

## Research Article

# Contrast-Based 3D/2D Registration of the Left Atrium: Fast versus Consistent

Matthias Hoffmann,<sup>1</sup> Christopher Kowalewski,<sup>2</sup> Andreas Maier,<sup>1,3</sup> Klaus Kurzidim,<sup>2</sup>  
Norbert Strobel,<sup>4</sup> and Joachim Hornegger<sup>1,3</sup>

<sup>1</sup>Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, 91058 Erlangen, Germany

<sup>2</sup>Cardiac Arrhythmia Service, Hospital Barmherzige Brueder, 93049 Regensburg, Germany

<sup>3</sup>Erlangen Graduate School in Advanced Optical Technologies (SAOT), 91052 Erlangen, Germany

<sup>4</sup>Siemens Healthcare GmbH, 91301 Forchheim, Germany

Correspondence should be addressed to Matthias Hoffmann; [matthias.hoffmann@cs.fau.de](mailto:matthias.hoffmann@cs.fau.de)

Received 30 October 2015; Revised 7 February 2016; Accepted 7 February 2016

Academic Editor: Richard H. Bayford

Copyright © 2016 Matthias Hoffmann et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For augmented fluoroscopy during cardiac ablation, a preoperatively acquired 3D model of a patient's left atrium (LA) can be registered to X-ray images recorded during a contrast agent (CA) injection. An automatic registration method that works also for small amounts of CA is desired. We propose two similarity measures: The first focuses on edges of the patient anatomy. The second computes a contrast agent distribution estimate (CADE) inside the 3D model and rates its consistency with the CA as seen in biplane fluoroscopic images. Moreover, temporal filtering on the obtained registration results of a sequence is applied using a Markov chain framework. Evaluation was performed on 11 well-contrasted clinical angiographic sequences and 10 additional sequences with less CA. For well-contrasted sequences, the error for all 73 frames was  $7.9 \pm 6.3$  mm and it dropped to  $4.6 \pm 4.0$  mm when registering to an automatically selected, well enhanced frame in each sequence. Temporal filtering reduced the error for all frames from  $7.9 \pm 6.3$  mm to  $5.7 \pm 4.6$  mm. The error was typically higher if less CA was used. A combination of both similarity measures outperforms a previously proposed similarity measure. The mean accuracy for well contrasted sequences is in the range of other proposed manual registration methods.

## 1. Introduction

Atrial fibrillation is the most common heart arrhythmia affecting around 2.2 million people in the USA. A possible treatment option is catheter ablation, which is a minimally invasive procedure. Depending on the ablation device used, it is carried out using either electroanatomic mapping systems, an X-ray guided approach, or a combination of both. X-ray guidance is, for example, required if a cryoballoon is used for ablation as current cryoballoons cannot be directly located by electroanatomic mapping systems. In this paper, we focus on X-ray guided approaches. One weakness of X-ray imaging is poor soft-tissue contrast. As a consequence, the left atrium (LA) can only be seen if contrast agent (CA) is injected. However, to reduce the risk of contrast-induced nephropathy, physicians try to keep the use of CA to a minimum often

highlighting only a part of the left atrium. To provide orientation to the physician when no CA is present, a model of the patient's LA, for example, generated by a preoperative CT or MRI scan of the patient can be overlaid [1]. The use of such an overlay was found to reduce procedure time and fluoroscopy time significantly [2]. Since a preprocedural 3D scan is often acquired to obtain prior knowledge about a patient's anatomy of the left atrium and to rule out unusual pulmonary vein configurations, it is readily available for augmented fluoroscopy applications. In general, however, the coordinate systems of the preprocedurally acquired 3D heart model and the patient during the intervention differ, and a registration step is usually needed. Today, this registration is usually carried out manually. If the 3D heart model was acquired using CT, several authors proposed to use further landmarks in addition to the LA for registration having

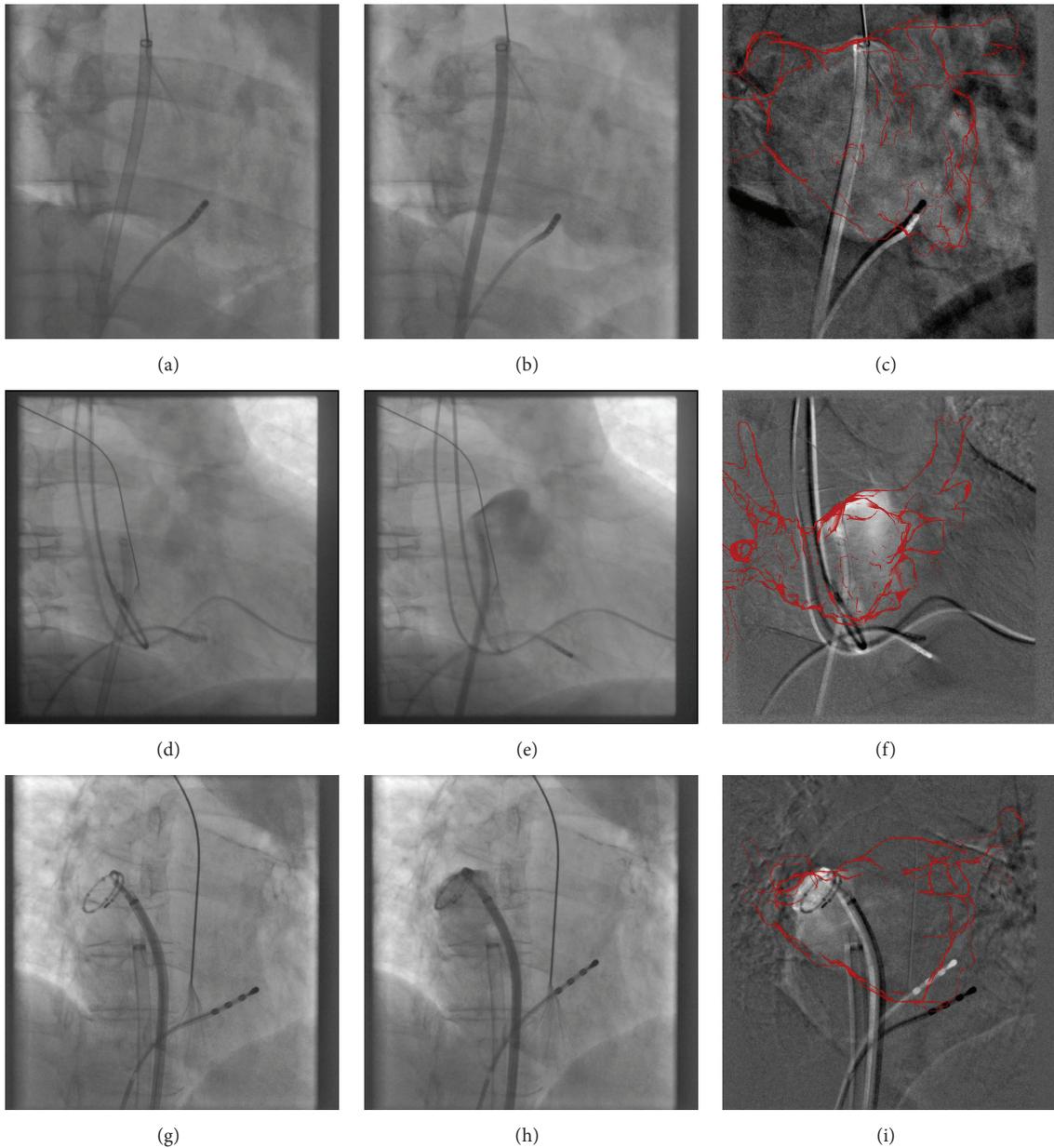


FIGURE 1: Three different contrast agent injections. The left column shows an uncontrasted frame, the middle column shows a contrasted frame of the sequence, and the rightmost column shows a subtraction of both images. The manually established ground-truth position of the left atrium is overlaid with red. The LA in image (c) appears only faintly contrasted and the CA just flows into the ventricle. On the other hand, image (f) shows disturbing motion artifacts caused by leads and wires associated with an implanted device. In image (i), the contrast agent is injected into a pulmonary vein and catheter motion artifacts are visible. While the spine in the center of this image vanishes, soft-tissue motion artifacts related to the lungs are visible at the left and right border, respectively.

obtained associated 3D models by segmentation. Examples are the carina [3], the coronary sinus (CS) [4, 5], the spine [6], or a combination of spine and heart anatomy [7]. For 3D volumes acquired by MRI, for example, MRI angiographies (MRAs), however, a segmentation of the CS or the carina can be very challenging or has not yet been adopted for routine clinical procedures. Fortunately, registration based on contrast injection has been shown to be a fast and accurate alternative [8]. A good moment to obtain contrast-enhanced X-ray images is after the transseptal puncture, in particular if

physicians use contrast injections to verify puncture success [9]. Example images of such injections are shown in Figure 1 to illustrate the images available. Since manual registration either needs the attention of the treating physician or requires a trained assistant, more automatic registration methods are desirable.

(1) *Related Work.* There is a significant body of research on registration of 3D objects to 2D fluoroscopic images, for example, for bones [10–12], implants, joints [13], or vessels

[14]. An overview is given by Markelj et al. [15]. Compared to implants, a registration of the LA is more complicated for at least two reasons: First, for implants and bones, all parts of the object are visible. During a contrast injection, however, only a part of the left atrium may appear under X-ray.

Second, depending on the amount of CA injected, the overall LA visibility may be poor. For example, in our case EP physicians use contrast at the beginning of the procedure to verify the success of the transseptal puncture. It may also be used later to enhance the anatomy, for example, to make sure that a catheter, for example, a circumferential mapping catheter or a cryoballoon catheter, was placed correctly. This is different to vessel angiography. In this case, higher amounts of contrast are injected to derive diagnostic information, for example, about a stenosis. For ablation procedures in the LA, the CA density and so the visibility of the LA under X-ray may be poorer as the small amount of contrast is injected into the high blood volume of the LA. As a consequence, further effort is needed to develop robust registration methods that can also be applied if CA is used sparingly.

Currently, very few publications deal with registration based on contrast agent: In a first approach towards automatic LA registration, Thivierge-Gaulin et al. [16] tried to find a 3D pose of a model such that its projected shadow matches the contrasted area in a selected image, enhanced by digital subtraction angiography (DSA), best. However, if only a small amount of contrast agent is injected into a somewhat large chamber such as the left atrium, this approach will not lead to a distinct optimum, because the anatomy may not be fully opacified.

Based on CT images, a second approach by Zhao et al. [17] relied on digitally rendered radiographies (DRRs) of the segmented left atrium. The rendered image was compared to a DSA image using normalized gradient correlation. This approach uses a weighting scheme that puts the focus on the roof of the LA. How well this method performs for injections into other areas of the LA is still unclear.

Besides contrast-based registration, 3D data and 2D data can also be aligned using devices. A feature-based method by Sra et al. [4] uses a segmentation of the coronary sinus (CS) in a CT volume to register a 3D LA model to a single 2D fluoroscopic image. A further approach by Brost et al. [5] uses a segmentation of the CS in an MRI volume and a 3D reconstruction of the CS catheter from two fluoroscopic images. Unfortunately, how well the CS can be extracted from a 3D MRI data set depends very much on the MRI scan protocol. Furthermore, due to the strong motion of the CS, it is difficult to relate the position of the CS catheter to the position of the LA, especially for patients having no sinus rhythm but heart arrhythmia [18, 19].

(2) *Contributions and Outline of This Work.* At the beginning of Section 2, we discuss the use of the normalized cross-correlation for registration of a 3D object to 2D X-ray projections based on extracted contrast agent. Afterwards, we propose two new registration techniques for contrast agent-based registration. They require a segmented left atrium as input and can, in contrast to [17], accommodate both CT and MRI data. The first method is described in Section 2.2.

Here, we take explicitly apparent edges extracted from a 3D model and compare them to LA edges present in the fluoroscopic images as proposed by [10, 11] for automatic registration of bones and by [20] for manual registration of the LA, respectively. Although our method is conceptually similar to a DRR generation followed by gradient correlation, this calculation can be carried out on a GPU much more quickly than a DRR.

The second contribution described in Section 2.3 is the introduction of a novel similarity measure for biplane X-ray that is tailored to cases in which only parts of an object are visible. Based on a 3D model of the LA, our second method estimates the contrast agent distribution inside the 3D object from a simultaneously acquired pair of fluoroscopic images taken under two different view angles. Then we evaluate how consistent the contrast agent distribution estimate (CADE) is with the acquired fluoroscopic images. As the CADE depends on the transformation used for registration, the transformation leading to the most plausible CADE is used as final position estimate. A brief description of both methods was previously published by us [21].

In Section 2.4 we propose to treat the registration results for each frame of the sequence no longer as independent. Based on our previous publication [22], we use a Markov chain approach to exploit the temporal dependency between successive frames instead.

In Section 3, we evaluate a similarity measure based on the projected shadow which is close to the approach by Thivierge-Gaulin et al. [16] and the two new methods as well as combinations of them. The results are discussed in Section 4.

## 2. Registration Method

For registration, two X-ray sequences showing a CA injection are used. These sequences are acquired simultaneously from two different angles using a biplane system. For each plane, the projection matrix that describes the X-ray camera setup is known. We denote the associated projection operator by  $P_A$  and  $P_B$  for the A-plane and the B-plane of the system, respectively. We also assume that a 3D model of the patient's LA is available, either as a triangle mesh or as a binary volume, as they can be converted into each other.

*2.1. Contrast Agent Extraction.* The contrasted area is found based on a difference image (DSA)  $\mathbf{I}_{\text{DSA}} = \mathbf{I}_u - \mathbf{I}_c$ ,  $\mathbf{I} \in \mathbb{R}^{m \times n}$  involving a frame  $\mathbf{I}_c$  that contains contrast agent and an uncontrasted frame  $\mathbf{I}_u$ . To distinguish between contrasted and uncontrasted frames, either manual annotation or a learning based method can be used, for example, the method described in [23]. Depending on the chosen contrasted frame  $\mathbf{I}_c$ ,  $\mathbf{I}_{\text{DSA}}$  may contain artifacts, for example, due to motion of the diaphragm or from catheters, if they are at different positions in  $\mathbf{I}_u$  and  $\mathbf{I}_c$ . Such motion artifacts depend, unlike the information about contrast agent, to a large degree on the choice of  $\mathbf{I}_u$ . For example, if the catheters in  $\mathbf{I}_u$  are at the same position as in  $\mathbf{I}_c$ , their intensities cancel out. Otherwise,  $\mathbf{I}_{\text{DSA}}$  has high positive values at the position of the catheter in  $\mathbf{I}_c$ .

and high negative values at the position of the catheter in  $\mathbf{I}_u$ . To keep motion artifacts to a minimum, we propose a best reference selection, which chooses an appropriate reference frame  $\widehat{\mathbf{I}}_u$  that matches the chosen contrasted frame  $\mathbf{I}_c$  as much as possible. Out of all uncontrasted frames, that frame  $\widehat{\mathbf{I}}_u$  is selected which minimizes the  $L_1$ -norm of the resulting DSA image:

$$\widehat{\mathbf{I}}_u = \arg \min_{\mathbf{I}_u} \sum_{x=0}^n \sum_{y=0}^m |\mathbf{I}_u(x, y) - \mathbf{I}_c(x, y)|. \quad (1)$$

By following (1), we get frames for which the catheters and the diaphragm cancel out as much as possible. See Figure 3(b) for an example. In  $\mathbf{I}_{\text{DSA}}$ , only pixels with positive values contain contrast agent. To extract them, we set the intensity of pixels with negative value to 0. Afterwards, we compute a filtered image  $\mathbf{I}_f$  by applying a median filter with a large kernel size. Smaller structures, for example, caused by motion artifacts that remained despite the optimized choice of  $\mathbf{I}_u$ , do not pass this filter, and the noise in the contrasted area is reduced as well. Finally, a binary image  $\mathbf{I}_{\text{thr}}$  of the filtered image  $\mathbf{I}_f$  is computed using a threshold at  $\mu_f + \sigma_f$ , where  $\mu_f$  and  $\sigma_f$  denote the mean and standard deviation of  $\mathbf{I}_f$ , respectively. Thus, a contrasted pixel  $\mathbf{p} \in \mathbb{R}^2$  is indicated by  $\mathbf{I}_{\text{thr}}(\mathbf{p}) = 1$ .

A previous approach [16] tried to find a transformation  $T$  of the 3D model such that the projected shadows  $\mathbf{S}_T^A, \mathbf{S}_T^B$  of the model into the A-plane and the B-plane of a biplane C-arm system fit best to the contrasted region. Using the normalized cross-correlation (NCC), denoted as  $\rho_n$ , of two images  $\mathbf{I}_1, \mathbf{I}_2$  with corresponding mean values  $\mu_1, \mu_2$  and standard deviations  $\sigma_1, \sigma_2$

$$\rho_n(\mathbf{I}_1, \mathbf{I}_2) = \sum_{x=0}^n \sum_{y=0}^m \frac{(\mathbf{I}_1(x, y) - \mu_1) \cdot (\mathbf{I}_2(x, y) - \mu_2)}{\sigma_1 \cdot \sigma_2}, \quad (2)$$

the similarity of the projected shadow and  $\mathbf{I}_{\text{DSA}}$  can be measured. Instead of  $\mathbf{I}_{\text{DSA}}$ , one can also use the binary version of it, that is,  $\mathbf{I}_{\text{thr}}$ . Then a registration transformation can be estimated by maximizing either one of the two functions

$$\begin{aligned} \rho_{\text{shad}}^{\text{DSA}}(\mathbf{I}_{\text{DSA}}^A, \mathbf{I}_{\text{DSA}}^B, T) &= \rho_n(\mathbf{I}_{\text{DSA}}^A, \mathbf{S}_T^A) \cdot \rho_n(\mathbf{I}_{\text{DSA}}^B, \mathbf{S}_T^B), \\ \rho_{\text{shad}}^{\text{thr}}(\mathbf{I}_{\text{thr}}^A, \mathbf{I}_{\text{thr}}^B, T) &= \rho_n(\mathbf{I}_{\text{thr}}^A, \mathbf{S}_T^A) \cdot \rho_n(\mathbf{I}_{\text{thr}}^B, \mathbf{S}_T^B). \end{aligned} \quad (3)$$

**2.2. Edge Feature.** Unfortunately, a registration approach only based on contrasted area has multiple solutions if the amount of CA is so little that it does not completely fill the LA; see Figure 2. Fortunately, CA is often injected against the roof or into the pulmonary veins. This results in perceivable edges of the contrasted area which can then be used as registration features. Edge-based registration can be carried out using only the silhouette boundary of the projected object [13] (see Figure 2) or all apparent edges [10, 11] (see Figure 3(f)).

We decided to consider all apparent edges to improve robustness against injections of small amounts of CA, as the silhouette-based approach requires that the LA is contrasted in its entirety. For a partially contrasted left atrium, internal

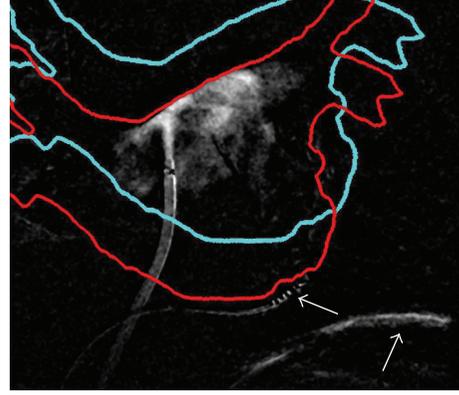


FIGURE 2: A correct (red) and wrong (cyan) registration result. In both cases, the contrasted area is fully inside the projection shadow represented by the colored outlines. Both registration results lead to a similar NCC value when using an area-based feature for automatic registration. Note that motion artifacts could be kept to a minimum thanks to the best reference frame selection. Only some residual artifacts remained in the vicinity of the moving coronary sinus (CS) catheter and the diaphragm; see white arrows.

contours may, however, also appear in the fluoroscopic images. This was already found to be beneficial for manual LA registration [20]. Instead of considering edges implicitly by comparing the DSA image to a DRR using gradient correlation [17], we computed them explicitly. To extract edges in the fluoroscopic images, we used the filtered image  $\mathbf{I}_f$ . After applying a median filter, edge-like variations *inside* the contrasted areas may remain. They, however, correspond rarely to anatomical structures and would trigger a response, if an edge filter was applied. To obtain an edge response only at the boundaries of the contrasted area, the image needs to be homogenized before applying an edge filter. Using a simple threshold method would result in a loss of the intensity drop-off at the boundary which provides important information about the edge intensity. Therefore, we weigh all image pixels by a sigmoid function

$$\mathbf{I}_{\text{sig}}(x, y) = \frac{1}{1 + e^{-(\mathbf{I}_f(x, y) + t) \cdot s}}. \quad (4)$$

The value of  $t$  is set to  $\mu_f - \sigma_f$ , and the parameter  $s$  depends on the pixel intensity range of the input image. An example of  $\mathbf{I}_{\text{sig}}$  is given in Figure 3(d). Finally,  $\mathbf{I}_{\text{sig}}$  is filtered using a derivative of Gaussian (DOG) filter to obtain the edge image  $\mathbf{I}_{\text{DOG}}$ . The kernel size of the DOG filter is set to a large value to get a smooth similarity measure; see Figure 3(e).

The projection of the 3D triangle mesh edges into 2D is done differently than in [10, 20]. We rendered the whole surface mesh and, depending on the viewing direction  $\mathbf{d}$  and the surface normal  $\mathbf{n}$  at a point, we set the opacity  $o$  of projected triangles to  $o = 1 - |\mathbf{d} \circ \mathbf{n}|$ ; see Figure 3(f) for an example. By doing so, areas that are parallel to the imaging plane are rendered transparent while areas with a normal vector orthogonal to the viewing direction are rendered opaque. The similarity between edges extracted from the

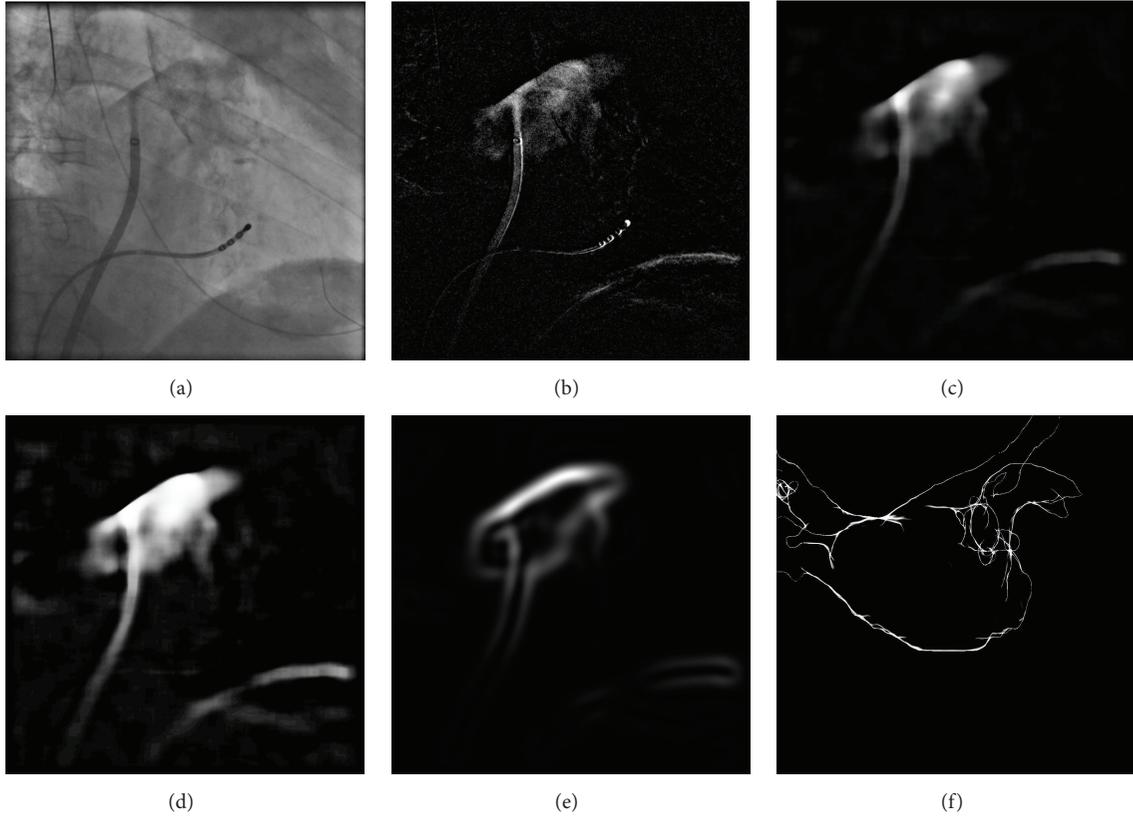


FIGURE 3: Using the original image (a), a DSA image (b) is computed. After median filtering (c), the remaining motion artifacts from the catheter have vanished. Only the boundary at the diaphragm remained. Afterwards, all pixels are weighted by a sigmoid function (d) to get a more homogeneous value distribution inside the contrasted area. Edges are extracted using derivatives of Gaussian with a large kernel size (e). Finally, the similarity to the rendered edges (f) is evaluated.

fluoroscopic images and the edge images  $\mathbf{E}_T^A, \mathbf{E}_T^B$  rendered from the 3D model transformed by  $T$  is measured by

$$\rho_{\text{edge}}(\mathbf{I}_{\text{DOG}}^A, \mathbf{I}_{\text{DOG}}^B, T) = \rho_n(\mathbf{I}_{\text{DOG}}^A, \mathbf{E}_T^A) \cdot \rho_n(\mathbf{I}_{\text{DOG}}^B, \mathbf{E}_T^B). \quad (5)$$

**2.3. Contrast Agent Distribution Estimation (CADE).** Previous approaches [16, 17] for LA registration searched for a rigid transformation of the LA such that either its projected shadow or its DRR fit to the contrasted area in both fluoroscopic images. 3D information was taken into account insofar as the resulting projections came from the same 3D position of the model. However, such an approach does not necessarily guarantee that corresponding objects in both fluoroscopic images are matched to the same 3D structure of the LA. More precisely, the registration result could be such that, in plane A, the contrast agent is located in a left pulmonary vein (PV) whereas, in plane B, the contrasted area corresponds to a right PV. This is possible as for a given 2D registration in one plane the 2D registration in the other plane has one degree of freedom, which corresponds to an out-of-plane motion in the first plane. An illustration of this problem can be found in Figure 4.

To solve this problem, we compute for a given transformation  $T$  a CADE inside the LA using binary reconstruction. Then,  $T$  is optimized such that the contrast agent distribution

estimate is most consistent with the projection images. More precisely, a voxel  $\mathbf{v}$  is estimated as contrasted if it satisfies all of the following conditions: (a) the voxel  $\mathbf{v}$  transformed by  $T$  is projected on a contrasted pixel in plane A, (b)  $\mathbf{v}$  is projected on a contrasted pixel in plane B, and (c)  $\mathbf{v}$  is part of the left atrium (as contrast agent can only be found inside the LA). To compute the CADE, we define the indicator function

$$\chi(\mathbf{v}) = 1 \iff \mathbf{v} \in \mathbb{R}^3 \text{ is inside the left atrium.} \quad (6)$$

Given the binary images  $\mathbf{I}_{\text{thr}}^A$  and  $\mathbf{I}_{\text{thr}}^B$  with corresponding projection operators  $P_A, P_B$  and the indicator function  $\chi(\mathbf{v})$ , the CADE  $C_T^{3D}$  for a transformed voxel,  $T(\mathbf{v})$ , can be computed as

$$C_T^{3D}(\mathbf{v}) = \mathbf{I}_{\text{thr}}^A(P_A(T(\mathbf{v}))) \cdot \mathbf{I}_{\text{thr}}^B(P_B(T(\mathbf{v}))) \cdot \chi(\mathbf{v}) \quad (7)$$

for a given rigid transformation  $T$ . This product is the mathematical equivalent of the three conditions introduced above.

If  $T$  is chosen suboptimally, the resulting 3D CADE will be inconsistent with the CA observed in the 2D images. That is, a pixel in the 2D image is contrasted but no corresponding voxel along its projection ray is estimated as contrasted. This can be due to the following reasons as shown in Figure 5: (a) the projection ray from a contrasted pixel does not intersect

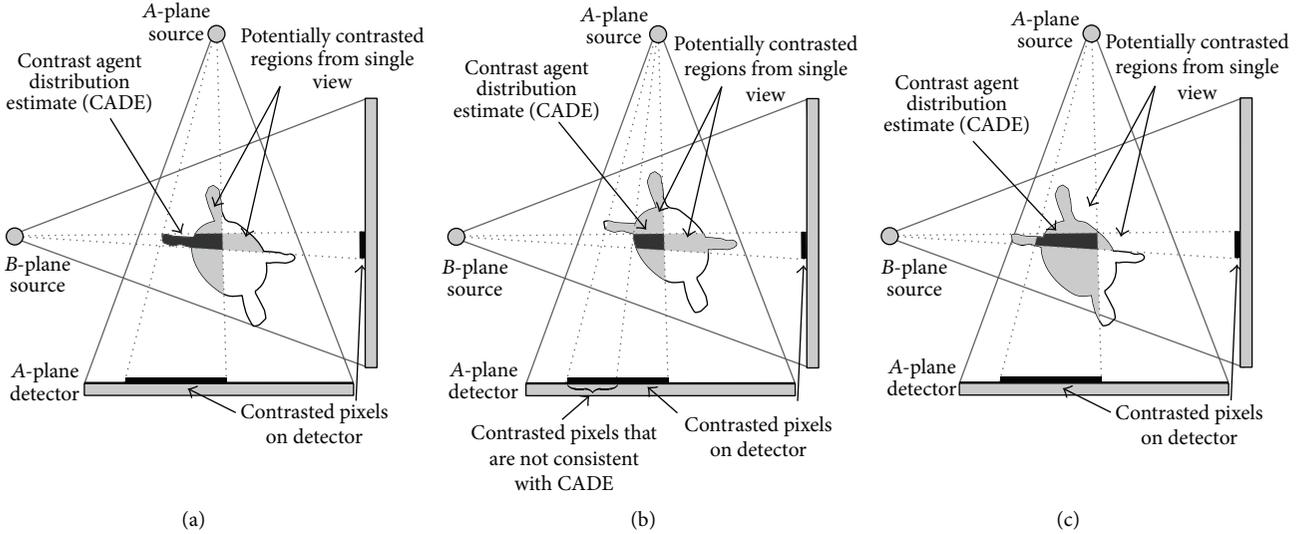


FIGURE 4: Top-down view on a left atrium and associated contrast-enhanced projections showing a correct registration (a) and two misregistrations (b) and (c). When looking at the contrasted pixels on each detector independently, all LA positions seem to be feasible. Considering the LA position in (b), the area, which can contain contrast agent according to the combination of both detectors, is partially outside the LA. For the CADE, however, only contrast agent inside the LA is considered. So the registration in (b) gives rise to contrasted pixels on the A-plane detector that cannot be explained by the CADE. This will result in a low CADE value and indicate misregistration. There exist, however, also misregistrations that can lead to a consistent CADE as shown in (c). This is why a combination of CADE and edge-based methods yields improved results.

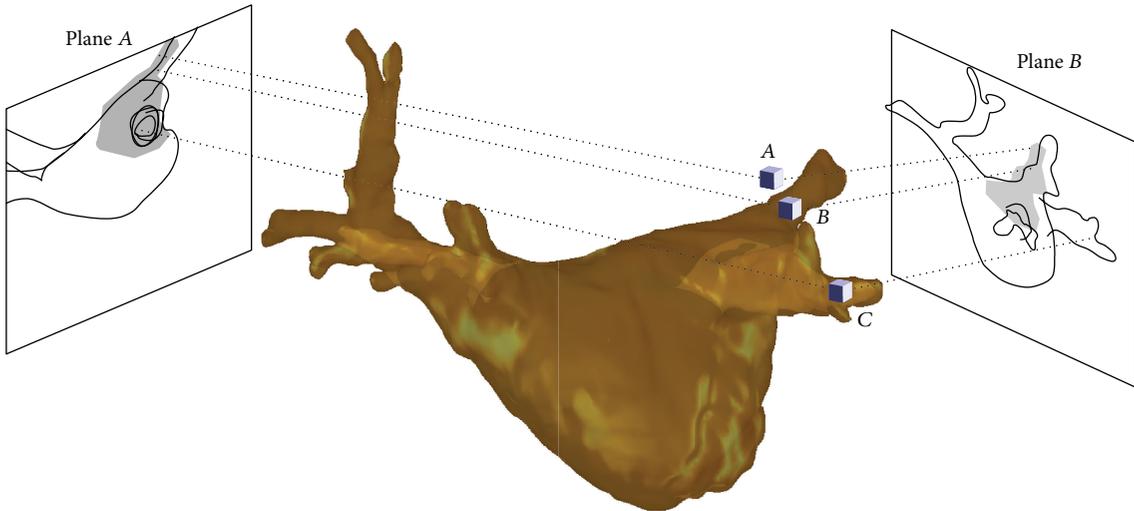


FIGURE 5: For the transformation shown, only voxel B is estimated as containing CA as it is inside the LA and contrasted in both planes. Voxel C is only contrasted in plane A but not in plane B. Voxel A is filled with contrast in both planes, but it is outside of the LA and therefore considered as uncontrasted. If the LA is moved such that its projections in A and B cover the contrasted area, this voxel will also be considered as contrasted, hence, increasing the consistency with the CADE.

the left atrium as the LA has not been placed at the proper position yet; (b) the projection ray hits the LA, but all voxels intersected by this ray cannot contain CA because their corresponding pixels in the other plane are uncontrasted. Additional inconsistencies are introduced by pixels which are erroneously labeled as contrasted, for example, due to motion artifacts. To verify the validity of the CADE, we compute binary 2D images  $C_T^A$ ,  $C_T^B$  by forward projecting all contrasted

voxels in  $C_T^{3D}$  using  $P_A, P_B$ . We assess the consistency of the CADE for the given transformation  $T$  by computing the similarity between the fluoroscopic images and the projected CADE by

$$\rho_{CADE}(\mathbf{I}_{thr}^A, \mathbf{I}_{thr}^B, T) = \rho_n(\mathbf{I}_{thr}^A, C_T^A) \cdot \rho_n(\mathbf{I}_{thr}^B, C_T^B). \quad (8)$$

Alternatively, the number of contrasted voxels inside the volume could be maximized. The set of potentially contrasted

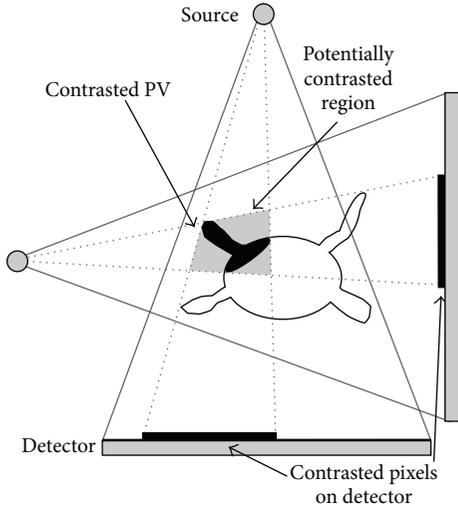


FIGURE 6: Top-down view on a left atrium with a contrasted pulmonary vein. The gray region denotes the area which may contain contrast agent based on the contrast agent observed in the 2D images. If the number of contrasted voxels in the volume was to be optimized, the optimization process would find a transformation such that as much of the potentially contrasted region as possible was included within the LA causing a registration bias towards larger LA regions.

voxels is defined by the intersection of the projection ray bundles of all contrasted pixels from views A and B, respectively. Maximizing the number of contrasted voxels would lead to a transformation such that as much as possible of this set is contained inside the LA. If, for example, a pulmonary vein was filled with CA, the transformation would be chosen such that the set of possible contrasted pixels was located in the main body of the LA which provides the most space to include most possibly contrasted voxels; see Figure 6. To avoid this registration bias towards large structures inside the LA, we decided to optimize the consistency to the 2D images.

**2.4. Confidence-Based Temporal Markov Filtering.** To exploit dependencies between successive frames of a sequence, we model the position of the LA model in 3D as a time-dependent continuous Markov chain of the first order as proposed by us previously [22]. The states are transformations, that is, a translation and a rotation of the model. The parameter  $\tau_i$  denotes the actual transformation of the LA in the  $i$ th contrasted frame,  $T_i$  refers to the estimate of the transformation for the  $i$ th frame, and  $P(\tau_i = T_i)$  is the probability that for frame  $i$  the transformation  $T_i$  is observed. For convenience, we define  $P(T_i) = P(\tau_i = T_i)$ . The transition probability from one state into another is independent of the frame number  $i$ . It is denoted as  $P(T_i \rightarrow T_{i+1})$ . Finally, a sequence of transformations  $T_1, \dots, T_n$  is to be determined such that the term

$$P(T_1, \dots, T_n) = P(T_1) \cdot \prod_{t=2}^n (P(T_{t-1} \rightarrow T_t) \cdot P(T_t)) \quad (9)$$

is maximized.

Transition probabilities and state probabilities control the filtering differently: Due to the transition probabilities, small motions of the LA from a frame to its next frame are preferred. This results in an averaging of the LA transformation over time. The state probability determines the impact of the averaging, that is, how much the registration result is changed by the filtering. We will model the state probability based on a confidence measure such that frames with high-confidence registration results are less subject to temporal filtering. If the confidence value is low, that is, the estimated error of the registration result is higher, we want to rely more on the results of the previous and next frame. Thus, high temporal averaging is performed for frames with a low confidence value. The state probabilities and the transition probability are determined as follows.

**2.4.1. State Probability.** The contrast agent visible in frame  $i$  determines the probability of the LA for being transformed by  $T_i$  in this frame. If the dependency on other frames is ignored, the most probable transformation  $T_i'$  is the transformation obtained by optimizing the similarity measures  $\rho$  defined in (3), (5), and (8) or a combination of them.

The similarity measure  $\rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, T_i')$  depends on the contents of images  $\mathbf{I}^{A,i}, \mathbf{I}^{B,i}$ . Zhao et al. suggested using the value  $\rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, T_i')$  as a confidence measure: from all  $N$  registration results the frame  $i$  where  $\rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, T_i')$  is maximum should be selected as registration result for the complete sequence. We confirmed a correlation of  $\rho$  and the registration error. The value of  $\rho$  can therefore be used as a confidence measure. By linear regression, a function  $e(\rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, T_i'))$  can be determined to estimate the error based on the value of  $\rho$ . With the transformation  $T_i'$  of the LA found during optimization for frame  $i$ , the probability  $P(T_i)$  can be modelled as normal distribution

$$P(T_i) = \mathcal{N}(T_i; T_i', \Sigma_i). \quad (10)$$

The confidence range is incorporated by setting the covariance matrix  $\Sigma_i$  to  $\mathbf{1} \cdot e(\rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, T_i'))$ .

**2.4.2. Transition Probability.** The transition probability  $P(T_{t-1} \rightarrow T_t)$  states how likely a movement of the LA is from frame  $t-1$  to  $t$ . Over multiple breathing cycles, the LA moves about a mean position. So the mean transformation is the identity. The likelihood of a transformation change depends on the magnitude of the change. The larger the change in translation and rotation is, the less likely the transformation transition is. To account for different frame rates, we consider the translational and rotational velocity  $\mathbf{v} \in \mathbb{R}^6$  which comprises three translational velocities and three rotational velocities. In a training step, velocity vectors are computed for annotated data. The covariance matrix  $\Sigma_v$  of the velocities gives an estimate for how likely a transition from one transformation to the next is. We model the probability of a transition  $T_{t-1} \rightarrow T_t$  as a normal distribution

$$P(T_{t-1} \rightarrow T_t) = \mathcal{N}((T_t - T_{t-1}) \cdot r; 0, \Sigma_v). \quad (11)$$

Besides the frame rate  $r$ , the transition probability depends also on the current breathing phase. Compared to breathing

motion, cardiac motion can be neglected as it rather deforms the LA. If information on the breathing phase can be estimated [24], superior motion should be, for example, more likely during the inhale phase. During exhalation phase, inferior motion should have a higher probability. To account for breathing motion, for example, the mean value and the covariance matrix could be estimated separately for each stage of the breathing cycle.

$$T_1^*, \dots, T_n^* = \arg \max_{T_1, \dots, T_n} P(T_1, \dots, T_n), \quad (12)$$

$$= \arg \max_{T_1, \dots, T_n} - \frac{1}{2} \left( (T_1 - T_1')^T \Sigma_1^{-1} (T_1 - T_1') + \sum_{i=2}^n \left( (T_i - T_i')^T \Sigma_i^{-1} (T_i - T_i') + r \cdot (T_i - T_{x_{i-1}})^T \Sigma_v^{-1} (T_i - T_{i-1}) \cdot r \right) \right). \quad (13)$$

This convex optimization problem can be solved using the BFGS method.

### 3. Experiments and Results

We retrospectively evaluated our method on 21 clinical biplane X-ray sequences from 10 different patients. All of the patients provided their informed consent for the analysis of their clinical data. For all patients, a segmentation of their left atrium from a preoperatively acquired MRI scan was available as a triangle mesh. When using the CADE measure, this mesh was converted into a binary volume. The data set comprised 11 sequences showing an initial contrast agent injection where 15 mL CA was injected into the LA center through a sheath to verify the success of the transseptal puncture. Besides the sheath, only a coronary sinus catheter was present. There were 10 more sequences showing subsequent injections to verify catheter placement. In these cases, about 10 mL CA was injected and additional catheters were present. All X-ray angiography sequences were acquired during normal breathing using a standard acquisition protocol. This resulted in a total number of 133 contrasted frames. For our experiments, the first contrasted frame was determined manually. For 129 frames, reference registrations performed by three clinical experts were available; for the remaining four frames registrations performed by two clinical experts were available. Reference 3D registrations were established by manually shifting the mesh in each of two orthogonal views. The resulting 3D translation was refined this way until a good match to the associated contrast distributions as seen in the two 2D X-ray projections had been found. These reference registrations covered only 3D translation for the following reasons: First, patients were positioned head first, supine, during both preoperative and intraoperative imaging. This patient positioning rules out large degrees of rotation a priori. Furthermore, given the small amounts of contrast injected in our cases, the remaining small rotations were very difficult to detect. This made it practically impossible for our clinical experts to reliably correct them. This is why we decided to evaluate our approach without rotation.

2.4.3. *Most Probable State Sequence.* The most likely state sequence,

can be found using a log-likelihood method: By applying the logarithm to (9), we get

As initialization for optimization, the 3D model was placed at that 3D position which corresponded to the centers of both 2D images. In some cases, the initialization was more than 30 mm away from the correct solution and beyond the capture range for gradient-based methods. Therefore we applied an octree-like coarse-to-fine scheme where we evaluated several positions at a coarse resolution. At positions in space that yielded a good similarity value, we performed subsequent evaluations at an increasingly finer resolution. The 3D translation  $\hat{\mathbf{t}} = \arg \max_{\mathbf{t}} \rho(\mathbf{t})$  found by the optimization process of the respective objective function  $\rho$  was compared to the mean translation vector  $\mathbf{t}^*$  of the three manual registration results. The distance  $\|\hat{\mathbf{t}} - \mathbf{t}^*\|_2$  was used as error measure. The significance of the results was measured using a Wilcoxon signed-rank test [25] and a significance level of  $p = 0.05$ .

All sequences contained 12-bit images of size  $1024 \times 1024$  pixels; all image processing, including rendering from the 3D model, was performed on the full image size. The kernel size for median filter was 30 pixels; the value  $s$  of (4) was 0.1. The standard deviation of the DOG filter was 24 pixels.

In the evaluation, we compared the similarity measures  $\rho_{\text{shad}}^{\text{thr}}$ ,  $\rho_{\text{shad}}^{\text{DSA}}$ ,  $\rho_{\text{CADE}}$ , and  $\rho_{\text{edge}}$  and the combined similarity measures  $\rho_{\text{shad}}^{\text{thr}} + \alpha \rho_{\text{edge}}$ ,  $\rho_{\text{shad}}^{\text{DSA}} + \alpha \rho_{\text{edge}}$ , and  $\rho_{\text{CADE}} + \alpha \rho_{\text{edge}}$ . We investigated different weightings. Giving both terms equal weights, that is, setting  $\alpha = 1$ , turned out to be a good choice. We did two types of evaluation: an evaluation on all frames and an evaluation using only a single frame of each sequence, namely, the one which provides the best similarity measure.

3.1. *Evaluation on All Frames.* We computed a registration for each contrasted frame and compared the result to the manual registration of the physicians. These results are clinically relevant if the physician requires a registration for a desired frame of his choice, for example, depending on the breathing phase. As potentially every frame could be selected by the physician, the overall registration accuracy should be high. The overall accuracy is also of importance if the results are postprocessed by temporal filtering.

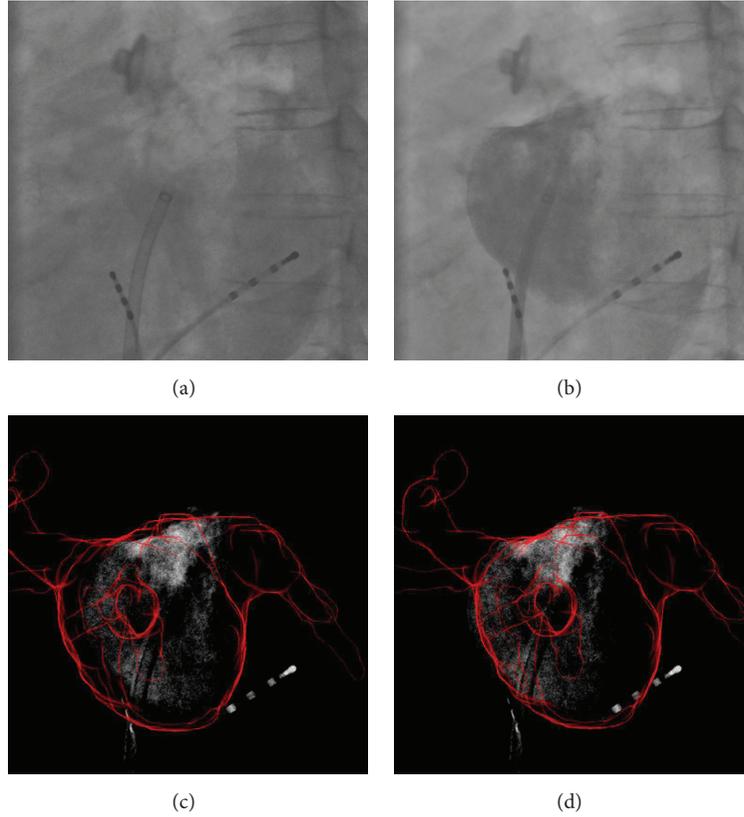


FIGURE 7: Uncontrasted (a) and contrasted (b) frame of an X-ray angiography. The registration result when using  $\rho_{\text{shad}}^{\text{DSA}}$  (c) had an error of 6.7 mm. By using  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  (d), the left border of the LA model fits better to the left edge of the CA and the error reduced to 3.1 mm.

TABLE 1: Translation errors for all frames.

Objective function	Initial injection	Subsequent injections
Thivierge-Gaulin	11.4 ± 8.8 mm	10.4 ± 5.4 mm
$\rho_{\text{shad}}^{\text{DSA}}$	9.3 ± 6.9 mm	9.8 ± 4.8 mm
$\rho_{\text{shad}}^{\text{thr}}$	9.9 ± 7.9 mm	12.1 ± 9.1 mm
$\rho_{\text{edge}}$	12.0 ± 8.9 mm	14.6 ± 8.7 mm
$\rho_{\text{CADE}}$	<b>8.7 ± 6.4 mm</b>	<b>9.0 ± 5.9 mm</b>
$\rho_{\text{shad}}^{\text{DSA}} + \rho_{\text{edge}}$	8.3 ± 6.6 mm	9.6 ± 5.4 mm
$\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$	8.3 ± 6.7 mm	10.5 ± 7.6 mm
$\rho_{\text{CADE}} + \rho_{\text{edge}}$	<b>7.9 ± 6.3 mm</b>	<b>8.8 ± 6.7 mm</b>
Clinical experts	3.3 ± 2.7 mm	3.1 ± 1.7 mm

The evaluation results are presented in Table 1. The overall interuser variability observed in the manual registrations was  $3.2 \pm 2.3$  mm. Considering all sequences,  $\rho_{\text{CADE}}$  performed significantly better than  $\rho_{\text{shad}}^{\text{thr}}$ ,  $\rho_{\text{shad}}^{\text{DSA}}$ , and the state-of-the-art method by Thivierge-Gaulin et al. [16]. Although  $\rho_{\text{edge}}$  gave significantly worse results than all other measures, a combination with  $\rho_{\text{edge}}$  could improve the results of  $\rho_{\text{shad}}^{\text{thr}}$  and  $\rho_{\text{shad}}^{\text{DSA}}$  significantly. For subsequent injections,  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  performed significantly better than  $\rho_{\text{shad}}^{\text{thr}}$ ,  $\rho_{\text{shad}}^{\text{DSA}}$ , their respective combinations with  $\rho_{\text{edge}}$ , and the method by Thivierge-Gaulin et al. [16].

For  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$  and  $\rho_{\text{CADE}} + \rho_{\text{edge}}$ , the error distribution for each sequence is shown in Figure 8, first for all initial injections and then for subsequent injections. In Figure 9, the error distribution is given for each frame number as counted after the CA injection. An example for a result is shown in Figure 7.

**3.2. Evaluation on a Single Frame.** Figure 8 indicates that many sequences have at least one frame with a registration error of less than 5 mm. In fact, there exists such a frame for over 75% of all sequences when using  $\rho_{\text{shad}}^{\text{DSA}}$  or  $\rho_{\text{shad}}^{\text{thr}}$  and for over 85% of the sequences when using  $\rho_{\text{CADE}}$ ,  $\rho_{\text{shad}}^{\text{DSA}} + \rho_{\text{edge}}$ ,  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$ , or  $\rho_{\text{CADE}} + \rho_{\text{edge}}$ . Zhao et al. [17] proposed to automatically find a good frame in each sequence and consider only these single good frames for evaluation. In other words, from all frames of a sequence, a single frame was automatically chosen for registration. From a clinical point of view, this would, however, only make sense if the physician accepted an automatically selected frame for registration or if further motion compensation steps for the other frames were done, for example, using device tracking [26].

Zhao et al. suggested selecting this single frame as follows: For each frame  $i \in [1, m]$  out of the  $m$  contrasted frames, the transform  $\mathbf{t}_i$  that maximizes the similarity measure  $\rho$  for this frame is computed. In a second step, going through all frames of a sequence, the frame with the highest overall similarity

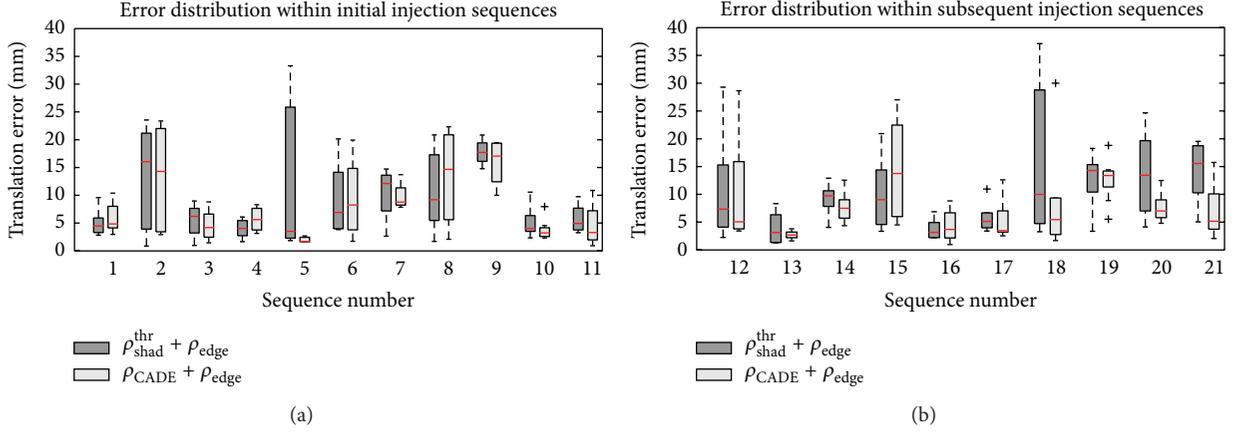


FIGURE 8: Distribution of the errors within each sequence. While there is no significant difference for initial injections (a),  $\rho_{CADE} + \rho_{edge}$  performs significantly better than  $\rho_{shad}^{thr} + \rho_{edge}$  for subsequent injections (b). Subsequent contrast injections are characterized by a smaller amount of contrast and the presence of more catheters making registration more difficult. Example images of sequence 1 are shown in Figures 2 and 3; example images of sequence 2 are shown in Figure 1(a). Here, contrast agent is very faint and in some frames contrasted is already ejected into the left ventricle. Sequence 5 is shown in Figure 1(d). In this case,  $\rho_{shad}^{thr} + \rho_{edge}$  got confused by edges caused by motion artifacts.

measure is picked. The associated transform is denoted by  $\hat{\mathbf{t}}_i$ . The error to the average manual registration result  $\mathbf{t}_i^*$  is then computed as

$$\|\hat{\mathbf{t}}_i - \mathbf{t}_i^*\|_2 \quad \text{with } \hat{\mathbf{t}}_i = \arg \max_i \rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, \mathbf{t}_i), \quad \mathbf{t}_i = \arg \max_{\mathbf{t}} \rho(\mathbf{I}^{A,i}, \mathbf{I}^{B,i}, \mathbf{t}). \quad (14)$$

The evaluation results are presented in Table 2. Recalling Section 3.1, where we found that the CADE method yielded

$$\|\hat{\mathbf{t}}_i - \mathbf{t}_i^*\|_2 \quad \text{with } \hat{\mathbf{t}}_i = \arg \max_i \rho_{shad}^{thr}(\mathbf{I}_{thr}^{A,i}, \mathbf{I}_{thr}^{B,i}, \mathbf{t}_i), \quad \mathbf{t}_i = \arg \max_{\mathbf{t}} \rho_{CADE}(\mathbf{I}_{thr}^{A,i}, \mathbf{I}_{thr}^{B,i}, \mathbf{t}). \quad (15)$$

The results of this frame selection method are denoted by  $\rho_{CADE} \triangleright \rho_{shad}^{thr}$ . We did the same also for the combination with  $\rho_{edge}$ . Although only 10 sequences were available for initial contrast injections,  $\rho_{shad}^{DSA}$ ,  $\rho_{shad}^{thr} + \rho_{edge}$ , and  $(\rho_{CADE} + \rho_{edge}) \triangleright (\rho_{shad}^{thr} + \rho_{edge})$  gave significant better results when compared to  $\rho_{shad}^{thr}$ . Also the performance of  $\rho_{edge}$  was significantly less. For subsequent injections, our proposed method  $\rho_{CADE} \triangleright \rho_{shad}^{thr}$  and the corresponding combination with  $\rho_{edge}$  outperformed the method by Thivierge-Gaulin et al.

**3.3. Temporal Filtering.** The error estimate function  $e(\rho(\cdot))$  and the transition probability covariance matrix  $\Sigma_v$  were trained in a leave-one-patient-out cross-validation. The results for the temporal filtered frames are given in Table 3. For a combination of the CADE and edge similarity measure, Markov filtering reduced the mean error to  $5.7 \pm 4.6$  mm for initial injections and  $6.6 \pm 3.6$  mm for subsequent injections. The respective median errors were 4.0 mm and 5.7 mm.

good results for all frames, the CADE performance for a single frame was, however, not that good. This is why we investigated if a different single-frame selection strategy would lead to better results. Instead of using  $\rho_{CADE}$  both for translation estimation and for best-frame selection, we used  $\rho_{CADE}$  only for estimating the translation  $\mathbf{t}_i$  for each frame  $i$ . Among the obtained translations  $\mathbf{t}_i$ , we selected the frame  $i$  with the corresponding translation  $\hat{\mathbf{t}}_i$  such that  $\rho_{shad}^{thr}$  was maximized. So (14) was modified to

The results obtained by the Markov filtering were in all cases significantly better than the results without filtering.

**3.4. Runtime Performance.** The image preprocessing took 0.5 s on an Intel Xeon E3 with 3.4 GHz and 16 GB RAM. The evaluations of the similarity measures were performed completely on the GPU. On an NVIDIA GeForce GTX 660 the evaluation of  $\rho_{shad}^{DSA}$ ,  $\rho_{shad}^{thr}$ , or  $\rho_{edge}$  took 1.8 ms for a given translation and  $13.4 \pm 3.7$  ms for  $\rho_{CADE}$ . The whole registration for a single frame took 2.9 s for  $\rho_{shad}^{DSA}$  or  $\rho_{shad}^{thr}$  and  $21.5 \pm 5.9$  s for  $\rho_{CADE}$ . For a combination with  $\rho_{edge}$  it took 5.8 s and  $27.1 \pm 5.9$  s, respectively. The runtime of the Markov filtering was  $173 \pm 124$  ms per sequence.

## 4. Discussion and Conclusions

**4.1. Similarity Measures.** Our novel CADE-based method outperformed the shadow based similarity measures  $\rho_{shad}^{DSA}$  and  $\rho_{shad}^{thr}$ , especially for reregistration sequences where only

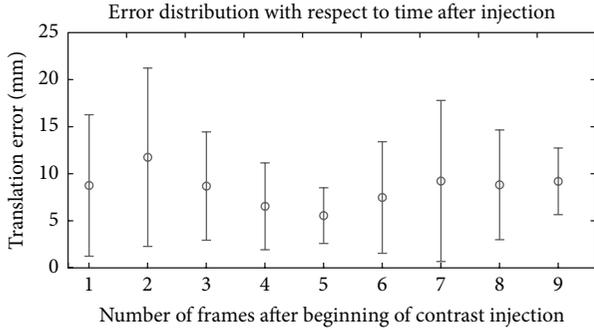


FIGURE 9: Average registration errors of  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  depending on the position of the frame after start of the contrast agent injection calculated over all sequences. Frame 1 denotes the first frame which contained contrast agent. The average error at the second frame after the CA injection was  $11.8 \text{ mm} \pm 9.5 \text{ mm}$ ; the average error at the fifth frame was  $5.5 \text{ mm} \pm 3.0 \text{ mm}$ . The frame rate of all sequences was 7.5 fps. Assuming a heart rate of 60 bpm, one heart cycle takes approximately 7 frames. So a first portion of CA is ejected into the ventricle not later than at the 7th frame. Note that the number of sequences used for calculation was not constant for each frame number as several sequences had fewer than nine contrasted frames.

TABLE 2: Translation errors for a single automatically chosen frame as described in (14) and (15).

Objective function	Initial injection	Subsequent injections
Thivierge-Gaulin	$8.3 \pm 8.2 \text{ mm}$	$10.6 \pm 4.4 \text{ mm}$
$\rho_{\text{shad}}^{\text{DSA}}$	$5.1 \pm 3.8 \text{ mm}$	$9.6 \pm 5.1 \text{ mm}$
$\rho_{\text{shad}}^{\text{thr}}$	$7.1 \pm 4.3 \text{ mm}$	$8.1 \pm 6.6 \text{ mm}$
$\rho_{\text{edge}}$	$14.1 \pm 10.0 \text{ mm}$	$12.1 \pm 9.1 \text{ mm}$
$\rho_{\text{CADE}}$	$6.0 \pm 5.0 \text{ mm}$	<b><math>7.0 \pm 3.8 \text{ mm}</math></b>
$\rho_{\text{CADE}} \triangleright \rho_{\text{shad}}^{\text{thr}}$	<b><math>5.0 \pm 5.0 \text{ mm}</math></b>	$7.7 \pm 4.9 \text{ mm}$
$\rho_{\text{shad}}^{\text{DSA}} + \rho_{\text{edge}}$	$5.1 \pm 4.2 \text{ mm}$	$8.5 \pm 4.7 \text{ mm}$
$\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$	$4.8 \pm 4.6 \text{ mm}$	$7.8 \pm 5.7 \text{ mm}$
$\rho_{\text{CADE}} + \rho_{\text{edge}}$	$6.1 \pm 5.8 \text{ mm}$	$7.6 \pm 5.5 \text{ mm}$
$(\rho_{\text{CADE}} + \rho_{\text{edge}}) \triangleright (\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}})$	<b><math>4.6 \pm 4.0 \text{ mm}</math></b>	<b><math>7.3 \pm 5.2 \text{ mm}</math></b>

a small amount of contrast agent was used. The problem with nondistinct optima for shadow based similarity measures which was already sketched in Figure 2 becomes also apparent in the sections of the optimization function in Figure 10 for  $\rho_{\text{shad}}^{\text{thr}}$ : for a small amount of contrast agent like in Figure 10(a), the optimum is not at the position of the ground truth, but about 15 mm off. Also for secondary injections into small structures such as the pulmonary veins, the objective function for  $\rho_{\text{shad}}^{\text{thr}}$  in Figure 10(c) has a large plateau. In both cases,  $\rho_{\text{CADE}}$  is more distinct, although clear extrema as for bones [12] are not obtained. Especially for secondary injections where CA was often injected into pulmonary veins, registration errors leading to inconsistent results could be avoided by  $\rho_{\text{CADE}}$ . For well-contrasted sequences, the improvement by  $\rho_{\text{CADE}}$  was less as inconsistencies played only a minor role and the shape of the contrasted region in the image was more similar to the projected shadow of the left

TABLE 3: Translation errors for temporal filtered frames.

Objective function	Initial injection	Subsequent injections
Thivierge-Gaulin	$9.5 \pm 6.2 \text{ mm}$	$7.2 \pm 3.0 \text{ mm}$
$\rho_{\text{shad}}^{\text{DSA}}$	$7.5 \pm 4.1 \text{ mm}$	$7.7 \pm 3.5 \text{ mm}$
$\rho_{\text{shad}}^{\text{thr}}$	$7.5 \pm 4.8 \text{ mm}$	$9.1 \pm 5.6 \text{ mm}$
$\rho_{\text{edge}}$	$9.7 \pm 7.1 \text{ mm}$	$12.9 \pm 7.6 \text{ mm}$
$\rho_{\text{CADE}}$	$6.1 \pm 4.1 \text{ mm}$	<b><math>6.2 \pm 3.1 \text{ mm}</math></b>
$\rho_{\text{shad}}^{\text{DSA}} + \rho_{\text{edge}}$	$6.5 \pm 4.3 \text{ mm}$	$7.4 \pm 3.3 \text{ mm}$
$\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$	$6.1 \pm 4.4 \text{ mm}$	$8.1 \pm 3.7 \text{ mm}$
$\rho_{\text{CADE}} + \rho_{\text{edge}}$	<b><math>5.7 \pm 4.6 \text{ mm}</math></b>	$6.6 \pm 3.6 \text{ mm}$

atrium. This becomes also apparent in Figure 10(b) where the shapes of the objective function of  $\rho_{\text{shad}}^{\text{thr}}$  and  $\rho_{\text{CADE}}$  look very similar.

We found that the similarity measure using explicit apparent edges,  $\rho_{\text{edge}}$ , yielded poor results when used on its own. A possible reason is that the objective function of  $\rho_{\text{edge}}$  has several local optima and also the global optimum does not necessarily correspond to the ground-truth position. However,  $\rho_{\text{edge}}$  can improve results significantly when combined with other similarity measures. It often has a more distinct local optimum at the position of the ground truth that facilitates fine registration. This is important as the other similarity measures define a rather plateau-shaped optimal region.

**4.2. Time-Dependency.** Figure 9 shows that the accuracy of the registration depends on when the frame was recorded after contrast agent injection. The best results were achieved at the 5th frame. This was in many cases the frame before the ejection into the ventricle; that is, it contains the most contrast agent. Especially for the first frames, where only little CA was present, registration accuracy was lower. The first frame is an exception, as it has a lower mean error than the second frame. This is probably because the previous frame, that is, the last uncontrasted frame, is used as mask frame for DSA computation which leads to low motion artifacts and a better extraction of the contrasted region.

**4.3. Best-Frame Selection.** We found that, for our registration to work best, a well opacified frame should be selected from the sequence. This strategy was evaluated in the context of the single-frame evaluation by selecting the frame which had the best objective function value. For methods like  $\rho_{\text{shad}}^{\text{DSA}}$  and  $\rho_{\text{shad}}^{\text{thr}}$ , which are based on the projection shadow, the best frame usually corresponded to the most contrasted frame. The measure  $\rho_{\text{CADE}}$ , however, is not related to the amount of contrast agent, but it depends on consistency. When relying on  $\rho_{\text{CADE}}$  for best-frame selection, we get the frame yielding the most consistent registration. This is, however, not necessarily the frame with the most contrast agent. And it is usually the frame with the most contrast agent which provides the most information for registration. We believe that this is the reason why the errors in Table 2 for  $\rho_{\text{CADE}}$  are higher. If, however, from the translations estimated by  $\rho_{\text{CADE}} +$

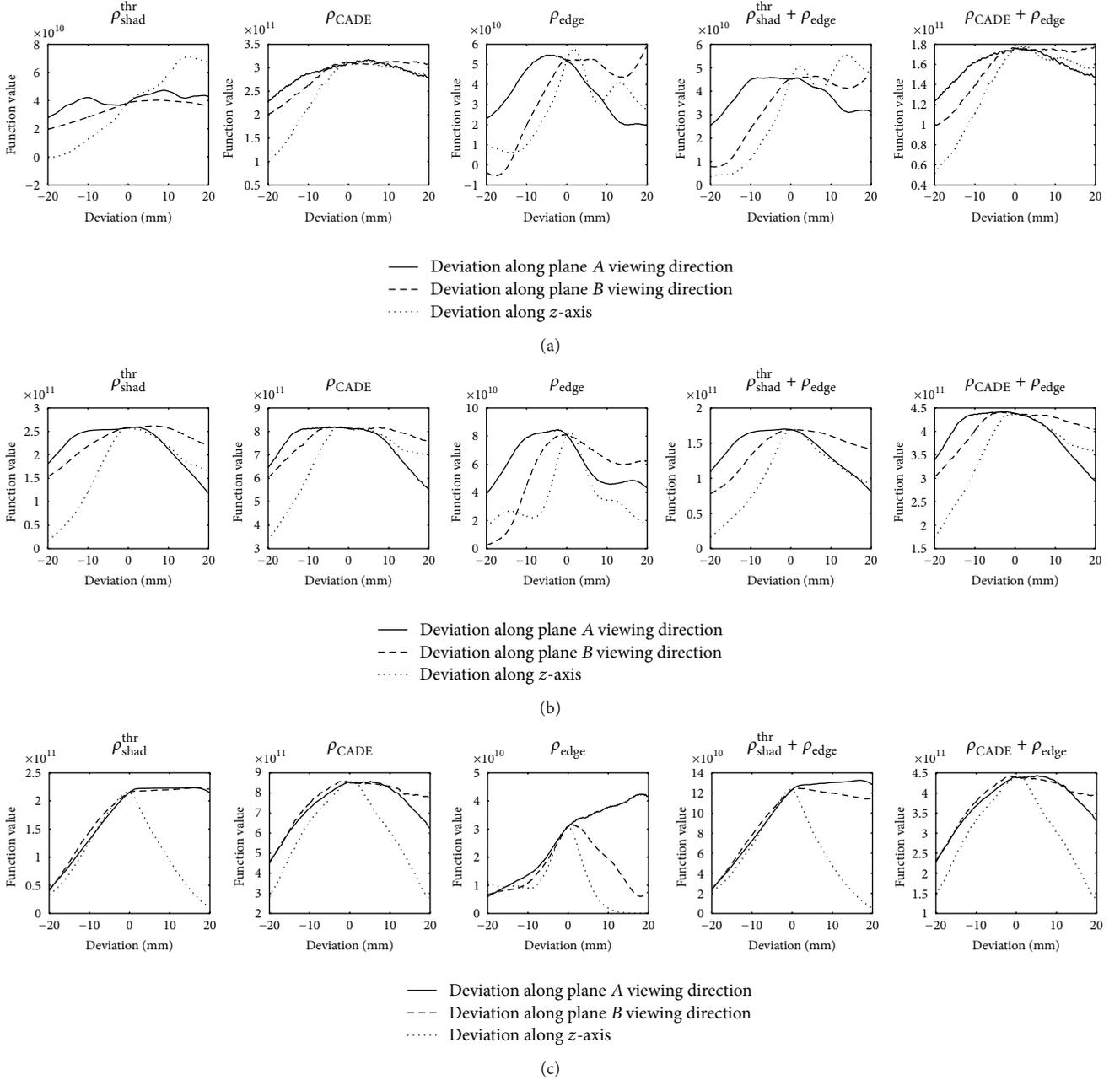


FIGURE 10: Optimization function values along sections through the ground-truth position. The sections were done along the viewing direction of the detector in plane A, the corresponding orthogonal direction in the transverse plane, and the  $z$ -axis of the patient coordinate system which corresponds to the cranial-caudal axis. The intersections are shown for three different frames: (a) one of the first frames of an initial injection (see Figure 3), (b) a frame of the same injection but with large parts of the left atrium contrasted, and (c) a secondary injection into a pulmonary vein (see Figure 1(g)). The plots for  $\rho_{\text{shad}}^{\text{DSA}}$  were left out as they look very similar to those for  $\rho_{\text{shad}}^{\text{thr}}$ .

$\rho_{\text{edge}}$ , the single frame was selected based on  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$ , better results were achieved as now again the frames with the most contrast agent were selected.

Sometimes, the automatic best-frame selection does not provide a frame with a satisfactory result. Still, 85% of all sequences contain at least one frame that is below a clinical relevant threshold of 5 mm [27, 28]. To benefit from

the fact that at least one frame with a good registration result is likely to be found, the frame selection could be performed manually by stepping through the frames and the associated registration results. The user can then quickly select the frame with the best registration result. Although this implementation still involves user input, the required user interaction is less than for fully manual registration.

For the remaining 15% of the cases or if higher accuracy level, for example, 3 mm [29], is required, small manual adjustment may be necessary in these cases.

**4.4. Temporal Filtering.** Temporal filtering based on the Markov chain could reduce the error significantly. Particularly high errors for frames at the beginning or at the end of the contrast injection were reduced. If a registration is not desired for an arbitrary frame but rather for a frame determined by the physician, for example, based on the breathing phase, temporal filtering should be applied.

**4.5. Runtime Performance.** It is notable that the runtime for  $\rho_{\text{CADE}}$  was considerably higher compared to, for example,  $\rho_{\text{shad}}^{\text{thr}}$ . This is because the contrast agent distribution needs to be updated in each iteration before projecting it. Moreover, this part of the implementation has not yet been optimized. Although the fast runtime of, for example,  $\rho_{\text{shad}}^{\text{thr}}$  will remain out of reach, we are confident that the runtime for  $\rho_{\text{CADE}}$  could be reduced to a clinically acceptable level.

**4.6. Impact of Clinical Issues.** In general, the performance for initial registration was better, as the amount of contrast agent was higher and fewer catheter artifacts were present. We found by visual inspection that also the amount of breathing motion had an impact on the registration accuracy as it caused motion artifacts in the DSA computation. As a consequence, if the patient was not anesthetized, acquisition of the CA injection should be performed under breath-hold or shallow breathing. For cases with intubation, also jet ventilation [30] or a small period of apnea could be applied to reduce breathing motion. Since our data was not acquired using a dedicated DSA program, the different brightness levels within a sequence changed. This interfered with the subtraction image computation. Although intuitively appealing, our data did not allow us to make a statement if the registration accuracy depends on the part of the LA in which CA was injected.

For initial contrast injections, an average accuracy of  $4.6 \pm 4.0$  mm and a median error of 3.7 mm could be achieved using  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  together with the best-frame selection approach. This error is in the range of the interuser variability of  $3.2 \pm 2.3$  mm. The faster registration method using  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$  reached an accuracy of  $4.8 \pm 4.6$  mm. These numbers are also similar to the accuracy reported when performing manual registration based on segmentations of the CS, the whole heart, and the spine [7] or a segmentation of the CS when 3D and 2D data are in the same breathing and cardiac phase [31] which was achieved by ventilation and rapid pacing of anesthetized patients. However, our method does not require structures other than the LA to be segmented and requires no anesthesia. Compared to a registration based on the CS alone [32], the error was reduced by over 50%.

The average accuracy for all frames after applying temporal filtering is 5.7 mm for  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  and 6.1 mm for  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$ . Though the results of  $\rho_{\text{CADE}} + \rho_{\text{edge}}$  are close to the 5 mm threshold [27, 28], it remains open if this is sufficient for

a clinical application, but we believe that these results should at least provide users with an acceptable initial estimate for further manual adjustments.

**4.7. Future Work.** By now, the registration method was evaluated only for left atria. In a next step, the performance of this approach to contrast-based registration of the right atrium should be evaluated. A registration based on the right atrium would also provide a registration for the LA which is available for transseptal puncture or one of the ventricles could be assessed. Due to the lack of pulmonary veins and arteries, the ventricles have a more simple anatomical structure but are subject to stronger cardiac motion. How strongly this affects the registration accuracy is open.

## 5. Conclusions

Compared to the approach by Zhao et al. [17], the use of special weights for different heart regions is not needed for any of the proposed approaches. In addition, for  $\rho_{\text{shad}}^{\text{thr}} + \rho_{\text{edge}}$ , a time-consuming DRR generation can be avoided. As a result, a registration approach based on a combination of shadow and edge features can be computed fast. If sufficient computational power was available, the novel CADE-based measure, which takes consistency into account, should be used as it improves results significantly, especially when very small amounts of contrast agent are injected.

## Disclosure

The concepts and information presented in this paper are based on research and are not commercially available.

## Conflict of Interests

Andreas Maier and Joachim Hornegger have no conflict of interests. Matthias Hoffmann, Christopher Kowalewski, and Klaus Kurzdin are supported by Siemens Healthcare GmbH, Forchheim, Germany. Norbert Strobel is an employee of Siemens Healthcare GmbH, Forchheim, Germany.

## Acknowledgments

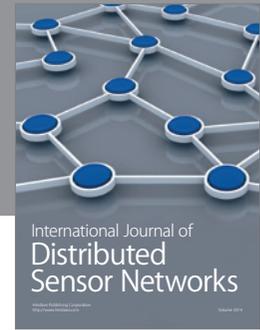
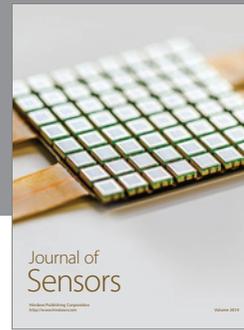
This work was supported by the German Federal Ministry of Education and Research (BMBF) in the context of the initiative Spitzencluster Medical Valley-Europäische Metropolregion Nürnberg, Project Grants nos. 13X1012A and 13EX1012E, respectively. Additional funding was provided by Siemens Healthcare GmbH. Andreas Maier and Joachim Hornegger gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative.

## References

- [1] S. De Buck, F. Maes, J. Ector et al., "An augmented reality system for patient-specific guidance of cardiac catheter ablation

- procedures,” *IEEE Transactions on Medical Imaging*, vol. 24, no. 11, pp. 1512–1524, 2005.
- [2] J. Sra, G. Narayan, D. Krum et al., “Computed tomography-fluoroscopy image integration-guided catheter ablation of atrial fibrillation,” *Journal of Cardiovascular Electrophysiology*, vol. 18, no. 4, pp. 409–414, 2007.
  - [3] J. H. Li, M. Haim, B. Movassaghi et al., “Segmentation and registration of three-dimensional rotational angiogram on live fluoroscopy to guide atrial fibrillation ablation: a new online imaging tool,” *Heart Rhythm*, vol. 6, no. 2, pp. 231–237, 2009.
  - [4] J. Sra, D. Krum, A. Malloy et al., “Registration of three-dimensional left atrial computed tomographic images with projection images obtained using fluoroscopy,” *Circulation*, vol. 112, no. 24, pp. 3763–3768, 2005.
  - [5] A. Brost, F. Bourier, L. Yatziv et al., “First steps towards initial registration for electrophysiology procedures,” in *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling*, K. H. Wong and D. R. Holmes III, Eds., vol. 7964 of *Proceedings of SPIE*, Lake Buena Vista, Fla, USA, February 2011.
  - [6] S. Knecht, H. Skali, M. D. O’Neill et al., “Computed tomography-fluoroscopy overlay evaluation during catheter ablation of left atrial arrhythmia,” *Europace*, vol. 10, no. 8, pp. 931–938, 2008.
  - [7] F. Bourier, T. Reents, S. Ammar-Busch et al., “Transseptal puncture guided by CT-derived 3D-augmented fluoroscopy,” *Journal of Cardiovascular Electrophysiology*, 2016.
  - [8] F. Bourier, D. Vukajlovic, A. Brost, J. Hornegger, N. Strobel, and K. Kurzydum, “Pulmonary vein isolation supported by MRI-derived 3D-augmented biplane fluoroscopy: A feasibility study and a quantitative analysis of the accuracy of the technique,” *Journal of Cardiovascular Electrophysiology*, vol. 24, no. 2, pp. 113–120, 2013.
  - [9] T. Feldman and W. G. Fisher, “Transseptal puncture,” in *Problem Oriented Approaches in Interventional Cardiology*, A. Colombo and G. Stankovic, Eds., chapter 17, pp. 203–218, CRC Press, New York, NY, USA, 2007.
  - [10] A. Guéziec, P. Kazanzides, B. Williamson, and R. H. Taylor, “Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot,” *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pp. 715–728, 1998.
  - [11] A. Hamadeh, S. Lavalley, and P. Cinquin, “Automated 3-dimensional computed tomographic and fluoroscopic image registration,” *Computer Aided Surgery*, vol. 3, no. 1, pp. 11–19, 1998.
  - [12] H. Livyatan, Z. Yaniv, and L. Joskowicz, “Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT,” *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1395–1406, 2003.
  - [13] B. L. Kaptein, E. R. Valstar, B. C. Stoel, P. M. Rozing, and J. H. C. Reiber, “A new model-based RSA method validated using CAD models and models from reversed engineering,” *Journal of Biomechanics*, vol. 36, no. 6, pp. 873–882, 2003.
  - [14] T. Benseghir, G. Malandain, and R. Vaillant, “Iterative closest curve: a framework for curvilinear structure registration application to 2D/3D coronary arteries registration,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*, vol. 8149 of *Lecture Notes in Computer Science*, pp. 179–186, Springer, Berlin, Germany, 2013.
  - [15] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, “A review of 3D/2D registration methods for image-guided interventions,” *Medical Image Analysis*, vol. 16, no. 3, pp. 642–661, 2012.
  - [16] D. Thivierge-Gaulin, C. Chou, A. Kiraly, C. Chef d’Hotel, N. Strobel, and F. Chieriet, “3D-2D registration based on mesh-derived image bisection,” in *Biomedical Image Registration*, pp. 70–78, Springer, 2012.
  - [17] X. Zhao, S. Miao, L. Du, and R. Liao, “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,” in *Proceedings of the 38th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’13)*, pp. 934–938, Vancouver, Canada, May 2013.
  - [18] H. U. Klemm, D. Steven, C. Johnsen et al., “Catheter motion during atrial ablation due to the beating heart and respiration: impact on accuracy and spatial referencing in three-dimensional mapping,” *Heart Rhythm*, vol. 4, no. 5, pp. 587–592, 2007.
  - [19] S. Kaeppler, A. Brost, M. Koch et al., “Motion estimation model for cardiac and respiratory motion compensation,” in *Information Processing in Computer-Assisted Interventions*, P. J. Abolmaesumi, Ed., vol. 7330 of *Lecture Notes in Computer Science*, pp. 94–103, Springer, Berlin, Germany, 2012.
  - [20] M. Hoffmann, F. Bourier, N. Strobel, and J. Hornegger, “Structure-enhanced visualization for manual registration in fluoroscopy,” in *Bildverarbeitung für die Medizin 2013: Algorithmen—Systeme—Anwendungen. Proceedings des Workshops vom 3. bis 5. März 2013 in Heidelberg*, Informatik Aktuell, pp. 241–246, Springer, Berlin, Germany, 2013.
  - [21] M. Hoffmann, C. Kowalewski, A. Maier, K. Kurzydum, N. Strobel, and J. Hornegger, “3-D/2-D registration of cardiac structures by 3-D contrast agent distribution estimation,” <http://arxiv.org/abs/1601.06062>.
  - [22] M. Hoffmann, N. Strobel, J. Hornegger, and A. Maier, “Contrast-based registration of left atria to fluoroscopic image sequences by temporal markov filtering and motion regularization,” in *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI ’16)*, 2016, <http://www5.cs.fau.de/Forschung/Publikationen/2016/Hoffmann16-CRO.pdf>.
  - [23] M. Hoffmann, S. Müller, K. Kurzydum, N. Strobel, and J. Hornegger, “Robust identification of contrasted frames in fluoroscopic images,” in *Bildverarbeitung für die Medizin 2015*, H. Handels, T. M. Deserno, H. P. Meinzer, and T. Tolxdorff, Eds., Informatik Aktuell, pp. 23–28, Springer, Berlin, Germany, 2015.
  - [24] P. Fischer, T. Pohl, and J. Hornegger, “Real-time respiratory signal extraction from X-ray sequences using incremental manifold learning,” in *Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI ’14)*, IEEE, Ed., pp. 915–918, Beijing, China, April-May 2014.
  - [25] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.
  - [26] A. Brost, R. Liao, N. Strobel, and J. Hornegger, “Respiratory motion compensation by model-based catheter tracking during EP procedures,” *Medical Image Analysis*, vol. 14, no. 5, pp. 695–706, 2010.
  - [27] A. P. King, R. Boubertakh, K. S. Rhode et al., “A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures,” *Medical Image Analysis*, vol. 13, no. 3, pp. 419–431, 2009.
  - [28] R. J. Housden, M. Basra, Y. Ma et al., “Three-modality registration for guidance of minimally invasive cardiac interventions,” in *Functional Imaging and Modeling of the Heart*, vol. 7945 of *Lecture Notes in Computer Science*, pp. 158–165, Springer, Berlin, Germany, 2013.

- [29] F. Bourier, R. Fahrig, P. Wang et al., "Accuracy assessment of catheter guidance technology in electrophysiology procedures," *Journal of Cardiovascular Electrophysiology*, vol. 25, no. 1, pp. 74–83, 2014.
- [30] J. S. Goode Jr., R. L. Taylor, C. W. Buffington, M. M. Klain, and D. Schwartzman, "High-frequency jet ventilation: utility in posterior left atrial catheter ablation," *Heart Rhythm*, vol. 3, no. 1, pp. 13–19, 2006.
- [31] M. Hoffmann, N. Strobel, and A. Maier, "Registration of atrium models to C-arm X-ray images based on devices inside the coronary sinus and the esophagus," in *Bildverarbeitung für die Medizin 2016*, 2016, <http://www5.cs.fau.de/Forschung/Publicationen/2016/Hoffmann16-ROA.pdf>.
- [32] F. Bourier, A. Brost, L. Yatziv, J. Hornegger, N. Strobel, and K. Kurzidim, "Coronary sinus extraction for multimodality registration to guide transseptal puncture," in *Proceedings of the 8th Interventional MRI Symposium-Book of Abstracts*, T. Kahn, F. A. Jolesz, and J. S. Lewin, Eds., pp. 311–313, Leipzig, Germany, September 2010.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

