

Robust Multi-View 2-D/3-D Registration Using Point-To-Plane Correspondence Model

Roman Schaffert, Jian Wang, Peter Fischer, Andreas Maier, and Anja Borsdorf

Abstract—In minimally invasive procedures, the clinician relies on image guidance to observe and navigate the operation site. In order to show structures which are not visible in the live X-ray images, such as vessels or planning annotations, X-ray images can be augmented with pre-operatively acquired images. Accurate image alignment is needed and can be provided by 2-D/3-D registration. In this paper, a multi-view registration method based on the point-to-plane correspondence model is proposed. The correspondence model is extended to be independent of the used camera coordinates and different multi-view registration schemes are introduced and compared. Evaluation is performed for a wide range of clinically relevant registration scenarios. We show for different applications that registration using correspondences from both views simultaneously provides accurate and robust registration, while the performance of the other schemes varies considerably. Our method also outperforms the state-of-the-art method for cerebral angiography registration, achieving a capture range of 18 mm and an accuracy of 0.22 ± 0.07 mm. Furthermore, investigations on the minimum angle between the views are performed in order to provide accurate and robust registration, while minimizing the obstruction to the clinical workflow. We show that small angles around 30° are sufficient to provide reliable registration results.

Index Terms—rigid 2-D/3-D registration, multi-view, point-to-plane correspondence model, spine registration

I. INTRODUCTION

OVERLAYS of pre-operative 3-D images on intra-operative 2-D fluoroscopy images are often used in interventional radiology to provide guidance to the physician. Typical use cases are displaying of anatomical structures which are not visible in the fluoroscopy images (e.g. blood vessels [1] or soft tissue such as the heart [2]) or markers defined before the intervention (e.g. bifurcation points which must not be obstructed or screw placement information for spine surgery). Furthermore, depth information can be visualized in an intuitive manner [3]. An accurate alignment of the images is crucial for the overlay. In order to achieve a high accuracy, 2-D/3-D image registration methods are employed to align the images with an estimated transformation. The topic of 2-D/3-D registration has been widely studied and a large number of methods were proposed. For a comprehensive overview of 2-D/3-D registration, we refer the reader to the overview by Markelj et al. [4] and

R. Schaffert and A. Maier are with the Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany. e-mail: roman.schaffert@fau.de.

J. Wang, P. Fischer and A. Borsdorf are with Siemens Healthineers AG, 91301 Forchheim, Germany.

A. Maier is also with the Graduate School in Advanced Optical Technologies (SAOT), 91058 Erlangen, Germany

for a more general overview of registration methods to Liao et al. [5].

This paper focuses on intrinsic rigid 2-D/3-D registration. This type of methods can be further categorized based on the type of image information used for the registration and how the dimensional correspondence between the 3-D and the 2-D images is established [4], [5]. More specifically, Markelj et al. [4] suggest the division into feature-based, intensity-based and gradient-based methods. While feature-based registration relies on extracted features, intensity-based and gradient-based registration is performed by defining and optimizing a similarity measure based on the intensities or gradients of the images. To obtain the dimensional correspondence of the input images, either projection of the 3-D image, back-projection of the 2-D image or a rough reconstruction in case of multiple 2-D images can be used [4]. Gradient-based approaches were demonstrated to have a high accuracy and enable X-ray to MR registration [6]–[8]. However, the robustness of gradient-based methods tends to be low [4] and a feature-based alignment is often introduced as an initial step to improve the robustness [8], [9]. The most common approach to 2-D/3-D registration is the combination of intensity-based registration and the projection strategy [4]. Here, digitally reconstructed radiographs (DRRs) are generated from the 3-D image and the similarity to the 2-D image is optimized, e.g. by Kubias et al. [10].

More recently, methods have been proposed which combine the matching of local structures with maximizing a similarity measure defined for the images. Mitrović et al. [11] propose a method for registration of cerebral angiography data. They extract centerline points together with the direction and radius of the vessel. The centerline points are projected onto the 2-D image and a similarity is computed by matching vessel orientations and the gradient orientation in neighborhoods around the projected point in the 2-D image. A high robustness of the method is demonstrated. Špiclin et al. [12] follow a more general approach to incorporate the neighborhood structure into the similarity estimation. Here, covariance matrices are estimated for both images. The covariances from the 2-D image are back-projected and compared to the 3-D covariances. To enable fast registration, only regions with high gradients in 3-D are considered. This method achieves a high robustness and is not limited to a specific use case.

The effect of the angle between the two views on the registration accuracy is investigated by Uneri et al. [13]. The authors demonstrate that relatively small angles around 20° are enough to achieve an optimal accuracy, which is not further improved for larger angles. However, a single initial position is used and robustness is not evaluated.

For multi-view registration, Kaiser et al. [14] propose a scheme which relies on a centerline of the registered structure. It allows to iteratively align the appearance of the object in one view, while preserving the independent in-plane parameters in the other view. The authors demonstrate faster and more accurate registration of the transesophageal echocardiography (TEE) probe compared to iterative in-plane registration in case of non-orthogonal views.

A general approach to increase the robustness is to apply a global optimization strategy when optimizing the similarity measure. As this requires a high number of evaluations of the similarity measure, different strategies for fast DRR generation have been investigated. Duménil et al. [15] propose a decomposition of the motion allowing changes in the DRR caused by translational and in-plane rotational motions to be approximated as 2-D transformations. This enables an efficient grid-search approach. Otake et al. [16] propose to use a highly parallel approach to enable fast DRR generation in combination with a multistart strategy and the CMA-ES [17] as a local optimizer. While the authors report an impressive capture range, the search space is adjusted to the expected range of motion for the individual motion parameters in order to limit the computational cost, allowing large motion mainly for translations along the spine and out-of-plane translations.

Schmid et al. [18] use 3-D contour generator points [19], i.e. points representing a contour in the projection of the volume onto the image plane, as a basis for the registration. Correspondences for these points in the 2-D X-ray image are established by performing local patch matching between a DRR and the X-ray image. A motion is computed based on forces defined for the correspondences. The physics-based motion estimation enables inter-object constraints between multiple registered structures. To achieve a high accuracy, a gradient-based similarity measure is optimized in a second step. Similarly, Wang et al. [20] propose a method where correspondences for contour generator points in the X-ray image are established. Considering the fact that displacement along a contour is hardly observable, the authors propose a point-to-plane correspondence (PPC) model that takes the displacements perpendicular to contours as input and efficiently constrains both in-plane and out-of-plane 3D transformations. A high robustness and accuracy is observed for the method and no refinement step is needed. Furthermore, a parallel implementation of the method on the GPU [21] enables real-time tracking of the patient. However, the PPC model is derived in the camera coordinate system and cannot be directly used for multi-view registration. While a high 3-D accuracy using single-view registration can be achieved [20], single-view registration is generally considered as an ill-posed problem [12] and is especially challenging, e.g. in case of limited field of view or small structures.

In this work, we propose a view-independent extension of the PPC model [20]. Considering the fact that distances are invariant to rigid transformations, we reformulate the PPC model to allow motion estimation directly in a chosen coordinate system related to the camera coordinates by a rigid transformation [22]. We propose to use the view-independent PPC model to enable multi-view registration using correspon-

dences from all views [22]. Alternatively, the registration can be obtained by performing a motion estimation step for each view and iterate over the views. Another strategy is to perform single-view registration for all views, select the most promising results and refine the out-of-plane parameters using the other view(s) [22]. We compare our proposed method to these strategies. Regularized motion estimation, which has shown to lead to increased robustness for single-view registration [23], is further investigated as a means to increase robustness. We evaluate the different approaches on a diverse set of animal and clinical data. To minimize the obstruction of the clinical workflow, a small angle between the views is desirable. We investigate the effect of the angular distance of the used views in order to establish the minimal angle for accurate and robust registration [22].

II. RELATED WORK

In this work, we extend the registration method described by Wang et al. [20]. In the following section, we describe the PPC model, as well as the PPC-based registration framework [20].

A. Point-To-Plane Correspondence Model

The PPC model is introduced for tracking of differential motion [24] and was shown to achieve a high accuracy as well as robustness for single-view 2-D/3-D registration [20]. The idea behind the PPC model is to measure the local misalignment between a 2-D X-ray image I^{FL} and a 3-D volume V and to estimate a 3-D rigid motion to compensate for it. As contours of high-contrast structures, e.g. bones, are well distinguishable in I^{FL} , they are used as the basis for the misalignment estimation.

In the volume V , the surface of high-contrast structures is extracted and used as the basis for the registration. From the surface, contour generator [19] points $\{\mathbf{w}_i \in \mathbb{R}^3\}$ are selected, i.e. a set of points \mathbf{w}_i representing the apparent contour in the projection of V under an initial registration transformation $\mathbf{T}^{init} \in \mathbb{R}^{4 \times 4}$ and the given projection geometry. The PPC model is defined in the camera coordinate system \mathcal{C} .

The points $\{\mathbf{w}_i\}$ are projected onto a normalized image plane parallel to the detector and with a distance of 1 to the source, yielding a set of projected points $\{\mathbf{p}_i \in \mathbb{R}^3\}$. The local misalignment is measured for all $\{\mathbf{p}_i\}$ by finding correspondences $\{\mathbf{p}'_i \in \mathbb{R}^3\}$ in I^{FL} . Following the assumption that displacements along the contour cannot be detected locally, only motion perpendicular to the contour, i.e. along the contour normal, is considered. For a contour point \mathbf{p}_i , the normal corresponds to the projection of the intensity gradient \mathbf{g}_i at $V(\mathbf{w}_i)$ onto the image plane. The correspondence \mathbf{p}'_i is searched along this direction only.

After the correspondences are established, the motion of \mathbf{w}_i is still not known in two directions: along the back-projection ray of \mathbf{p}'_i and along the contour. As the contour is perpendicular to the gradient \mathbf{g}_i as well as the viewing ray through \mathbf{w}_i , the contour direction can be expressed as $\mathbf{w}_i \times \mathbf{g}_i$. Therefore, we can only say that the point is located on a plane Π_i with a normal $\mathbf{n}_i \in \mathbb{R}^3$ spanned by these two directions, i.e.

$$\mathbf{n}_i = (\mathbf{w}_i \times \mathbf{g}_i) \times \mathbf{p}'_i . \quad (1)$$

To achieve local alignment, the point \mathbf{w}_i has to be located on the corresponding plane Π_i , i.e.

$$\text{dist}(\Pi_i, \mathbf{w}_i + d\mathbf{w}_i) = 0 , \quad (2)$$

where $\text{dist}(\cdot, \cdot)$ indicates point-to-plane distance and $d\mathbf{w}_i$ is the displacement of \mathbf{w}_i . As Π_i contains the back-projection of \mathbf{p}' and passes through the origin, Eq. (2) can be expressed as

$$\mathbf{n}_i^\top (\mathbf{w}_i + d\mathbf{w}_i) = 0 . \quad (3)$$

In case of rigid registration, all points are displaced according to a global transformation consisting of a translation and a rotation of V . The global motion is denoted as $\delta\mathbf{v} = (\delta\omega^\top, \delta\nu^\top)^\top$, where $\delta\omega \in \mathbb{R}^3$ is the rotational motion component in the axis-angle representation and $\delta\nu \in \mathbb{R}^3$ is the translational component.

The PPC model was originally developed for the tracking case and a small amount of motion is assumed [24]. Using the Rodrigues rotation formula and the small angle assumption, the displacements of the individual points can be expressed linearly in the motion vector as

$$d\mathbf{w}_i = \delta\omega \times \mathbf{w}_i + \delta\nu . \quad (4)$$

Using this simplified relation, Eq. (3) is reformulated as

$$((\mathbf{n}_i \times \mathbf{w}_i)^\top, -\mathbf{n}_i^\top) \delta\mathbf{v} = \mathbf{n}_i^\top \mathbf{w}_i . \quad (5)$$

To estimate the motion, the overall point-to-plane distance over all correspondences is minimized, i.e. $\mathbf{A}\delta\mathbf{v} = \mathbf{b}$ is solved. Every correspondence contributes one row to $\mathbf{A} \in \mathbb{R}^{N \times 6}$ as $\mathbf{a}_i^\top = ((\mathbf{n}_i \times \mathbf{w}_i)^\top, -\mathbf{n}_i^\top)$ and one element to $\mathbf{b} \in \mathbb{R}^N$ as $b_i = \mathbf{n}_i^\top \mathbf{w}_i$, where N is the number of correspondences.

B. Single-View Registration Method

In this section, we describe the registration method which uses the PPC model and can be used for initial single-view registration [20].

As a pre-processing step, a guided image filter [25] is applied to V in order to reduce the noise. Then, surface points are extracted from V using a 3-D Canny detector [26]. The actual registration is performed by iterating the following steps.

1) *Extraction of Apparent Contour Points:* The apparent contour points under the initial registration transformation T^{init} are selected using the perpendicularity condition [24]. For a point on the apparent contour, the surface normal (or intensity gradient \mathbf{g}_i) and the viewing ray through \mathbf{w}_i are perpendicular to each other. In practice, a perpendicularity threshold is applied, i.e. $\angle(\mathbf{g}_i, \mathbf{w}_i) \geq t_\theta$. The points fulfilling this criterion are further considered as the apparent contour point set $\{\mathbf{w}_i\}$.

2) *Correspondence Estimation:* After $\{\mathbf{w}_i\}$ is extracted, a correspondence for each \mathbf{w}_i in I^{FL} is established using patch matching. As a first step, V is uniformly divided into depth intervals and a gradient projection image is computed for every interval, leading to a set of depth layer images (DLs) $\{\nabla I_d^{\text{proj}}\}$, where d indicates the depth interval index [27]. For every projected point \mathbf{p}_i , a symmetric neighborhood with the

half size r_Ω around the point is used from the corresponding DL, denoted as $\Omega(\nabla I_d^{\text{proj}}, \mathbf{p})$. Patch matching is performed on candidate positions \mathbf{p}_i^c in a defined search range r_s both along the parallel and antiparallel direction of the image gradient $\nabla I_d^{\text{proj}}(\mathbf{p}_i)$ using the normalized gradient correlation (NGC) [28], i.e.

$$\mathbf{p}'_i = \underset{\mathbf{p}_i^c}{\operatorname{argmin}} \text{NGC}(\Omega(\nabla I_d^{\text{proj}}, \mathbf{p}_i), \Omega(\nabla I^{\text{FL}}, \mathbf{p}_i^c)) , \quad (6)$$

where $\Omega(\nabla I, \mathbf{x})$ is a patch from the gradient image ∇I around the position \mathbf{x} . Note that the image gradient $\nabla I_d^{\text{proj}}(\mathbf{p}_i)$ is used instead of \mathbf{g}_i , as it is obtained from an integration along depth and is more robust to noise [20]. Matches below a similarity threshold t_{NGC} are regarded as wrong and not considered for the motion estimation.

3) *Motion Estimation:* In order to estimate the motion, the optimization problem $\operatorname{argmin} \Sigma_i \mathbf{L}(\mathbf{a}_i \delta\mathbf{v} - b_i)$ is solved, where \mathbf{L} denotes the loss function. The normals \mathbf{n}_i are computed as

$$\mathbf{n}_i = (\mathbf{w}_i \times \nabla I_d^{\text{proj}}(\mathbf{p}_i)) \times \mathbf{p}'_i \quad (7)$$

instead of Eq. (1), as the estimation of $\nabla I_d^{\text{proj}}(\mathbf{p}_i)$ is more robust to noise and is also perpendicular to the contour [20]. To achieve a high robustness of the motion estimation to outliers, we follow Wang et al. [20] and use the maximum correntropy criterion for regression (MCCR) [29] as the loss function \mathbf{L} . The optimization problem is solved using the iteratively reweighted least squares (IRLS) scheme [30].

III. 2-D/3-D REGISTRATION USING VIEW-INDEPENDENT POINT-TO-PLANE CORRESPONDENCE MODEL

The PPC model and the registration framework proposed by Wang et al. [20] enable an accurate and robust single-view registration. However, single-view registration is challenging in some circumstances, e.g. with limited field-of-view. Therefore, we investigate multi-view registration using the PPC model. In the original formulation, the PPC model is defined in the camera coordinates of a view and correspondences of different views cannot be combined.

In this work, we reformulate the model in a coordinate system \mathcal{D} related to the camera coordinates \mathcal{C}_v of a view v by a rigid transformation in order to enable the simultaneous use of correspondences from different views. We propose to use correspondences from all views simultaneously and compare the proposed method to alternative multi-view registration schemes. Furthermore, we explore the effect of the angular distance between the views on the registration performance as well as the effect of regularized motion estimation. To enable meaningful regularization, we use the view-independent PPC model and move the origin to the center of the volume. This enables rotations without displacing the object and therefore no translations are needed to compensate this displacement.

A. View-Independent Point-To-Plane Correspondence Model

According to the PPC model, the motion is estimated by minimizing point-to-plane distances (see Eq. (2)). In order to obtain the extended, view-independent PPC model [22], we

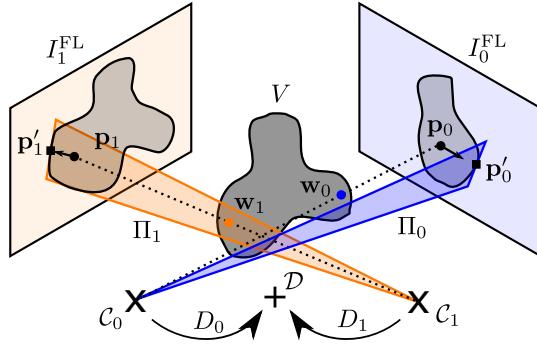


Fig. 1. Correspondences for two views. For the volume V , each view C_j and the corresponding image I_j^{FL} , one correspondence is shown (consisting of point w_j , projected point p_j , found correspondence p'_j and the corresponding plane Π_j). The views are related to the coordinate system \mathcal{D} by D_j .

make use of the fact that distances are invariant under rigid transformations. Therefore, Eq. (2) can be reformulated as:

$$\text{dist}(D_v(\Pi_{i,v}), D_v(w_{i,v}) + D_v(dw_{i,v})) = 0 , \quad (8)$$

where $D_v(\cdot)$ denotes a rigid transformation. For every view v , a different $D_v(\cdot)$ is used so that $D_v(\cdot)$ transforms correspondences related to the view C_v to a common coordinate system \mathcal{D} . Analogous to Eq. (3), Eq. (8) can be expressed as

$$D_v(n_{i,v})^\top (D_v(w_{i,v}) + D_v(dw_{i,v})) - D_v(n_{i,v})^\top D_v(\mathbf{0}) = 0 , \quad (9)$$

where $D_v(\mathbf{0})$ is the origin of the camera coordinate system C_v expressed in \mathcal{D} . See Fig. 1 for an illustration of the used coordinate systems and transformations. Note that the transformed plane $D_v(\Pi_{i,v})$ does not pass through the origin of \mathcal{D} and the distance to the origin is accounted for by $D_v(n_{i,v})^\top D_v(\mathbf{0})$. The displacement of the point $D_v(dw_{i,v})$ is directly expressed in \mathcal{D} depending on the motion estimated in \mathcal{D} . Analogous to Eq. (4), the displacement is expressed as

$$D_v(dw_{i,v}) = d\hat{w}_{i,v} = \delta\hat{\omega} \times D_v(w_{i,v}) + \delta\hat{\nu} , \quad (10)$$

where $\delta\hat{\mathbf{v}} = (\delta\hat{\omega}^\top, \delta\hat{\nu}^\top)^\top$ is the motion vector estimated in \mathcal{D} . Combining Eq. (9) and Eq. (10), we obtain

$$((D_v(n_{i,v}) \times D_v(w_{i,v})))^\top, -D_v(n_{i,v})^\top \delta\hat{\mathbf{v}} = D_v(n_{i,v})^\top (D_v(w_{i,v}) - D_v(\mathbf{0})) , \quad (11)$$

which can be used to estimate the motion directly in \mathcal{D} .

B. Multi-View Registration Variants

In this section, the different multi-view registration schemes are discussed. All methods use the view-independent PPC model in order to center the registered object to the origin.

1) *Registration Using Multiple Views (PPC-M):* The PPC-M method uses the view-independent PPC model in order to allow for motion estimation using correspondences from multiple views. All correspondences are combined into a single system of equations and a motion $\delta\hat{\mathbf{v}}$ is computed to align the object in all views simultaneously [22]. For an illustration of the setup, see Fig. 1.

2) *Registration Using View Selection (PPC-S):* Another approach to multi-view registration is to perform a single-view registration for each view first. Then, the view which leads to the best registration is selected. The corresponding registration is then refined by iteratively performing one registration iteration for each view, restricting the motion in depth [22].

3) *Alternating Registration (PPC-A):* The last considered variant is to alternate between views on every iteration. It is not possible to estimate only the in-plane parameters for every iteration, as one rotational component is out-of-plane for both views. Instead, only motion in depth, i.e. along the viewing direction, is not estimated.

For all methods, the origin of \mathcal{D} is set to the center of the registered structure.

C. Regularized Motion Estimation

The L2-regularized least squares solution is computed as

$$\delta\hat{\mathbf{v}} = \underset{\delta\hat{\mathbf{v}}'}{\text{argmin}} \left(\frac{1}{N} \|\mathbf{A}\delta\hat{\mathbf{v}}' - \mathbf{b}\|_2^2 + \lambda \|\delta\hat{\mathbf{v}}'\|_2^2 \right) , \quad (12)$$

where λ is the weight of the regularizer. For meaningful regularization, the rotational and translational parameters have to be made independent of each other. This is achieved by centering the coordinate system, in which the motion is estimated, to the center of the registered structure [22].

IV. IMPLEMENTATION DETAILS

In this section, essential implementation details of the used registration method are discussed.

A. Correspondence Weighting

In order to increase the robustness of the motion estimation, the found correspondences are weighted by $s_{\text{NGC},i,v}$ according to the respective image patch similarity, i.e.

$$s_{\text{NGC},i,v} = \text{NGC}(\Omega(\nabla I_d^{\text{proj}}, \mathbf{p}_{i,v}), \Omega(\nabla I^{\text{FL}}, \mathbf{p}'_{i,v})) . \quad (13)$$

The weights are incorporated into the PPC model by using $\mathbf{A}_s = \text{diag}(\mathbf{s}) \cdot \mathbf{A}$ and $\mathbf{b}_s = \text{diag}(\mathbf{s}) \cdot \mathbf{b}$ in the motion estimation, where \mathbf{s} is a vector containing weights $s_{\text{NGC},i,v}$ for all correspondences and $\text{diag}(\cdot)$ is a diagonal matrix. To compute the IRLS weights, equations without the applied $s_{\text{NGC},i,v}$ are considered to ensure that the residual errors represent actual point-to-plane distances. The weights $s_{\text{NGC},i,v}$ are additionally multiplied to the IRLS weights in order to retain higher weights for correspondences with higher similarity.

B. Regularized Motion Estimation

Regularized motion estimation using the PPC model [23] is considered in order to further increase the robustness of the registration. To regularize the rotational and translational components equally, the rotational component is scaled to the same range as the translational component. The scaling factor s_ω is computed in a way to ensure that the maximum possible influence of a vector $\delta\omega$ with $\|\delta\omega\|_2 = 1$ on the residual averaged over all correspondences is equal to the maximum

influence of a translation vector $\delta\nu$ with $\|\delta\nu\|_2 = 1$. It is computed as

$$s_\omega = \frac{\sum_v^M \sum_i^{N_v} \|s_{i,v} \mathbf{n}_{i,v}\|_2}{\sum_v^M \sum_i^{N_v} \|s_{i,v} (\mathbf{n}_{i,v} \times \mathbf{w}_{i,v})\|_2} . \quad (14)$$

The scaling is applied by multiplying the components of \mathbf{A} corresponding to the rotation by s_ω and reversed by multiplying s_ω to the scaled rotation vector $\delta\tilde{\omega}$. The regularized motion estimation is combined with the motion estimation using MCCR.

C. Used Coordinate System

A common coordinate system \mathcal{D} is needed for the registration using PPC-M. To define \mathcal{D} , we start with the coordinate system \mathcal{C}_0 and shift the origin to the center of the registered structure. If a registration iteration is performed for a single view, the current view is treated as the only view \mathcal{C}_0 , i.e. the motion estimation is performed in the camera coordinates with the origin shifted to the object center.

D. Iterative Registration

The registration is performed iteratively. To ensure that correspondences are found for large misalignment, a multi-resolution scheme is utilized [20] to effectively enlarge the search range and the patch size. For single-view registration and the PPC-S method, the registration is performed on the lowest resolution level without allowing motion in depth first to improve the robustness of the method.

Following Wang et al. [20], we compute a quality measure Q_k , where k is the iteration index. This measure is used to detect convergence. Additionally, the result of every resolution level is selected from the iteration with the lowest value for Q_k . The quality measure is defined as

$$Q_k = \frac{\bar{r}_k}{\bar{s}_k + \eta} , \quad (15)$$

where \bar{r}_k is the mean residual over the used correspondences, \bar{s}_k the mean similarity and η a constant to avoid division by zero. To enable the computation for multi-view registration, these measures are defined as follows.

The mean image similarity is computed as

$$\bar{s}_k = \frac{1}{M} \sum_{v=1}^M \frac{1}{N_v} \sum_{i=1}^{N_v} \text{NGC}(\Omega(\nabla I_{d,v}^{\text{proj}}, \mathbf{p}_{i,v}), \Omega(\nabla I_v^{\text{FL}}, \mathbf{p}_{i,v})) , \quad (16)$$

where M is the number of views. The mean residual \bar{r}_k is defined differently depending on the used method. As all correspondences are used simultaneous for PPC-M, it is the resulting mean residual of the motion estimation. For PPC-A and the refinement step in PPC-S, the motion is computed sequentially for different views. In this case, the overall mean residual is computed as the mean residual over all views.

For each resolution level, the final result is selected as the iteration k leading to the lowest Q_k [20]. The registration is performed until the convergence criterion is met or the maximum number of iterations is reached. The convergence criterion is fulfilled if the difference between the quality

measure of two iterations and the difference between the old and new minimum quality measure are sufficiently small [20]. Additionally, convergence is only assumed if the value of \bar{s}_k exceeds a minimum threshold $t_c = 0.15$.

For each resolution level, the results are rejected if $\bar{s}_{k'} < t_{\bar{s}}$, where k' is the iteration leading to the lowest Q_k and $t_{\bar{s}}$ is a success threshold.

V. EXPERIMENTS AND RESULTS

A. Experimental Setup

A set of experiments are designed to compare the multi-view registration methods between each other as well as to other methods.

1) Evaluation Methodology and Metrics: We follow the standardized evaluation methodology proposed by van der Kraats et al. [31]. To evaluate the registration accuracy, we mainly rely on the mean target registration error (mTRE), which is defined as the distance of target points under the ground truth (GT) registration and the estimated registration. The mean re-projection distance (mRPD) is defined as the distance between the target points under the GT registration and the back-projection rays of the target points under the estimated registration. To assess the robustness of a registration method, success rate (SR) and capture range (CR) are used. The SR is defined as the number of registrations below a certain success criterion. In this work, we use a criterion of $\text{mTRE} \leq 2 \text{ mm}$. The CR is the lower bound of the first 1 mm interval for which less than 95 % of registrations are successful [11], [31].

For the gold standard (GS) vertebra dataset [32] (see Sec. V-A2d), we additionally report results following the evaluation methodology proposed by Tomažević et al. [7] to improve comparability. Here, the root mean square target registration error (Rms(TRE)) is reported as well as the root mean square angular error (Rms(γ)). The SR is computed using the maximum error of the target points instead of the mean value. We denote this variant of the SR as SR_m . Instead of a CR, the SR_m is given for different initial error intervals.

2) Datasets: The datasets that are used in the evaluation are summarized as follows:

a) GS Cerebral Angiography Data: It is a publicly available cerebral angiography dataset introduced by Mitrović et al. [11]. It consists of ten 3-D digital subtraction angiogram (DSA) images of the cerebral vessel tree of different patients. For every volume, a set of 2-D images of the anterior-posterior (AP) and in lateral (LAT) direction views are provided. Both 2-D DSA images as well as native images (images with contrast agent but without the background subtracted) are available. We focus on the native images, as they contain additional structures making the registration more challenging and therefore are well suited to evaluate the robustness of a registration method. For each 3-D image, 400 uniformly distributed random start positions are available [11]. Those are computed using the standardized methodology proposed by van de Kraats et al. [31] and are in the range of $[0, 20] \text{ mm}$, with translations in the range of $[-20, 20] \text{ mm}$ and rotations in the range of $[-10, 10] \text{ degrees}$.

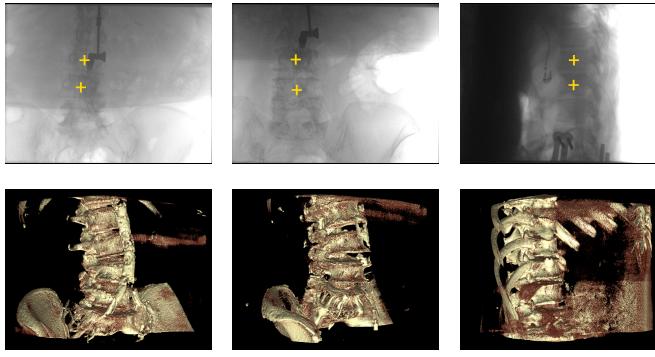


Fig. 2. Example of acquisitions in the spine dataset. The upper row shows samples of 2-D images I^{FL} used for registration, the lower row shows renderings of the corresponding 3-D images V . Vertebrae used in single-vertebra registration evaluation are marked with a yellow cross, best viewed in color.

Together with this dataset, Mitrović et al. [11] propose a registration method, denoted as MGP, which is specifically designed for vessel registration and is the state-of-the-art method for this dataset regarding robustness. By combining it with the backprojection gradient-based (BGB) registration method [7] (see Sec. V-A2d), the highest accuracy among all compared methods is achieved. The combined method is denoted as MGP + BGB. Recently, MGP was compared to different state-of-the-art registration methods for cerebral angiography data and again showed the highest robustness [33]. To evaluate the performance of our method, we compare it to MGP + BGB for this dataset.

b) Spine Data: The spine dataset consists of six clinical C-arm CT acquisitions of the lumbar and lumbar-thoracic regions of the spine. We register the acquired 2-D projection images to the reconstructed 3-D volumes. As the projection images are the basis for the reconstruction, the GT registration is known from the calibration of the system with an accuracy of ≤ 0.16 mm for the projection error in the iso-center. Also, we have many different views available and can perform registration from different angulations. Poisson-distributed noise is added to the 2-D images and the intensities are normalized to better simulate fluoroscopy images. Examples of the used images are shown in Fig. 2. The 2-D images have a resolution of 616×480 pixels with a pixel spacing of 0.608 mm. The 3-D images have slices with a resolution of 512×512 pixels and contain about 390 slices. The voxel spacing is 0.49 mm in all directions. For each volume, we define a volume of interest (VOI) containing the spine while excluding other structures like the pelvis or the ribs as much as possible.

We furthermore evaluate the registration performance for single vertebrae. This is important in practice as the spine may be deformed and the rigid motion assumption is only valid for individual vertebrae, which have to be registered individually. Furthermore, this enables us to observe the effect that the size of the structures has on the registration. To obtain the individual vertebrae, we split the VOI of the spine into smaller VOIs for individual vertebrae. Overlap of the VOIs is present as the border of the vertebrae cannot be divided cleanly by a bounding box border. For each acquisition, two

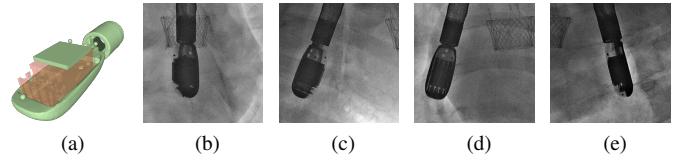


Fig. 3. TEE data used for registration. The mesh used in the registration (green) as well as the removed parts (red) are shown in (a), best viewed in color. Cropped fluoroscopy images (corresponding to 512×512 pixels in the original images) are shown in (b) - (e). In multi-view registration, images (b) and (c) as well as (d) and (e) are used as image pairs.

vertebrae are considered. For the four acquisitions of the lumbar region, L3 and L4 vertebrae are considered, while for lumbar-thoracic acquisitions, L1 and L2 are considered for one of the acquisitions and T11 and T12 for the other. We put an emphasis on single-vertebra registration, where single-view registration is often not sufficient and a second view is needed to reliably achieve an accurate registration.

For each spine as well as single vertebra used in registration, we generate 300 start positions which vary from the ground truth position with translations in the range of [-30, 30] mm and rotations in the range of [-15, 15] degrees, yielding uniformly distributed initial mTRE in the range of [0, 30] mm. Target points are uniformly distributed in the VOIs representing the registered structures.

c) Transesophageal Echocardiography (TEE) Probe: The geometry of the TEE probe is fixed and therefore a single 3-D image can be used for registration on different patients. We use a pre-processed high-resolution CT image of the probe with a size of $160 \times 144 \times 572$ voxels and a spacial resolution of 0.1 mm in all directions. We use a dataset similar to Kaiser et al. [14]. Fluoroscopic images with a size of 1024×1024 and a spacial resolution of 0.184 mm are used as I^{FL} . Manual annotation is available and is used as the ground truth.

A high-quality mesh representation of the probe is generated by Kaiser et al. [34]. We use the mesh vertices of their mesh instead of Canny-based surface extraction and substitute the gradients $\{g_i\}$ with the vertex normals. To obtain a sufficient number of contour points, sub-division of the mesh faces is performed first. We observe that the TEE probe contains repetitive structures which can adversely affect the registration robustness and remove these structures from the mesh (see Fig. 3 (a)). As the same mesh can be used for all registrations of the TEE probe, this needs to be performed only once and would not introduce a manual step to the registration in clinical practice. We perform experiments on porcine data. We use a total of 5 image pairs with 90° between the images. For sample images, see Fig. 3 (b) - (e).

While methods specially adjusted to the registration of the TEE probe achieve a high robustness and accuracy [14], [35], our focus here is to compare the different multi-view registration schemes and optimizing the registration for the probe is planned for future work.

d) GS Vertebra Dataset: We furthermore perform an evaluation on the GS single-vertebra dataset published by Tomažević et al. [32]. Here, a phantom is created by placing

a cadaver spine into a tube and applying markers to the tube which are visible in the 2-D X-ray images as well as the reconstructed volume. VOIs around the vertebrae are defined for registration. As the markers are outside of the spine, no markers are present in the VOIs. The dataset consists of 18 2-D images acquired from different angles and five volumes containing the VOIs for the individual vertebrae. For each vertebra, 450 start positions in the range [-20, 20] mm and [-51.7, 51.7] degrees are defined as well as the combination of 2-D views to be used for the registration of each start position. The backprojection gradient-based (BGB) registration method, proposed by Tomažević et al. [7], is evaluated on this dataset. Due to the high accuracy it achieves, BGB is used as a refinement step for other methods, e.g. by Mitrović et al. [11] and Špiclin et al. [12]. As the volumes in this dataset have an anisotropic resolution, we resample them to an isometric voxel spacing of 0.75 mm.

B. Used Parameters

The PPC-based registration methods have a number of parameters. For many parameters, default values can be defined which work well for all datasets. However, some parameters need to be selected depending on the registered data. The values of the relevant parameters are summarized in Tab. I. The adjusted parameters, t_θ , number of DLs, and minimum used resolution level RL_{min} , are all concerned with the size and shape of the registered structures.

The number of DLs is fixed to five following Wang et al. [20] except for the TEE probe. As the probe is a very small structure with a simple shape and few overlapping structures, a single DL is used. The value of t_θ is adjusted manually to obtain contours without large holes while preserving thin contours. In Fig. 4, the selection of t_θ is depicted for the angiography dataset as an example. We select a low value of 80° for the TEE-probe. This is done due to the fact that mesh vertices are used as contour points. As these points are sparser compared to the Canny points, a larger angle is needed to obtain a sufficient number of points.

The minimum resolution level L_m is selected as an image resolution where the registered object is still well visible. We observe that selecting a resolution level which is too low leads to a failure of the registration method on the lowest resolution level, even for small initial misalignment. Therefore, we verify the selected resolution level by performing a few registrations with small initial misalignment.

C. Evaluation Discussion

In the following section, the results of our experiments are discussed.

1) Selection of Registration Success Criterion: As a multi-resolution scheme is applied, the higher resolution levels (larger images) depend on the results of the lower resolution levels. However, large errors may be introduced in small-misalignment cases due to the large search range for patch matching, which offers more potentially wrong correspondences. To obtain optimal performance, we expect a minimum alignment quality after each resolution level. If the

TABLE I
PARAMETERS FOR PPC-BASED REGISTRATION METHODS AND CORRESPONDING VALUES USED IN THE EVALUATED DATASETS. SCALING FOR THE DIFFERENT RESOLUTION LEVELS IS 0.1, 0.25, 0.5, 0.75 1.0 OF THE ORIGINAL IMAGE SIZE, WHERE THE LOWEST LEVEL USED FOR EACH DATASET IS DENOTED IN THE TABLE AS WELL AS THE CORRESPONDING IMAGE SIZE ON THE LARGER AXIS OF THE LARGEST IMAGE USED IN THE DATASET (IN PARENTHESES). PARAMETER VALUES WHICH DO NOT CHANGE FROM THE DATASET LOCATED IN THE NEIGHBOR LEFT ROW ARE DENOTED WITH “ \leftarrow ” TO EMPHASIZE THAT THEY ARE SHARED BETWEEN DIFFERENT DATASETS. $t_{NGC, RL \neq 1}$ AND $t_{NGC, RL = 1}$ DENOTE THE VALUES OF t_{NGC} FOR ALL RESOLUTION LEVELS EXCEPT THE HIGHEST, AND THE HIGHEST RESOLUTION LEVEL, RESPECTIVELY.

Parameter	Spine	GS Vertebra	GS Angio	TEE
$t_{\bar{s}}$	0.15	\leftarrow	\leftarrow	\leftarrow
r_Ω [pix]	12	\leftarrow	\leftarrow	\leftarrow
r_s [pix]	16	\leftarrow	\leftarrow	\leftarrow
$t_{NGC, RL \neq 1}$	0.1	\leftarrow	\leftarrow	\leftarrow
$t_{NGC, RL = 1}$	0.2	\leftarrow	\leftarrow	\leftarrow
t_θ [°]	87	\leftarrow	86	80
# DLs	5	\leftarrow	\leftarrow	1
RL_{min}	0.25 (154)	0.25 (390)	0.1 (248)	0.25 (256)

image similarity is below a threshold $t_{\bar{s}}$, the results of the resolution level are rejected. To validate this approach and select a value for $t_{\bar{s}}$, different thresholds are applied and the results are compared. This experiment is performed on single vertebrae, as registration of small structures is challenging and the effect of the $t_{\bar{s}}$ is evident. For each vertebra, 60 start positions are generated and the SR is examined to evaluate the robustness.

The results are shown in Fig. 5. We observe that the robustness is improved for increasing values of $t_{\bar{s}}$. For PPC-M, an plateau with an SR of 88 %–89 % is achieved for $t_{\bar{s}} \in [0.1, 0.2]$. For PPC-S, the influence of $t_{\bar{s}}$ is minimal for values below 0.25, with $t_{\bar{s}} = 0$ leading to an SR of 84 % and the values of 0.1 and 0.15 leading to an SR slightly above 85 %. For PPC-A, the maximum robustness is achieved for $t_{\bar{s}} = 0.1$ (77 %), while $t_{\bar{s}} = 0$ leads to an SR of 71 %. However, if $t_{\bar{s}}$ is chosen too high (above 0.25), the robustness is decreased for all methods.

We furthermore observe that for PPC-M as well as PPC-S, the effect of $t_{\bar{s}}$ is small for initial errors in the range [0, 10] mm. For large initial misalignment in the range [20, 30] mm, a higher $t_{\bar{s}}$ only slightly increases the SR for PPC-M and mostly has adversarial effects on the registration. The highest increase in robustness is observed in the range [10, 20] mm. For PPC-M, a large increase in SR is observed for $t_{\bar{s}} \geq 0.1$, increasing from 86 % for $t_{\bar{s}} = 0$ to 95 % for $t_{\bar{s}} = 0.1$ and 96 % for $t_{\bar{s}} = 0.15$. For PPC-S, the effect is smaller, leading to an SR of 88 % for $t_{\bar{s}} \in [0.1, 0.2]$. For PPC-A, $t_{\bar{s}}$ has a high effect in the range of [0, 20] mm. Here, the maximum is very pronounced and the value of $t_{\bar{s}}$ has to be selected carefully. In general, the results suggest that $t_{\bar{s}}$ has to be selected in a way to reject clearly misaligned cases while allowing for rough and imperfect alignment. For all following experiments, we choose a value of $t_{\bar{s}} = 0.15$ if not stated otherwise.

2) Random Studies: In this section, we discuss the registration results on the evaluated datasets. The results for the compared multi-view schemes are shown in Fig. 6.

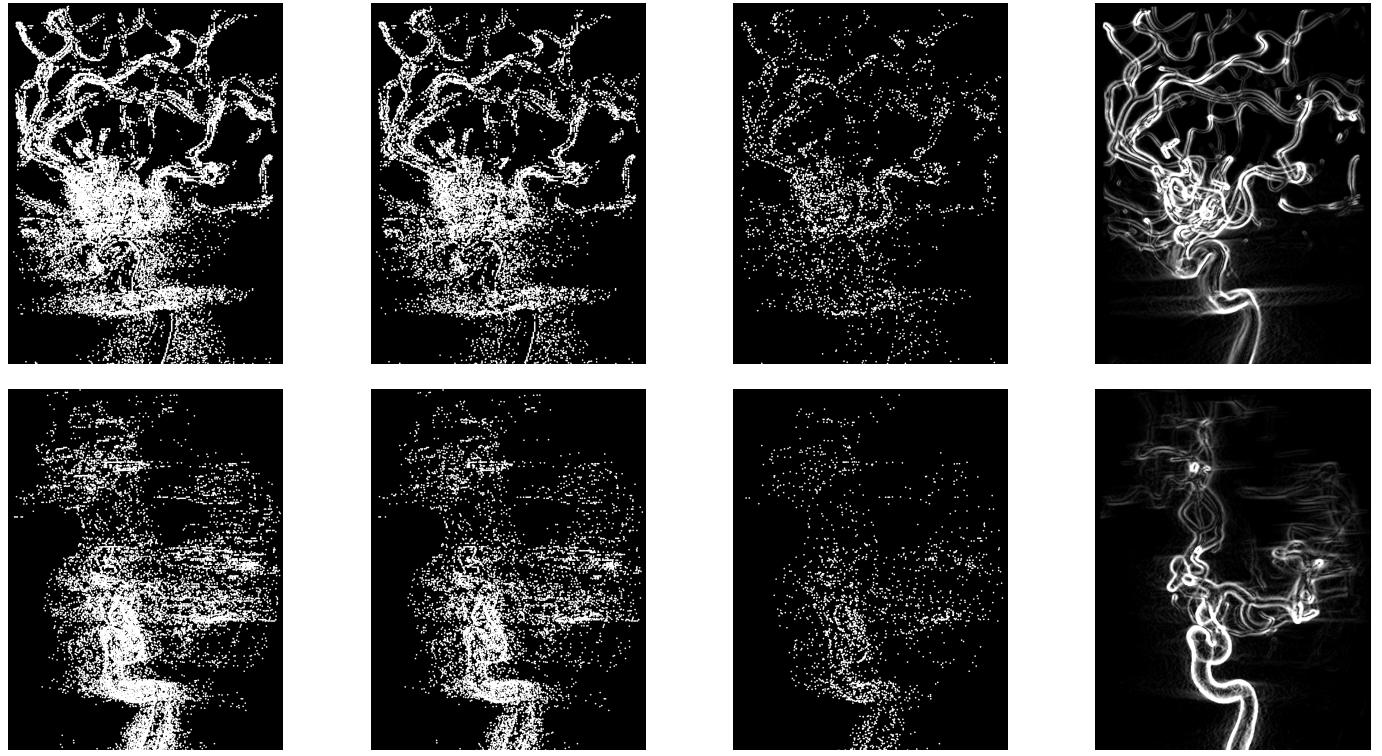


Fig. 4. Contour points selected for different values of t_θ for the LAT (top) and AP (bottom) views of the first volume from the angiography dataset. Selected and projected contour points are shown for 83° , 86° , and 89° (from left to right), 86° being the selected value. For reference, the magnitude of the projected gradient of the volume is also shown for both views (right). While for 89° , the contour points do not cover some parts of the vessels, 83° leads to an increased number of points without considerably increasing the quality of the contour.

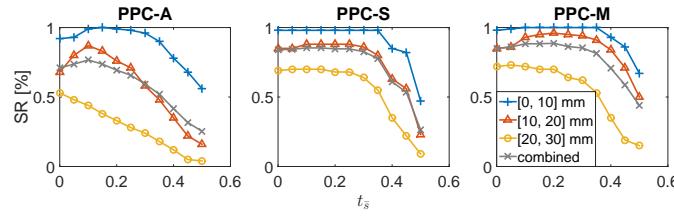


Fig. 5. Success rate (SR) for different values of the success threshold t_s for single-vertebra registration. The results are presented for different subintervals of the initial mTRE. The evaluation was performed with 60 start positions per vertebra in the range $[0, 30]$ mm.

TABLE II
EVALUATION RESULTS FOR THE CEREBRAL ANGIOGRAPHY DATASET [11]
USING NATIVE 2-D IMAGES FOR THE MULTI-VIEW SCENARIO. RESULTS
FOR mTRE INCLUDE MEAN AND STANDARD DEVIATION. RESULTS FOR
MGP + BGB ARE INCLUDED FROM [11] FOR EASIER COMPARISON.

Method	mTRE [mm]	SR [%]	CR [mm]
MGP + BGB [11]	0.23 ± 0.10	90.7	12
PPC-A	0.23 ± 0.07	86.1	9
PPC-S	0.23 ± 0.07	99.6	20
PPC-M	0.22 ± 0.07	98.4	18

a) *GS Cerebral Angiography*: The results for the GS angiography data are shown in Table II. We observe that both PPC-M and PPC-S outperform the reference method (MGP + BGB) regarding robustness by a large margin: CR of 18 mm and 20 mm for PPC-M and PPC-S, respectively, com-

pared to 12 mm for MGP + BGB and SR of 99.6 % and 98.4 % respectively for PPC-M and PPC-S, compared to 90.7 % for MGP + BGB. Also, PPC-S leads to a slightly higher robustness (SR of 99.6 % and CR of 20 mm) compared to PPC-M (SR of 98.4 % and CR of 18 mm). However, PPC-M is slightly more accurate, achieving an mTRE of 0.22 ± 0.07 mm. In general, the differences regarding accuracy and robustness are small between PPC-M and PPC-S. We observe that compared to PPC-M and PPC-S, PPC-A is far less robust (SR of 86.1 % and CR of 9 mm) and does not outperform the reference method. One possible reason is the fact that registrations for both views are performed sequentially and large errors in any view lead to a failed registration, resulting in the low robustness. All PPC-based methods achieve an accuracy on par with or slightly better than MGP + BGR, without relying on a refinement step.

b) *Whole Spine*: The results for spine registration evaluation are shown in Table III. Similarly to the angiography dataset, we observe that PPC-A is the least robust method, achieving an SR of 92.0 % and a CR of 20 mm. The CR is improved only slightly over single-view registration using the AP view. The highest robustness is achieved by PPC-M (SR of 98.2 % and CR of 25 mm), followed by PPC-S (SR of 97.1 % and CR of 25 mm). The accuracy of the methods is comparable. While PPC-A achieves 0.46 ± 0.12 mm, both PPC-M and PPC-S achieve 0.47 ± 0.11 mm. For PPC-M, the error is decreased by 0.55 mm compared to using only the AP view and by 0.19 mm using the LAT view. This shows that for the spine, high accuracy can be already achieved

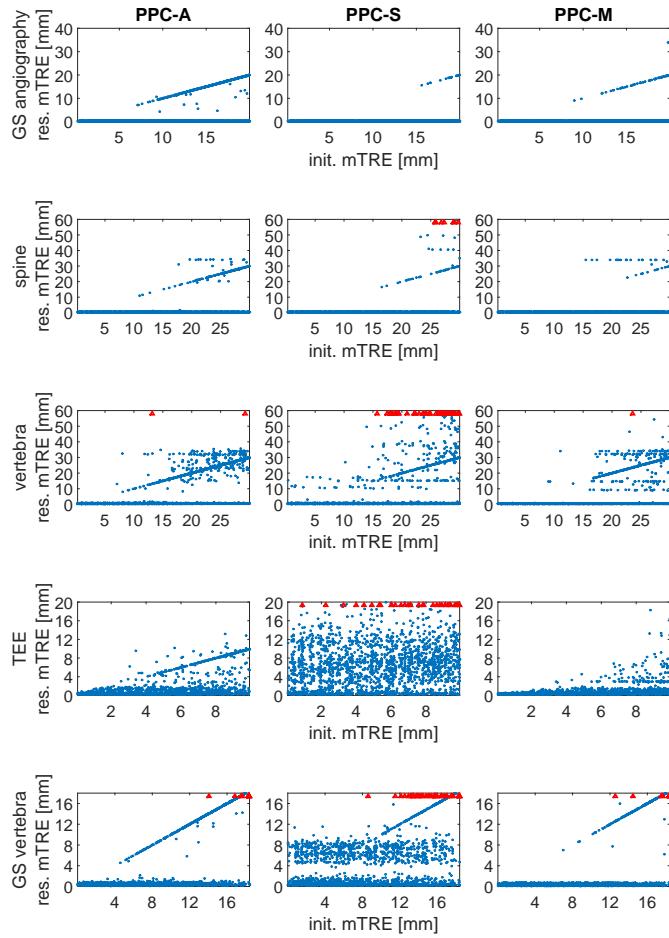


Fig. 6. Scatter plots showing the initial mTRE and the mTRE after registration for all cases for the different data sets and the compared methods. Red triangles indicate registration cases where the resulting error is outside the shown range. Note the different scale of plots for different datasets. For PPC-S, many failed cases lead to large errors outside of the plotted range (red triangles). This is due to the unconstrained single-view registration, which is performed in PPC-S and can lead to large errors for the out-of-plane translation.

TABLE III
EVALUATION RESULTS FOR THE SPINE DATASET. RESULTS FOR mTRE INCLUDE MEAN AND STANDARD DEVIATION. PPC INDICATES SINGLE-VIEW REGISTRATION.

Structure	Views	Method	mTRE [mm]	SR [%]	CR [mm]
Spine	AP	PPC	1.02±0.40	88.8	16
	LAT	PPC	0.66±0.33	84.9	8
	AP & LAT	PPC-A	0.46±0.12	92.0	20
		PPC-S	0.47±0.11	97.1	25
Single Vertebrae	AP	PPC	1.34±0.49	6.8	0
	LAT	PPC	1.12±0.48	44.9	0
	AP & LAT	PPC-A	0.50±0.15	71.6	11
		PPC-S	0.51±0.15	87.6	16
		PPC-M	0.47±0.12	86.5	16

using single-view registration for some views, and the gains of multi-view registration are limited for these views.

c) *Single Vertebrae*: Results for single-vertebra registration are shown in Table III. We observe that for single-view registration, the accuracy of the registration is decreased considerably compared to the full spine for single-view regis-

TABLE IV
EVALUATION RESULTS FOR REGISTRATION OF THE TEE PROBE. RESULTS FOR MTRE INCLUDE MEAN AND STANDARD DEVIATION. PPC INDICATES SINGLE-VIEW REGISTRATION.

Method	mTRE [mm]	SR [%]	CR [mm]
PPC	1.11±0.51	13.0	0
PPC-A	0.56±0.40	79.0	4
PPC-S	0.61±0.45	26.1	0
PPC-M	0.51±0.29	94.2	7

tration. For multi-view registration, the accuracy is decreased for PPC-A and PPC-S (to 0.50 ± 0.15 mm and 0.51 ± 0.15 mm, respectively). PPC-M achieves an accuracy of 0.47 ± 0.12 mm, on par with spine registration. In general, the robustness of the registration is decreased compared to the whole spine. The highest robustness is achieved by PPC-S, resulting in a SR of 87.6 % and a CR of 16 mm. PPC-M achieves an only slightly lower SR of 86.5 % and the same value for the CR, namely 16 mm.

The SR (71.6 %) as well as the CR (11 mm) are considerably smaller for PPC-A. For PPC-S, we observe failed cases for relatively small initial errors (see Fig. 6). This indicates that the refinement step at of the PPC-S method cannot handle the increased out-of-plane error due to the smaller structures compared to the registration of the whole spine. This is further substantiated by our experiments on convergence (see Sec. V-C5), where we show that high errors are present for PPC-S until the refinement phase and by the results for the TEE probe as well as the GS vertebra dataset.

d) *TEE Probe*: The results for the TEE probe are shown in Table IV. We observe that here, PPC-M clearly outperforms PPC-S regarding robustness, leading to a CR of 7 mm and an SR of 94.2 %, compared to 0 mm and 26.1 % for PPC-S. PPC-A achieves a CR of 4 mm and an SR of 79.4 %, outperforming PPC-S, but leading to a decrease of the SR of around 15 percentage points and a decrease of the CR of 3 mm compared to PPC-M. The best accuracy is achieved by PPC-M (0.51 ± 0.29 mm). While PPC-A is more robust compared to PPC-S, it does not reach the robustness of PPC-M, again showing the instability of the method. The low robustness of PPC-S can be explained by the fact that large out-of-plane errors are present in single-view registration, which cannot be recovered using the second view and the largest resolution level, i.e. full image resolution. We also observe a large number of cases with errors around 15 mm or below, even for small initial errors (see Fig. 6). This shows that the error is introduced by the method regardless of a good initial alignment, which is typical for translational out-of-plane errors in small structures. Furthermore, we observe a relatively low accuracy for all methods, with errors up to 2 mm mTRE for a large set of cases (see Fig. 6). Compared to the other data sets, no clear margin between successful and failed cases exists. This shows that rough alignment does not guarantee convergence. We observe that one reason is the symmetry of the probe. While the outline of the probe visually seems to be well aligned after registration, a rotation around the centerline can be observed. Note that while such errors are present for all compared methods, PPC-M shows a higher

TABLE V

EVALUATION RESULTS FOR REGISTRATION OF THE GS VERTEBRA DATASET [32]. RESULTS FOR THE BGB METHOD [7] FROM THE ORIGINAL PUBLICATION ARE PROVIDED WITH THE DATASET AND INCLUDED FOR EASE OF COMPARISON. THE SR_M (SR USING THE MAXIMUM TARGET POINT DISTANCE INSTEAD OF THE MEAN VALUE) IS GIVEN FOR DIFFERENT INTERVALS OF INITIAL ERRORS AND RANGES OF ROTATION. RESULTS FOR mTRE INCLUDE MEAN AND STANDARD DEVIATION.

Method	Accuracy After Registration				SR _M [%]			mTRE [mm]	SR [%]	CR [mm]
	Rms(TRE) [mm]	Max(TRE) [mm]	Rms(γ) [$^{\circ}$]	Max(γ) [$^{\circ}$]	[0, 6] mm [0, 17.2] $^{\circ}$	[6, 12] mm [17.2, 34.4] $^{\circ}$	[12, 18] mm [34.4, 51.7] $^{\circ}$			
BGB	0.35	1.59	0.36	0.94	94.6	47.8	11.2	0.32 ± 0.13	51.4	3
PPC-A	0.45	1.78	0.62	3.64	98.9	78.8	35.1	0.40 ± 0.20	71.2	6
PPC-S	0.54	1.98	1.22	5.42	62.3	60.4	40.7	0.49 ± 0.32	56.6	0
PPC-M	0.39	0.92	0.34	0.72	100	97.2	65.0	0.35 ± 0.17	87.3	8

accuracy for small initial errors, especially below 2 mm. Here, PPC-M achieves an accuracy of 0.35 ± 0.12 mm, while PPC-A achieves 0.44 ± 0.28 mm.

e) *GS Vertebrae*: The results for the GS vertebra dataset are shown in Tab. V. Compared to BGB, PPC-M achieves a higher robustness, leading to an overall SR of 87.3 %, compared to 51.4 % for BGB. While BGB leads to a higher mean accuracy (Rms(TRE) of 0.35 mm compared to 0.39 mm for PPC-M), the results are comparable and PPC-M achieves a lower maximum error value (0.92 mm vs. 1.59 mm for BGB). PPC-M outperforms both PPC-A and PPC-S regarding robustness and accuracy. We observe that for PPC-S, registration fails in many cases even for small initial misalignment (see Fig. 6), leading to an error of roughly 4 mm or above. This sharp border indicates that for larger out-of-plane errors of the selected view, the refinement step of PPC-S is not successful. While it may be possible to improve the results by performing a multi-resolution refinement, this would further increase the computational complexity of the method. PPC-S and PPC-A lead to an Rms(TRE) of 0.54 mm and 0.45 mm, respectively. Interestingly, PPC-S leads to a considerably lower accuracy in comparison to PPC-A, although the refinement step of PPC-S is identical to PPC-A. This indicates that the large out-of-plane errors present for PPC-S before the refinement can affect the accuracy of the result as well as the robustness.

3) *Regularized Motion Estimation*: We further investigate the possibility to increase the registration robustness using L2-regularized motion estimation. We perform experiments for the single-vertebra and TEE probe datasets, as these small structures lead to a lower robustness compared to larger structures and may benefit mostly from the regularized motion estimation. We generate two start positions per mm for the vertebrae (a total of 60 start positions) and 10 start positions per mm for the TEE probe (a total of 100 start positions) for each tested image combination. We investigate on the registration robustness using different regularizer weights. We do not enforce a minimum alignment quality using $t_{\bar{s}}$ to avoid interdependencies between these parameters, as a high λ may lead to lower similarity values \bar{s} due to incomplete convergence while the registration error is actually decreased. Furthermore, estimation of translations in depth is not disabled for the lowest resolution level for PPC-S and single-view registration, as it is the task of the regularizer to ensure robust motion estimation.

The results for the single-vertebra and TEE probe registration with regularization are shown in Fig. 7. In both cases, the

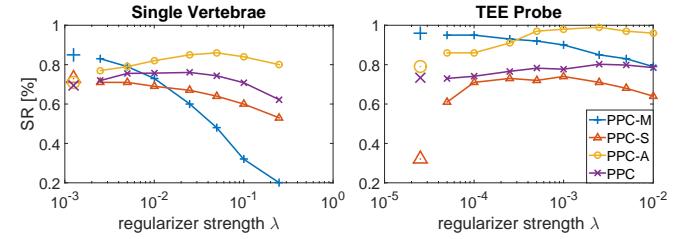


Fig. 7. SR for different values of the regularizer weight λ . PPC indicates single-view registration averaged over both views and with a success criterion of 2 mm mRPD. Larger symbols indicate results without regularization, where no success criterion was applied and estimation of translational motion in depth not restricted for PPC-S and single-view (see Sec. V-C3).

robustness of PPC-M is decreased by increasing λ . For PPC-A and single-view registration, an increased robustness can be observed as well as for PPC-S for registration of the TEE probe. For the TEE probe, the SR for single-view registration is increased only slightly, while the SR for PPC-S is improved considerably. As the SR for PPC-S is very low without regularization compared to single-view registration, this indicates large out-of-plane errors, which cannot be recovered by the refinement step but can be avoided by the regularized motion estimation. The large effect of the regularizer on PPC-S registration substantiates the conclusion that high out-of-plane errors are responsible for the poor robustness of the method, as these errors are reduced by the regularizer.

In general, PPC-A outperforms the other methods for larger values of λ . For vertebra evaluation, it achieves an SR of 86 % for $\lambda = 0.05$. We performed an evaluation for $\lambda = 0.05$ on all start positions and achieved an SR of 82.8 % and a CR of 16 mm again using $t_{\bar{s}} = 0.15$ (SR of 84.2 % and a CR of 8 mm without $t_{\bar{s}}$).

For the TEE probe, the highest SR of 99 % is achieved by PPC-A for $\lambda = 2.5 \cdot 10^{-3}$. We again performed an evaluation on all start positions and achieved an SR of 97.0 % and a CR of 8 mm, using $t_{\bar{s}} = 0.15$ (SR of 96.9 % and a CR of 8 mm without $t_{\bar{s}}$).

The results indicate that while regularization may be beneficial for single-view registration and methods involving single-view registration, i. e. PPC-A and PPC-S, PPC-M is robust and regularization only hinders the convergence. Furthermore, regularization is beneficial for PPC-S for the TEE probe, while not leading to an advantage in vertebra registration. One possible reason is that out-of-plane motion is reduced by the regularization, preventing the out-of-plane errors to become

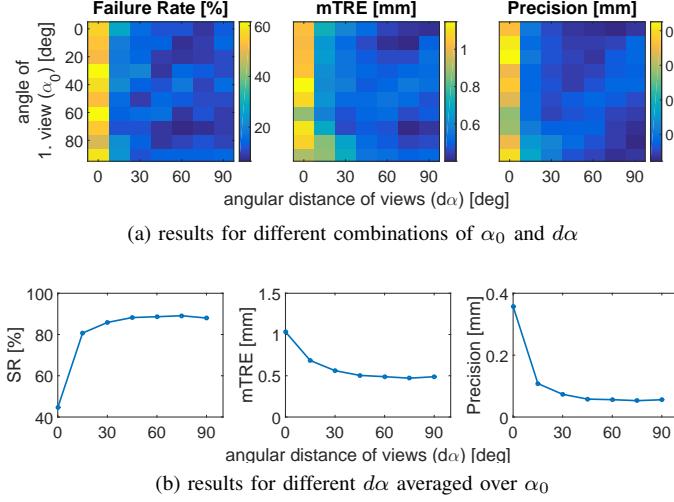


Fig. 8. Results of systematic angle evaluation for PPC-M and single-vertebra registration. Precision is the mean standard deviation of mTRE averaged over all datasets.

too large for the refinement step to recover.

4) Systematic Angle Evaluation: As small angles between the views decrease the obstruction of the clinical workflow, we systematically investigate the registration performance using the PPC-M method under different viewing angles. We perform the evaluation on single vertebra, as the effect of using a second view is large regarding both accuracy and robustness. Furthermore, images for many different angles are available. We vary both the angle of the first view α_0 and the angular distance between the views $d\alpha$. The results are shown in Fig. 8. An angle of α_0 roughly corresponds to the posterior-anterior (PA) view and $d\alpha = 0$ indicates single-view registration.

As indicated in Fig. 8, the SR of multi-view registration using $d\alpha = 15^\circ$ is largely increased compared to single-view registration (80.7 % vs. 44.6 %). It continues to increase for larger angles and reaches a value of 85.8 % for $d\alpha = 30^\circ$ and a plateau with an average SR of 88.5 % for $d\alpha \geq 45^\circ$. The accuracy (see Fig. 8 (b) and (e)) is also strongly increased for multi-view registration. The mTRE decreases from 1.03 mm for single-view registration to 0.68 mm for $d\alpha = 15^\circ$ and 0.50 mm for $d\alpha = 30^\circ$. From there, it decreases only slightly, achieving the minimum value (0.47 mm) for $d\alpha = 75^\circ$. A similar effect is observed for the precision, which is 0.36 mm for single-view registration and decreases to 0.11 mm for $d\alpha = 15^\circ$ and 0.06 mm for $d\alpha = 30^\circ$. It decreases slightly from there on and reaches a minimum of 0.05 mm for $d\alpha = 75^\circ$.

In general, we observe that the registration performance is increased for multi-view registration. However, a high accuracy is already achieved for $d\alpha = 30^\circ$. The SR is increased until $d\alpha = 45^\circ$ and varies only slightly for higher angles. In order to decrease the obstruction of the clinical workflow, $d\alpha = 30^\circ$ (or even $d\alpha = 15^\circ$) can be used, decreasing the registration performance only slightly. This shows that angles below 90° between the views are sufficient. While we show that 30° between the views is sufficient for single-vertebra registration, a smaller angle may suffice for larger and more complex structures. On the other hand, smaller or less complex

structures such as cervical vertebrae or the TEE probe may exhibit less change in appearance for small angle differences and a larger angle above 30° may be necessary for reliable registration.

Our findings are in line with Uneri et al. [13], where the authors demonstrate that high registration accuracy can be achieved for small angles, reaching the optimal accuracy for angles $\geq 20^\circ$ between the views. Note that the results are not directly comparable, as larger anatomical structures were registered by Uneri et al. and only accuracy was evaluated. Nevertheless, this shows that relatively small angles can be sufficient for optimal or near-optimal multi-view registration and that this effect is not specific to a single registration method.

5) Convergence: We compare the convergence of the different methods. The results for single-vertebra registration can be seen in Fig. 9. For PPC-M and regularized PPC-A, we observe that for every resolution level, the minimum mean and median error is achieved after a small number of iterations. For the lowest resolution level, the error is increased in some of the cases, leading to an increasing 95th percentile. For PPC-A without regularization, the error is increasing on the lowest resolution level for many cases (increasing 75th percentile and mean value). The median error is decreasing for the lowest resolution level. However, the convergence is slower compared to PPC-M and regularized PPC-A. Additionally, an oscillation is observed. This behavior is expected if one of the views leads to correct registration, while the other view introduces a higher error. For PPC-S, we observe a high median error which decreases steadily throughout the resolution levels. We observe an instability on the lowest resolution level when motion estimation in depth is enabled, which leads to a highly increased 95th percentile and mean value, while the median is slightly reduced and the 75th percentile does not change much. The high-error cases are recovered due to discarding of the results of this resolution level. Furthermore, cases with relatively high errors are present until the refinement step. This illustrates the challenge of PPC-S for the registration of small structures, where high out-of-plane errors are present.

6) Runtime: The runtime for different structures is shown in Tab. VI. We observe that in general, PPC-M and PPC-A lead to comparable runtime, while PPC-S has an increased runtime due to the refinement step and the additional iterations at lowest resolution levels without motion in depth. Furthermore, comparing the vertebra and the GS vertebra datasets, we observe that the registration is faster for the GS dataset. This is most probably due to the lower resolution of the used volume. As discussed in Sec. V-C5, only a few iterations are needed for each resolution level. However, more iterations are currently performed, especially for later resolution levels. Therefore, the runtime can be further improved by limiting the number of iterations or by carefully choosing a convergence criterion to avoid unnecessary iterations. Additional optimizations can also be investigated, e.g. reducing the 3-D resolution or limiting the number of contour points.

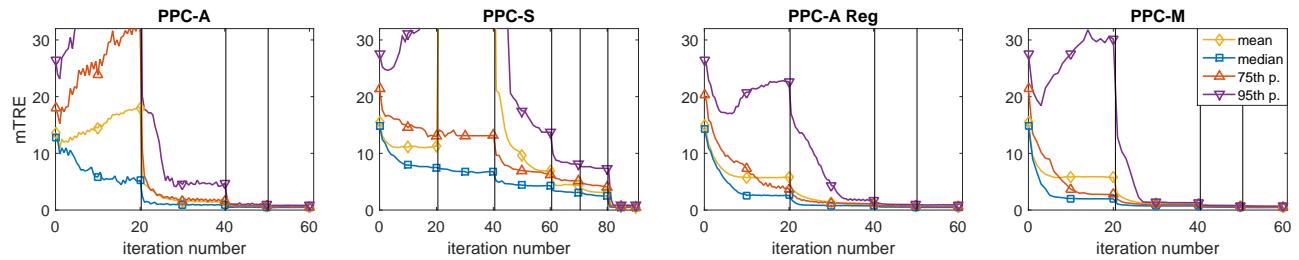


Fig. 9. Convergence evaluation for single-vertebra registration. Vertical lines indicate a change in the resolution level. Iteration number 0 indicates initial error. For PPC-S, registration is performed for the lowest resolution first without motion estimation in the viewing direction and then with. These steps are also separated. Last resolution level for PPC-S is the refinement step. Only the selected view is considered for PPC-S. For PPC-A, one iteration consists of two single-view registration steps. In case of convergence before the iteration limit, the result is set to a constant error equal to the last performed iteration. For each vertebra, 25 registrations for an initial mTRE in the range [5, 30] mm are considered.

TABLE VI

RUNTIME (MEAN VALUE AND STANDARD DEVIATION) FOR PPC-BASED MULTI-VIEW REGISTRATION METHODS FOR THE DIFFERENT DATASETS. EXPERIMENTS WERE PERFORMED ON A MACHINE WITH A XEON E5-2620 CPU WITH 2.4 GHz (BOOST UP TO 3.2 GHz), 6 CORES AND A NVIDIA QUADRO M4000 GPU (1664 CUDA CORES, 773 MHZ, MEMORY BANDWIDTH: 192.3 GB/S) OR ON A SLOWER SYSTEM. FOR EACH DATASET, ALL METHODS WERE RUN ON THE SAME SYSTEM.

Dataset	Method	Runtime [s]
angio	PPC-A	49.7±25.1
	PPC-S	65.2±12.3
	PPC-M	42.7±9.5
spine	PPC-A	35.3±11.0
	PPC-S	58.0±10.3
	PPC-M	40.4±9.6
vertebra	PPC-A	12.6±3.5
	PPC-S	17.9±4.1
	PPC-M	13.7±4.5
TEE probe	PPC-A	11.8±5.2
	PPC-S	13.7±5.5
	PPC-M	8.8±4.7
GS vertebra	PPC-A	7.8±1.3
	PPC-S	11.8±1.5
	PPC-M	7.0±1.4

VI. CONCLUSION

In this paper, the PPC model is extended to be independent of the camera coordinate system. The extension is based on the facts that the PPC model describes the distances of 3-D points to planes representing contours in the 2-D image and distances are invariant to rigid transformations. The view-independent PPC model enables motion estimation directly in a coordinate system related to the camera coordinates by a rigid transformation. Thereby, it enables motion estimation combining correspondences from multiple views. Additionally, it allows meaningful regularized motion estimation by allowing to shift the coordinate system to make the rotational and translational components of the motion independent from each other.

Different multi-view registration schemes are compared and investigations on the effect of regularized motion estimation are carried out. Experiments are performed using clinical data sets, i. e. a GS cerebral angiography dataset as well as a spine dataset and a GS vertebra dataset. For the spine dataset, the evaluation is performed for the whole visible region of the spine as well as for single vertebrae. Additionally, experiments for TEE probe registration are performed on a porcine dataset.

The results demonstrate that PPC-M, in contrast to PPC-S

and PPC-A, performs reliably for a wide range of use cases. While PPC-A in combination with regularized motion estimation and PPC-S can achieve comparable results or even slightly outperform PPC-M for some data sets, PPC-M performs well for all use cases and shows a fast convergence. It is furthermore not dependent on regularized motion estimation and achieves the highest accuracy among PPC-based methods for small structures.

PPC-M achieves a high robustness and accuracy for different use cases. For the GS angiography dataset, our method demonstrates a highly increased robustness comparing to the state-of-the-art reference method while maintaining a high accuracy without requiring a refinement step. On the GS vertebra dataset, we compare our method to BGB, which is often used as a refinement step in other methods due to the high accuracy it can achieve. We demonstrate a highly increased robustness of our method, while achieving a comparable accuracy.

We furthermore investigate the dependency of the registration on the angulation of the views and show that relatively small angles around 30° allow for robust and accurate registration for single vertebrae. While larger angles may improve the registration, the improvement is small and relatively small angles around 30°, or even 15°, may be advantageous to reduce the obstruction of the workflow while still allowing reliable registration.

VII. FUTURE WORK

The weights for the used correspondences are defined based on the local image similarity. However, more advanced weighting criteria can be considered, i. e. by extending our previously proposed learning-based weighting method [23] to multi-view registration. Furthermore, the success criterion and quality measure used for the individual resolution levels can be improved, and possibly learned. This would lead to more robust and faster registration and possibly enable an automatic detection of misregistrations. In general, the amount of heuristics used for the registration can be reduced. The registration of the TEE probe is challenging due to its symmetrical structure. As the probe contains symmetry-breaking structures, exploiting these structures is a promising direction. Another direction is to extend the method to other modalities, e. g. to MRI volumes or CAD models of implants.

While we show the effectiveness of the proposed method for different structures, the evaluation is not exhaustive and further experiments are needed to characterize the performance of the method for other use cases. Especially, evaluations can be performed for cervical and higher thoracic vertebrae in order to make sure that the proposed method performs well for those structures and to investigate whether the proposed view distance of 30° is sufficient for those structures. In order to further improve the robustness of the method, multi-start approaches can be investigated, e.g. as a way to better cope with the repetitive structure of the spine.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful comments that significantly improved the quality of this paper.

DISCLAIMER

The concept and software presented in this paper are based on research and are not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

REFERENCES

- [1] S. Rossitti and M. Pfister, "3D Road-Mapping in the Endovascular Treatment of Cerebral Aneurysms and Arteriovenous Malformations," *Interv Neuroradiol*, vol. 15, no. 3, pp. 283–290, 2009.
- [2] M. Hoffmann, C. Kowalewski, A. Maier, K. Kurzidim, N. Strobel, and J. Horngger, "Contrast-Based 3D/2D Registration of the Left Atrium: Fast Versus Consistent," *Int J Biomed Imaging*, vol. 2016, 2016.
- [3] J. Wang, M. Kreiser, L. Wang, N. Navab, and P. Fallavollita, "Augmented Depth Perception Visualization in 2D/3D Image Fusion," *Comput Med Imaging Graph*, vol. 38, no. 8, pp. 744–752, 2014.
- [4] P. Markelj, D. Tomažević, B. Likar, and F. Pernuš, "A Review of 3D/2D Registration Methods for Image-Guided Interventions," *Med Image Anal*, vol. 16, no. 3, pp. 642–661, 2010.
- [5] R. Liao, L. Zhang, Y. Sun, S. Miao, and C. Chefd'Hotel, "A Review of Recent Advances in Registration Techniques Applied to Minimally Invasive Therapy," *IEEE Trans Multimedia*, vol. 15, no. 5, pp. 983–1000, 2013.
- [6] C. Gendrin, P. Markelj, S. A. Pawiro, J. Spoerk, C. Bloch, C. Weber, M. Figl, H. Bergmann, W. Birkfellner, B. Likar *et al.*, "Validation for 2D/3D Registration II: The Comparison of Intensity-and Gradient-based Merit Functions Using a New Gold Standard Data Set," *Med Phys*, vol. 38, no. 3, pp. 1491–1502, 2011.
- [7] D. Tomažević, B. Likar, T. Slivnik, and F. Pernuš, "3-D/2-D Registration of CT and MR to X-Ray Images," *IEEE Trans Med Imag*, vol. 22, no. 11, pp. 1407–1416, 2003.
- [8] P. Markelj, D. Tomažević, B. Likar, and F. Pernuš, "Robust Gradient-Based 3-D/2-D Registration of CT and MR to X-ray Images," *IEEE Trans Med Imag*, vol. 27, no. 12, pp. 1704–1714, 2008.
- [9] H. Livyatyan, Z. Yaniv, and L. Joskowicz, "Gradient-Based 2-D/3-D Rigid Registration of Fluoroscopic X-ray to CT," *IEEE Trans Med Imag*, vol. 22, no. 11, pp. 1395–1406, 2003.
- [10] A. Kubias, F. Deinzer, T. Feldmann, D. Paulus, B. Schreiber, and T. Brunner, "2D/3D Image Registration on the GPU," *Pattern Recognition and Image Analysis*, vol. 18, no. 3, pp. 381–389, 2008.
- [11] U. Mitrović, Ž. Špiclin, B. Likar, and F. Pernuš, "3D-2D Registration of Cerebral Angiograms: A Method and Evaluation on Clinical Images," *IEEE Trans Med Imag*, vol. 32, no. 8, pp. 1550–1563, 2013.
- [12] Ž. Špiclin, B. Likar, and F. Pernuš, "Fast and Robust 3D to 2D Image Registration by Backprojection of Gradient Covariances," in *6th International Workshop on Biomedical Image Registration*. Springer, 2014, pp. 124–133.
- [13] A. Uneri, Y. Otake, A. S. Wang, G. Kleinszig, S. Vogt, A. J. Khanna, and J. H. Siewersden, "3D-2D Registration for Surgical Guidance: Effect of Projection View Angles on Registration Accuracy," *Phys Med Biol*, vol. 59, no. 2, pp. 271–287, 2014.
- [14] M. Kaiser, M. John, T. Heimann, A. Brost, T. Neumuth, and G. Rose, "2D/3D Registration of TEE Probe From Two Non-orthogonal C-arm Directions," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*. Springer, 2014, pp. 283–290.
- [15] A. Duménil, A. Kaladjii, M. Castro, C. Göksu, A. Lucas, and P. Haigron, "A Versatile Intensity-based 3D/2D Rigid Registration Compatible With Mobile C-arm for Endovascular Treatment of Abdominal Aortic Aneurysm," *Int J Comput Assist Radiol Surg*, vol. 11, no. 9, pp. 1713–1729, 2016.
- [16] Y. Otake, A. S. Wang, J. W. Stayman, A. Uneri, G. Kleinszig, S. Vogt, A. J. Khanna, Z. L. Gokaslan, and J. H. Siewersden, "Robust 3D-2D Image Registration: Application to Spine Interventions and Vertebral Labeling in the Presence of Anatomical Deformation," *Phys Med Biol*, vol. 58, no. 23, p. 8535, 2013.
- [17] N. Hansen, "The CMA Evolution Strategy: A Comparing Review," in *Towards a New Evolutionary Computation*. Springer, 2006, pp. 75–102.
- [18] J. Schmid and C. Chênes, "Segmentation of X-ray Images by 3D-2D Registration Based on Multibody Physics," in *12th Asian Conference on Computer Vision*, 2014.
- [19] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2003, p. 200.
- [20] J. Wang, R. Schaffert, A. Borsdorf, B. Heigl, X. Huang, J. Horngger, and A. Maier, "Dynamic 2-D/3-D Rigid Registration Framework Using Point-To-Plane Correspondence Model," *IEEE Trans Med Imag*, vol. 36, no. 9, pp. 1939–1954, 2017.
- [21] M. Utzschneider, J. Wang, R. Schaffert, A. Borsdorf, and A. Maier, "Real-Time-Capable GPU-Framework for Depth-Aware Rigid 2D/3D Registration," in *Bildverarbeitung für die Medizin 2017*. Springer, 2017, pp. 185–190.
- [22] R. Schaffert, J. Wang, P. Fischer, A. Borsdorf, and A. Maier, "Multi-View Depth-Aware Rigid 2-D/3-D Registration," in *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2017.
- [23] ———, "Metric-Driven Learning of Correspondence Weighting for 2-D/3-D Image Registration," in *German Conference on Pattern Recognition*. Springer, 2018, pp. 140–152.
- [24] J. Wang, A. Borsdorf, B. Heigl, T. Köhler, and J. Horngger, "Gradient-Based Differential Approach for 3-D Motion Compensation in Interventional 2-D/3-D Image Fusion," in *International Conference on 3D Vision*, 2014, pp. 293–300.
- [25] K. He, J. Sun, and X. Tang, "Guided Image Filtering," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [26] J. Canny, "A Computational Approach to Edge Detection," in *Readings in Computer Vision*, 1987, pp. 184–203.
- [27] J. Wang, A. Borsdorf, and J. Horngger, "Depth-Layer-Based Patient Motion Compensation for the Overlay of 3D Volumes Onto X-ray Sequences," in *Bildverarbeitung für die Medizin 2013*. Springer, 2013, pp. 128–133.
- [28] W. Wein, B. Roepke, and N. Navab, "2D/3D Registration Based on Volume Gradients," in *Medical Imaging. SPIE*, 2005, pp. 144–150.
- [29] I. Santamaría, P. P. Pokharel, and J. C. Principe, "Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization," *IEEE Trans Signal Process*, vol. 54, no. 6, pp. 2187–2197, 2006.
- [30] P. W. Holland and R. E. Welsch, "Robust Regression using Iteratively Reweighted Least-Squares," *Commun Stat Theory Methods*, vol. 6, no. 9, pp. 813–827, 1977.
- [31] E. B. van de Kraats, G. P. Penney, D. Tomažević, T. van Walsum, and W. J. Niessen, "Standardized Evaluation Methodology for 2-D-3-D Registration," *IEEE Trans Med Imag*, vol. 24, no. 9, pp. 1177–1189, 2005.
- [32] D. Tomažević, B. Likar, and F. Pernuš, "Gold Standard" Data for Evaluation and Comparison of 3D/2D Registration Methods," *Comput. Aided Surg.*, vol. 9, pp. 137–144, 02 2004.
- [33] U. Mitrović, B. Likar, F. Pernuš, and Ž. Špiclin, "3D-2D Registration in Endovascular Image-guided Surgery: Evaluation of State-of-the-art Methods on Cerebral Angiograms," *Int J Comput Assist Radiol Surg*, vol. 13, no. 2, pp. 193–202, 2018.
- [34] M. Kaiser, "Fusion of Interventional Ultrasound & X-ray," Ph.D. dissertation, Otto von Guericke University Magdeburg, 2016.
- [35] S. Miao, Z. J. Wang, and R. Liao, "A CNN Regression Approach for Real-time 2D/3D Registration," *IEEE Trans Med Imag*, vol. 35, no. 5, pp. 1352–1363, 2016.