the mTRE and mPD error (equations 4 and 5) are measured in millimeters (mm). Table 1 contains the results of the experiment.

Profiling of the mask generation process (implemented in OpenGL) reveals that it takes 0.4 millisecond on a NVIDIA Quadro 2000M to generate a 256x256 mask. This compares favorably to DRR generation implemented on GPU which takes 15 milliseconds to produce a 256x256 image [5].

Figures 6 and 7 show graphical examples of registration results with typical registration errors. Note that the ground truth is not unambiguously *visually* better than the registered result. This is due to the fact that it is very difficult to discern the LA and PVs' frontier under fluoroscopy, even when contrast liquid is injected.

Table 1. mTRE and mPD error after registration initialized with starting positions derived from the ground truth. The variability measure ($\pm\sigma$) is one sample standard deviation.

Case	mTRE (mm$\pm\sigma$)	mPD (mm$\pm\sigma$)
C037	6.76 ± 2.71	6.04 ± 2.11
C129	5.97 ± 2.18	5.83 ± 2.43
C130	6.37 ± 1.96	7.15 ± 2.07
C135	5.59 ± 2.27	4.87 ± 1.84
C137	5.91 ± 1.15	5.49 ± 1.14
C154	9.55 ± 3.42	8.58 ± 3.53
C200	5.76 ± 2.20	6.12 ± 2.23
Average	6.56 ± 2.67	6.30 ± 2.55

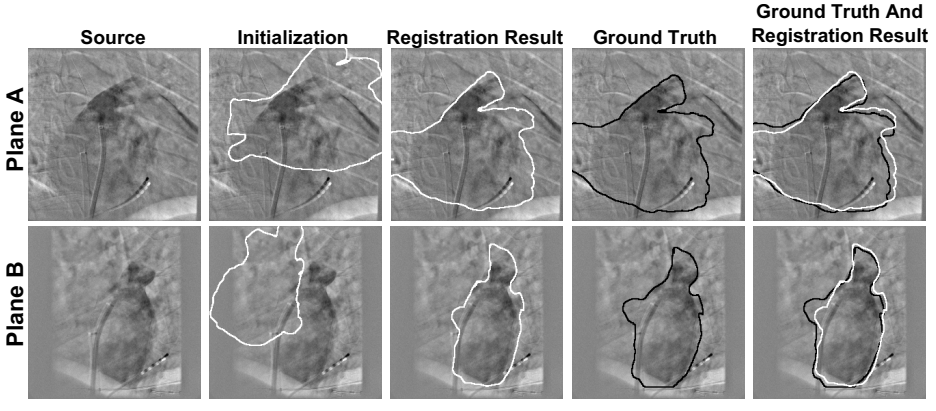

Fig. 6. Case C200 registration result compared with ground truth. The projection distance error for this registration is 7.23mm.

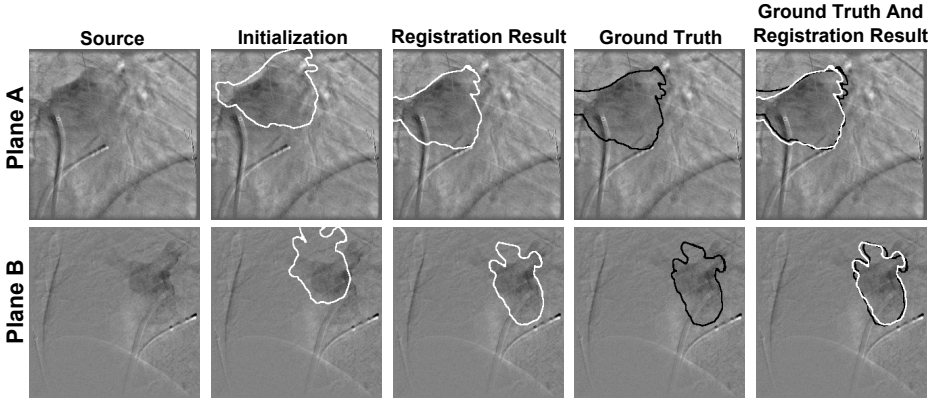

Fig. 7. Case C135 registration result compared with ground truth. The projection distance error for this registration is 4.69mm.

4 Discussion and Conclusion

We presented a mesh-based 2D/3D registration algorithm that can successfully register meshes derived from 3D volumes to fluoroscopic images. The algorithm has the potential to provide near real-time registration. It is especially useful in applications where a 3D mesh is available pre-operatively.

In cases of CA for AF, the fluoroscopic images must contain contrast liquid in order to be used for registration. This means that the algorithm cannot continuously update the registration during the whole operation. However, the registered LA mesh at the time of contrast liquid injection can be used as an initialization for follow-up tracking methods that do not require the presence of contrast liquid [6].

It is not clear if the main source of error is due to the algorithm itself, or to the conditions of the experiment. An important source of error could come from inexact projective geometry and ground truth positions since they were found by visual inspection. The difficulty to evaluate the registration result visually is highlighted in figures 6 and 7.

In the future, we plan to use fully calibrated projection systems and ground truth positions obtained by a medical expert. To get higher precision, we plan to modulate the local cost in function of the mesh's thickness. This will also allow bypassing the manual cutting of the PVs because their thinness will result in a low or null cost for that zone. We also plan to experiment with different similarity measures, including gradient correlation and histogram matching.

References

1. Sra, J., Narayan, G., Krum, D., Malloy, A., Cooley, R., Bhatia, A., Dhala, A., Blanck, Z., Nangia, V., Akhtar, M.: Computed tomography-fluoroscopy image integration-guided catheter ablation of atrial fibrillation. Journal of Cardiovascular Electrophysiology 18, 409–414 (2007)
2. Lemieux, L., Jagoe, R., Fish, D.R., Kitchen, N.D., Thomas, D.G.T.: A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. Medical Physics 21, 1749–1760 (1994)
3. Chan, T., Vese, L.: Active contours without edges. IEEE Transactions on Image Processing 10(2), 266–277 (2001)
4. van de Kraats, E., Penney, G., Tomazevic, D., van Walsum, T., Niessen, W.: Standardized evaluation methodology for 2-d-3-d registration. IEEE Transactions on Medical Imaging 24(9), 1177–1189 (2005)
5. Miao, S., Liao, R., Zheng, Y.: A hybrid method for 2-d/3-d registration between 3-d volumes and 2-d angiography for transcatheter aortic valve implantation (tavi). In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1215–1218 (2011)
6. Brost, A., Liao, R., Strobel, N., Hornegger, J.: Respiratory motion compensation by model-based catheter tracking during ep procedures. Medical Image Analysis 14(5), 695–706 (2010)