

Registration of 3D freehand ultrasound to a bone model for orthopedic procedures of the forearm

Journal Article**Author(s):**

Ciganovic, Matija; [Ozdemir, Firat](#) ; Péan, Fabien; Fuernstahl, Philipp; Tanner, Christine; [Goksel, Orcun](#) 

Publication date:

2018-06

Permanent link:

<https://doi.org/10.3929/ethz-b-000258148>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

International Journal of Computer Assisted Radiology and Surgery 13(6), <https://doi.org/10.1007/s11548-018-1756-0>



Registration of 3D freehand ultrasound to a bone model for orthopedic procedures of the forearm

Matija Ciganovic¹ · Firat Ozdemir¹ · Fabien Pean¹ · Philipp Fuernstahl² · Christine Tanner¹ · Orcun Goksel¹

Received: 27 January 2018 / Accepted: 26 March 2018 / Published online: 5 April 2018
© CARS 2018

Abstract

Purpose For guidance of orthopedic surgery, the registration of preoperative images and corresponding surgical plans with the surgical setting can be of great value. Ultrasound (US) is an ideal modality for surgical guidance, as it is non-ionizing, real time, easy to use, and requires minimal (magnetic/radiation) safety limitations. By extracting bone surfaces from 3D freehand US and registering these to preoperative bone models, complementary information from these modalities can be fused and presented in the surgical realm.

Methods A partial bone surface is extracted from US using phase symmetry and a factor graph-based approach. This is registered to the detailed 3D bone model, conventionally generated for preoperative planning, based on a proposed multi-initialization and surface-based scheme robust to partial surfaces.

Results 36 forearm US volumes acquired using a tracked US probe were independently registered to a 3D model of the radius, manually extracted from MRI. Given intraoperative time restrictions, a computationally efficient algorithm was determined based on a comparison of different approaches. For all 36 registrations, a mean (\pm SD) point-to-point surface distance of 0.57 (\pm 0.08) mm was obtained from manual gold standard US bone annotations (not used during the registration) to the 3D bone model.

Conclusions A registration framework based on the bone surface extraction from 3D freehand US and a subsequent fast, automatic surface alignment robust to single-sided view and large false-positive rates from US was shown to achieve registration accuracy feasible for practical orthopedic scenarios and a qualitative outcome indicating good visual image alignment.

Keywords Tracked ultrasound · Bone segmentation · Iterative refinement

Introduction

Orthopedic surgery is a branch of medicine that focuses on injuries of the musculoskeletal system, gaining ever more importance with increasing life expectancy and the elderly being particularly prone to bone injuries [1]. Fractures are a common case of orthopedic trauma, where the forearm is highly vulnerable to injuries [1,2]. Twenty-five percent of

fractures in pediatric population and 18% for the elderly are those of the distal radius [2,3].

A common step in the preoperative planning of orthopedic surgery is the acquisition of a computed tomography (CT) scan, due to its suitability for visualizing bone tissue. However, little information can be inferred from CT images about soft tissue details (e.g., muscles, ligaments, nerve fibers, and smaller blood vessels), which are potentially relevant as structures at risk during intervention and/or as soft tissue constraints on musculoskeletal function that should ideally be taken into account for an optimal surgical outcome. Ultrasound (US) can provide soft tissue information for preoperative planning as well as real-time guidance feedback for surgical navigation. Nevertheless, US allows limited field of view, often in 2D slices, where the bones also act as acoustic reflectors preventing visibility behind their top surface. Therefore, given an accurate registration technique, detailed and precise 3D bone models from other

✉ Matija Ciganovic
matija.ciganovic@vision.ee.ethz.ch

Orcun Goksel
ogoksel@vision.ee.ethz.ch

¹ Computer-Assisted Applications in Medicine (CAiM), ETH Zurich, Zurich, Switzerland

² Computer Assisted Research and Development (CARD), Balgrist University Hospital, University of Zurich, Zurich, Switzerland

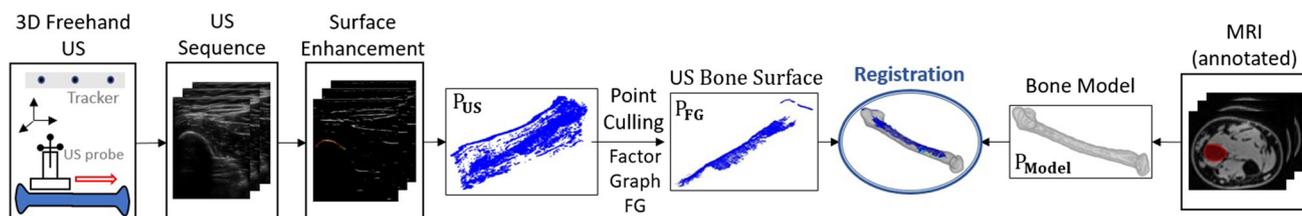


Fig. 1 Registration pipeline overview

imaging modalities can be enriched by ultrasonic soft tissue information as well as aligned to intraoperative coordinates for surgical guidance. We aim at such multi-modality fusion in this work.

Given the limited field of view of individual US images/volumes, more comprehensive information can be enabled by panoramic reconstruction of many such images [4,5]. For 3D reconstruction from 2D freehand US, which are conventionally acquired in clinics, image-based techniques [6] were also proposed for relative tracking. Nevertheless, optically tracked transducers [4,5] can provide a high reconstruction accuracy required in orthopedic interventions. Additionally, this absolute tracking enables use in surgical navigation.

US registration approaches in the literature can be divided into two categories: *Image-based* and *surface-based* approaches. *Image-based* approaches (e.g., [7–9]) directly register pre-processed US slices to pre-processed CT slices using a similarity metric between 2D images such as the *Linear Correlation of Linear Combinations* (LC^2) [9]. While image-based registration has the advantage that no bone surface segmentation is necessary, it can take a long time and is highly susceptible to noise in the images [10].

In *surface-based* registration, the anatomical structure (e.g., bone surface) is extracted from both modalities, and registration is performed on the two surfaces [11,12]. A popular point cloud registration algorithm for rigid transformations is Iterative Closest Point (ICP) [13]. This was used for a joint bone detection–registration of the femur between US and CT [14], in combination with morphological operators and a connected components approach [15]. In [16], Gaussian Mixture Models (GMMs) were used for modeling pelvic bone surfaces [17], where the CT surface was manually cropped to a region of interest prior to registration, and the US surface extracted using 3D local phase features. Statistical methods for the registration have also been proposed, e.g., using Unscented Kalman Filters [18,19], statistical shape models and GMMs of vertebrae [10], as well as statistical shape + pose + scale models of the lumbar spine [20]. But these require many observations to build model statistics and may not generalize well to other populations or to pathologies.

A subsequent application of image- and surface-based approaches has also been proposed to take advantage of each

method at a certain registration stage. In [21], US-to-CT registration of the lumbar spine was proposed using first a rigid intensity-based registration based on mutual information similarity (using the BRAINS module [22]), followed by Coherent Point Drift (CPD) [23] for surface-based registration. The US bone surfaces were extracted by applying the method from [24] to phase-filtered US images. In [25], first surface-based registration was employed using a simplex-based minimization algorithm (Controlled Random Search), which was then fine-tuned using an intensity-based approach with a proposed Linear Correlation of Linear Combination (LC^2) similarity metric, also taking into account the soft tissue deformation due to US transducer compression. To extract bone surfaces from US, they introduced the *bone confidence localizer*, using filtering and confidence maps [26].

In this work, we propose the extraction of approximate bone surfaces from tracked freehand US images using phase-symmetry-based bone surface extraction followed by a novel, factor-graph-based outlier removal approach. This US-extracted point cloud is then registered to a 3D bone model based on a point-set alignment algorithm. We adopted a multi-initialization framework [14] and have extended it by leveraging the US-physics-based fact that the bone surfaces have single-sided visibility in the images, which is herein shown to render the method more robust. Because this study is based on volunteer data, the 3D bone model was created from MRI instead of CT to avoid harmful radiation. An overview of our proposed method is seen in Fig. 1. Our algorithm was designed focusing on long bones and was evaluated on forearm images.

Methods

3D freehand US acquisition

In our method, we assume that the 3D freehand US forearm data have been acquired in a single sweep, starting distally at the wrist and moving proximally toward the elbow. To allow for a good registration, a good US response at the wrist should be ensured. We used the surface of the radius for the registration, as it offers more shape variability along its elongation than the ulna.

Bone surface delineation

As we perform surface-based registration, we have to extract the bone surface from both, a preoperative imaging modality (e.g., CT or MRI) yielding the *model surface*, as well as from US, yielding the *US surface*. Thereafter, the two extracted surfaces are registered.

We extracted the radius from the forearm MRI volume by manual segmentation, using the Medical Imaging Interaction Toolkit (MITK), resulting in a bone *volume*. The single-voxel thick layer extracted from it by means of morphological thinning comprises the model point cloud we hereinafter denote as $\mathcal{P}_{\text{Model}}$. This step is performed offline.

Extracting bone surfaces from US is one of the main challenges in surface-based registration. A popular method is *Phase Symmetry* (PS) [27], which is based on phase congruency [28]. While being fairly fast and reliably responding to bone surfaces, the method produces many false positives, i.e., points that are not actually part of the bone surface. Morphological thinning and selection of detections farthest from the transducer (denoted by PS_{\uparrow}) was proposed to reduce false positives [29]. Due to its high speed, we first use PS_{\uparrow} to segment each frame individually, resulting in a point cloud denoted by \mathcal{P}_{US} . To remove the remaining many false positives, we added another step, described next.

As we focus on the forearm, the shape of the bone surface to be extracted can be expected to be fairly elongated. Furthermore, assuming that the US data have been acquired using a single sweep, from its distal to its proximal end, \mathcal{P}_{US} can be expected to have a clear and distinct principal direction \mathbf{v}_{pri} . Moreover, since the shape of the forearm closely follows that of the radius, the unknown principal direction of only the true positives can be assumed to be fairly similar to \mathbf{v}_{pri} . Thus, false positives can be removed by starting at a small set of *seed* points $\mathcal{P}_{\text{seed}} \subset \mathcal{P}_{\text{US}}$ that are *known* to be part of the bone surface, and then propagating that information to other points in \mathcal{P}_{US} along \mathbf{v}_{pri} . The propagation of information from point to point can be achieved using graphical models. Specifically, we use a factor graph for that purpose.

To generate $\mathcal{P}_{\text{seed}}$, the user is asked to manually label some bone surface points among the points in \mathcal{P}_{US} in 2–4 frames, which span the 3D surface. The manually labeled points within each frame are then interpolated using cubic B-Splines, and all the points in \mathcal{P}_{US} from the same frame sufficiently close to the interpolated curve are considered bone points as well and thus added to $\mathcal{P}_{\text{seed}}$. Let us define binary label l_i for point $\mathbf{p}_i \in \mathcal{P}_{\text{US}}$ to indicate that \mathbf{p}_i is regarded as being part of the bone surface or not, i.e., $l_i = 1$ if \mathbf{p}_i is on the bone surface. Hence, we have $\forall \mathbf{p}_i \in \mathcal{P}_{\text{seed}} : l_i = 1$. We now use the locations of points in \mathcal{P}_{US} to form a factor graph FG, whose purpose is the propagation of $l_i = 1$ from $\mathcal{P}_{\text{seed}}$ along the remaining bone surface points in \mathcal{P}_{US} . Pairs of points that are close enough, and whose connecting vector

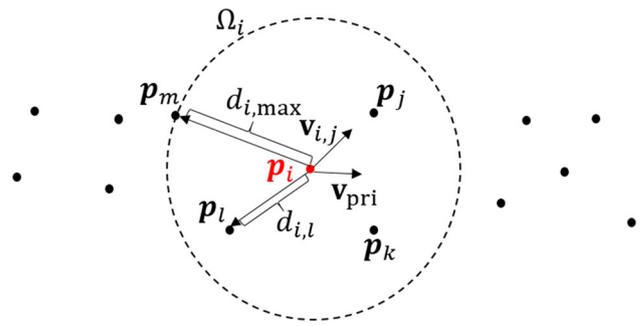


Fig. 2 Illustration of factor graph neighborhood

is sufficiently similar to \mathbf{v}_{pri} are neighbors in FG, so that they are likely to share labels after optimization. In more detail, FG is constructed as follows:

1. Compute the major extent of the long bone, \mathbf{v}_{pri} of \mathcal{P}_{US} using principal component analysis (PCA). Note that $\|\mathbf{v}_{\text{pri}}\| = 1$.
2. For each $\mathbf{p}_i \in \mathcal{P}_{\text{US}}$, find its n nearest neighbors in \mathcal{P}_{US} and denote this point set as Ω_i . For $\mathbf{p}_i \in \mathcal{P}_{\text{US}}$ and $\mathbf{p}_j \in \Omega_i$, define distances $d_{i,j} = \|\mathbf{p}_j - \mathbf{p}_i\|$, $d_{i,\max} = \max_j \{d_{i,j}\}$ and direction $\mathbf{v}_{i,j} = \frac{\mathbf{p}_j - \mathbf{p}_i}{d_{i,j}}$. See Fig. 2 for an illustration.
3. For all $\mathbf{p}_i \in \mathcal{P}_{\text{US}}$ and each $\mathbf{p}_j \in \Omega_i$, compute the following threshold:

$$t_{i,j} = |\mathbf{v}_{i,j}^T \mathbf{v}_{\text{pri}}|^{c_1} d_{i,\max} \tag{1}$$

where c_1 is constant. As $0 \leq |\mathbf{v}_{i,j}^T \mathbf{v}_{\text{pri}}| \leq 1$, threshold $t_{i,j}$ is a fraction of the maximum *neighbor* distance, based on whether the given \mathbf{p}_j is a neighbor in the principal bone direction.

4. For $\mathbf{p}_i \in \mathcal{P}_{\text{US}}$ and $\mathbf{p}_j \in \Omega_i$ add an edge (i, j) if $d_{i,j} \leq t_{i,j}$. Thus, the bigger $|\mathbf{v}_{i,j}^T \mathbf{v}_{\text{pri}}|$ (and thus the closer $\mathbf{v}_{i,j}$ is to \mathbf{v}_{pri}), the farther away \mathbf{p}_j can be from \mathbf{p}_i to still be connected to \mathbf{p}_i in FG.

We define unary potentials $\Psi(i)$ and pairwise potentials $\Psi(i, j)$ as follows:

$$\Psi(i) = \begin{cases} |l_i - 1| & \text{if } \mathbf{p}_i \in \mathcal{P}_{\text{seed}} \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

$$\Psi(i, j) = \begin{cases} \frac{d_{i,j}}{d_{i,\max}} & \text{if } l_i = l_j \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

Unary potential values for points in $\mathcal{P}_{\text{seed}}$ as defined in Eq. 2 are set such that it is unlikely that their label is changed during optimization. The pairwise potentials $\Psi(i, j)$ encode that pairs of points that are close to one another and whose connecting vector is sufficiently similar to \mathbf{v}_{pri} (which is how

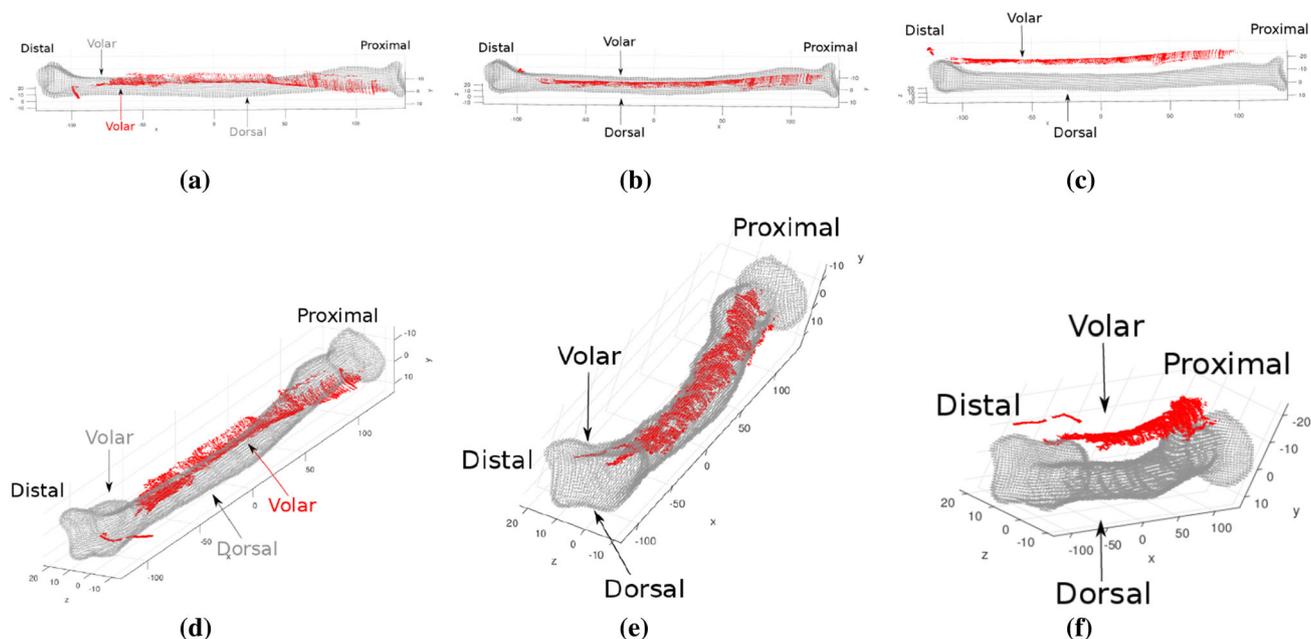


Fig. 3 $\mathcal{P}_{\text{Model}}$ (gray) and \mathcal{P}_{FG} (red) after various stages of our registration framework: **a, d** centered and aligned; **b, e** surfaces approximately correctly rotated around \mathbf{v}_{pri} ; **c, f** rotated as in **b, e** and translated before registration

we establish the neighborhood relations between the points in the first place) should have the same labels. The cost function of the factor graph is given by

$$C(\text{FG}) = \sum_{i=1}^{|\mathcal{P}_{\text{US}}|} \Psi(i) + \mu \sum_{i=1}^{|\mathcal{P}_{\text{US}}|} \sum_{j:p_j \in \Omega_i, d_i, j \leq t_{i,j}} \Psi(i, j) \quad (3)$$

where μ is the pairwise weighting constant. Eq. 3 is minimized using Tree-Reweighted Message Passing (TRW-S), and the resulting labels determine the extracted bone surface \mathcal{P}_{FG} .

Robust, multi-initialization registration scheme

Registration algorithms often suffer from local minima, leading to poor registration results if the two input surfaces are poorly aligned. We thus devised a multi-initialization scheme that runs a registration algorithm from different initial alignments of the two surfaces, with the aim to provide at least one which is sufficiently close to enable a good final result. For registration, we considered three algorithms: point-to-plane ICP [30], CPD [23], and the GMM-based algorithm as introduced in [17], which we will hereinafter denote as GMM. We found that, depending on which algorithm is used within our framework, different steps have to be taken to achieve good results within a reasonable time frame. In this section, we focus on registration using ICP, which is fully automatic and robust.

PCA Alignment. Long bones have a fairly straight and elongated surface (see Fig. 3). We thus obtain a first, crude alignment of \mathcal{P}_{FG} and $\mathcal{P}_{\text{Model}}$ by centering the two surfaces, and aligning their main principal axes extracted via PCA. Let us denote the direction of those aligned, principal axes as \mathbf{v}_{pri} . An example alignment using these steps can be seen in Fig. 3a, d.

Multi-initialization. After aligning the two surfaces with \mathbf{v}_{pri} , their individual rotation around \mathbf{v}_{pri} is arbitrary. If these individual angles are too far apart, a registration algorithm may converge to a poor alignment. To avoid this, we use a multi-initialization scheme, in which we run a registration algorithm after rotating \mathcal{P}_{FG} around \mathbf{v}_{pri} by $\alpha = i2\pi/n_{\text{rot}}$ for $i = 1, 2, \dots, n_{\text{rot}}$. Among these n_{rot} registration results \mathbf{T}_α , the one with the lowest surface registration error (SRE) E is chosen, i.e., $E(\mathcal{A}, \mathcal{B}) = \frac{1}{|\mathcal{A}|} \sum_a^{|\mathcal{A}|} \min_b \{d_{ab}\}$ for $\mathbf{p}_a \in \mathcal{A}$, $\mathbf{p}_b \in \mathcal{B}$ and $\mathbf{T} = \arg \min_{\mathbf{T}_\alpha} E(\mathbf{T}_\alpha(\mathcal{P}_{\text{FG}}), \mathcal{P}_{\text{Model}})$. An example alignment resulting from such n_{rot} iterations can be seen in Fig. 3b, e, where the rotation around \mathbf{v}_{pri} of the two surfaces is approximately correct.

Data-based translation. In US, only the part of the bone surface facing the transducer is visible. Because ICP establishes *hard* point correspondences between the two surfaces on a nearest-neighbor basis, ICP may snap \mathcal{P}_{FG} to the part of $\mathcal{P}_{\text{Model}}$ which is invisible in US. To handle this, we translate \mathcal{P}_{FG} , prior to running the multiple-initialization scheme described above, as follows:

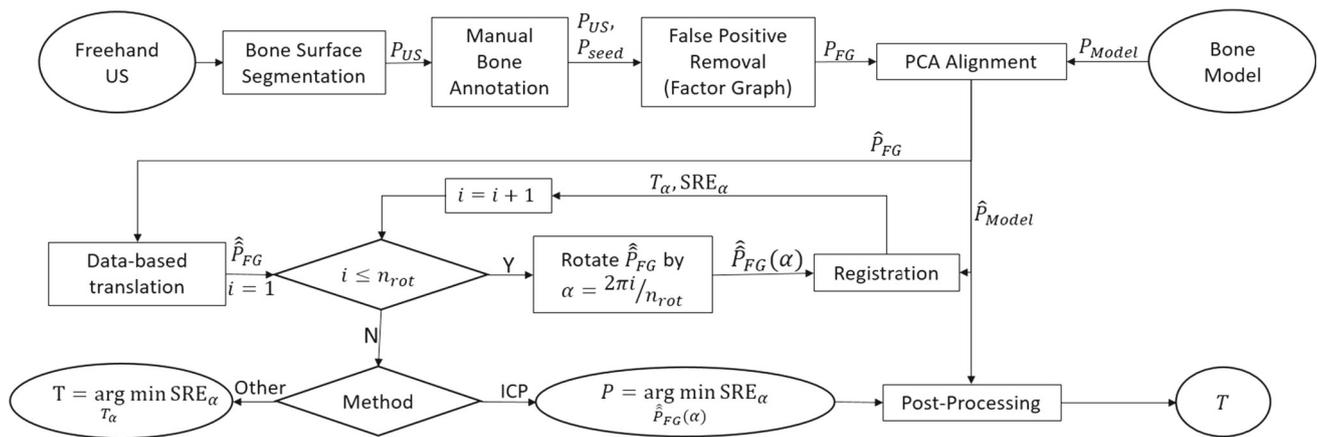


Fig. 4 Registration pipeline overview

1. Leveraging the information from US physics, translate \mathcal{P}_{FG} in the *opposite* of the average direction of the US scanlines for a specific distance d_{up} . This moves \mathcal{P}_{FG} *toward* the location of the US transducer during imaging, and for correct rotations of \mathcal{P}_{FG} around \mathbf{v}_{pri} this brings \mathcal{P}_{FG} closer to a part of \mathcal{P}_{Model} which is visible on US.
2. As stated in the “3D freehand US Acquisition” section, a good US response at the wrist should be ensured. Thus, the most distal point of \mathcal{P}_{FG} is likely to be near the wrist. As a consequence, we align the two distal ends of the two surfaces by translating \mathcal{P}_{FG} along \mathbf{v}_{pri} , which can improve the convergence rate of ICP.

- S2. Post-processing: Starting at the alignment as computed in S1, repeat ICP, while gradually lowering r in between repetitions.

Stop when the Hausdorff distance between the two surfaces after the *refined* alignment is below a specified threshold d_{in} .

Re-running ICP with a gradually smaller inlier ratio r in S2 will start to ignore points from \mathcal{P}_{FG} which are farther away from \mathcal{P}_{Model} and hence are likely false positives, thereby improving the registration.

Figure 4 shows an overview of the entire pipeline.

The initial alignment for the selected registration result, including the translation of \mathcal{P}_{FG} , can be seen in Fig. 3c, f.

Point Cloud Registration. A major issue in any registration algorithm is the presence of noise in the point clouds. While CPD and GMM are both designed to handle substantial amounts of noise and outliers (e.g., false positives), ICP is not. Hence, even small numbers of false positives can lead to suboptimal alignments with ICP if left unhandled. However, if \mathcal{P}_{FG} is sufficiently clean, ICP with multiple initializations as described above will provide a reasonably good *first* alignment. As the true positives (i.e., points correctly identified as bone surface points) outnumber the false positive in \mathcal{P}_{FG} , they dominate the registration, so that they end up in good alignment with \mathcal{P}_{Model} and consequently have a smaller distance to \mathcal{P}_{Model} than the false positives. Based on this observation, registration using ICP is performed in two stages:

- S1. Run ICP (inside the multi-initialization approach, see above) using a high *inlier ratio* $0 < r \leq 1$, such that only the point correspondences with a nearest-neighbor distance below the r th percentile will be considered for the least-squares estimation of the aligning transformation.

Results and discussion

Dataset

We acquired 36 tracked freehand in-vivo US-image sequences of the left radius from two volunteers. For acquisitions, an L14-5 linear-array US transducer was used with Sonix Touch and Tablet machines (Ultrasonix Medical Corporation, Richmond, BC, Canada). Imaging depths varied between 40 and 50 mm, at an imaging frequency of 6.66 or 10 MHz. The recording time of the sequences was within [19, 33] s, with a mean of 28 s; at a frame rate of [10.0, 23.2] fps, with a mean of 13.7 fps. Twenty sequences were recorded from a volar aspect, and 16 from a dorsal aspect. With these differing parameters, we aimed to capture a wide range of imaging settings and options for testing the generalization of our methods. US frames had an isotropic pixel resolution in the range of 65 and 81 μm . For the optical tracking of the US transducer, we used the Atracsys Easytrack 500. The calibration of the tracker was performed using the PLUS

Table 1 Parameter settings during grid searches

Grid search	$\mathcal{P}_{\text{Model}}$	N	n_{rot}	d_{up} (mm)	n	c_1	w	s
ICP factor graph	Full	All	18	15	10–500	2.5–20	–	–
ICP multi-init.	Full	All	1–18	0–20	100	10	–	–
CPD	Partial	100–2000	1	0	100	10	0–0.99	–
GMM	Partial	100–2000	18	15	100	10	–	0.01–10

toolkit [31]. US images were acquired using a custom C++ script running on the US machine.

The model surfaces were created using the respective volunteer’s water saturated Dixon MR sequence having an isotropic spatial resolution of 0.9 mm. The bone surface was extracted by manually annotating every third axial in-plane slice and converting these into a mesh through geometric 3D interpolation via MITK. Quality of the segmentations was ensured by visual inspection.

Evaluation metrics

Our goal was to assess the quality of registration by means of bone surface points that were not used in the registration process. To this end, for each recorded US-image sequence, we manually labeled the bone surface in 50 frames distributed across the entire elongation of the bone, which serve as the gold standard bone surface points in US, \mathcal{P}_{GS} . For this labeling, we used the Stradwin software.¹ After performing a registration, we align \mathcal{P}_{GS} with $\mathcal{P}_{\text{Model}}$ using the resulting transformation \mathbf{T} and refer to $E(\mathbf{T}(\mathcal{P}_{\text{GS}}), \mathcal{P}_{\text{Model}})$ by SRE_{GS} . Denoting the nearest-neighbor distances between the surface points $\mathbf{p}_a \in \mathbf{T}(\mathcal{P}_{\text{GS}})$ and $\mathbf{p}_b \in \mathcal{P}_{\text{Model}}$ as $d_{a,b}$, we also determined the 95%ile and maximum of $\{d_{a,b}\}$, referred below as $d_{\text{GS},95}$ and $d_{\text{GS},\text{max}}$, respectively.

Finally, we measured the registration time, where we *exclude* the US sequence segmentation time and the labeling time for the bone seed points, as they are identical for all considered registration methods. The mean segmentation time of the in-vivo US sequences (250–750 slices) by PS_{\uparrow} was 76.76 s (standard deviation ± 33.34 s). Manually labeling the bone seed points typically required less than 30 seconds. The mean time for creating and optimizing the factor graph was 2.10 (± 0.14) s and 1.01 (± 0.08) s, respectively.

Method parameters

To assess the robustness of our registration framework, we conducted a grid search over different subsets of parameters, which are summarized in Table 1. To obtain lower bounds on registration performances, the search was performed using

all datasets. To measure the influence of factor graph-related parameters on ICP-based registration, we performed a search over the parameters (n, c_1) , as shown in Fig. 5a, while keeping the values of n_{rot} and d_{up} fixed at 18 and 15 mm, respectively, which were empirically determined to be viable values. In all experiments, we set $\mu = \frac{|\mathcal{P}_{\text{US}}|}{|E|}$, where $|E|$ represents the number of edges in the factor graph. This gives equal weight per unary and pairwise potentials in Eq. (3). Figure 5b shows the influence of the parameters $(n_{\text{rot}}, d_{\text{up}})$ on ICP-based registration, with n and c_1 set to 100 and 10, respectively. ICP was performed with an initial inlier ratio $r = 0.99$, which was decreased in steps of 0.01 during post-processing. Finally, we used $d_{\text{in}} = 2$ mm as stopping criteria.

For rigid transformations, CPD and GMM both offer *one* free parameter. For CPD, w denotes the weight for noise and outliers, and for GMM, s defines the bandwidth of the Gaussian kernel used for the registration. For CPD and GMM, we considered only half of the model surface that corresponds to the US imaging direction (volar or dorsal), marked as “partial” in Table 1, since taking the entire model yielded significantly lower registration accuracy for these two algorithms, with their results incomparable to ICP. Under these circumstances, we found CPD to be sufficiently robust so that multiple initializations were not necessary. In contrast, our preliminary experiments had shown that GMM performs poorly without multi-initialization; thus, we performed GMM-based registration, similarly to ICP-based registration, with $(n_{\text{rot}}, d_{\text{up}})$ set to (18, 15 mm). Due to the high time complexities of CPD and GMM, it was necessary to down-sample \mathcal{P}_{FG} and $\mathcal{P}_{\text{Model}}$ to obtain results within reasonable computational times. Assuming smooth bone surfaces, we used farthest point sampling to accurately capture the surface shape as follows: starting with a subset consisting of the two points with the maximum distance from each other, the subset is iteratively increased by greedily choosing the point that is the farthest away from the current subset. With this, both surfaces were downsampled to N number of points. Thus, the grid searches for GMM and CPD were performed over N , as well as their respective free parameter, as shown in Fig. 5c, d.

To reduce the time and memory requirements of the registrations, \mathcal{P}_{US} was randomly downsampled to 90,000 points, while the model surfaces of the two subjects contained 23,320 and 20,435 points, respectively.

¹ <http://mi.eng.cam.ac.uk/~rwp/stradwin/>.

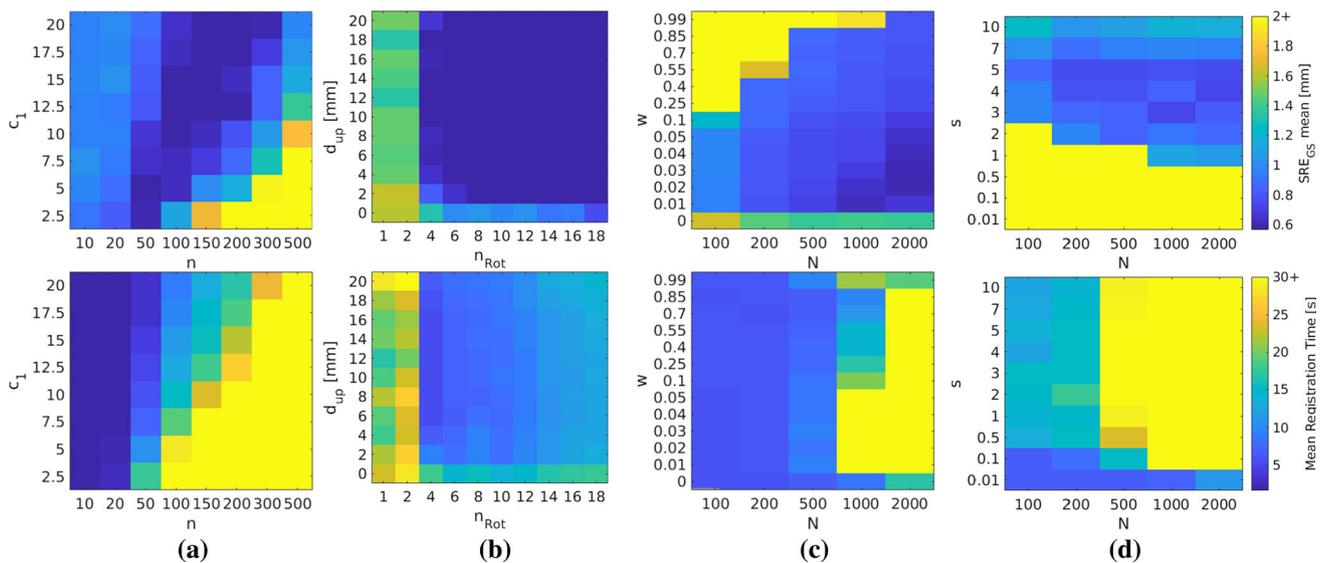


Fig. 5 Mean surface registration errors (top row) and registration times (bottom row) for varying parametrizations of **a** factor-graph outlier removal, applied prior to ICP, **b** ICP multi-initialization stage, **c** CPD, and **d** GMM. It is seen that ICP exhibits large regions that are robust to parametrization.

Table 2 Registration framework performance (mean ± standard deviation), where bold numbers indicate best results for each metric

Algorithm	SRE _{GS} (mm)	$d_{GS,95}$ (mm)	$d_{GS,max}$ (mm)	Time (s)
ICP	0.57 ± 0.08	1.09 ± 0.27	1.87 ± 0.60	15.00 ± 5.24
CPD ($N = 2000, w = 0.03$)	0.60 ± 0.1	1.24 ± 0.45	2.58 ± 1.78	226.89 ± 118.11
CPD ($N = 500, w = 0.01$)	0.71 ± 0.23	1.76 ± 1.22	4.01 ± 2.44	9.39 ± 3.14
GMM ($N = 1000, s = 3$)	0.73 ± 0.25	1.96 ± 1.67	4.46 ± 4.08	89.87 ± 51.49

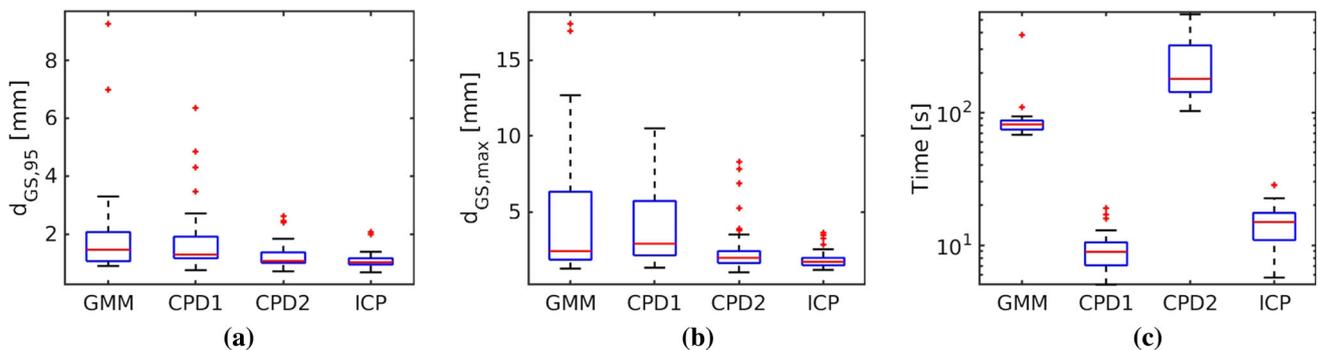


Fig. 6 Box plots showing the registration error (**a, b**) and run-time distributions (**c**) for parameters selected via grid search (CPD1: $N = 500, w = 0.01$, CPD2: $N = 2000, w = 0.03$). **a** $d_{GS,95}$. **b** $d_{GS,max}$. **c** Registration time

All experiments were conducted on a computer with an Intel i7-2600K CPU @ 3.40 GHz and 16 GB RAM. The registration pipeline was implemented in MATLAB (r2017a). We used the open-source implementation by Myronenko² [23] for CPD and the implementation by Jian³ for GMM [17]. The factor graph was implemented using OpenGM.⁴

Results

Figure 5 shows the results of the grid searches for the ICP-, GMM- and CPD-based registration. Table 2 states the performances of the ICP-based registration, along with selected configurations of the CPD- and GMM-based registration, while Fig. 6 shows registration error and time distributions of the same configurations. For CPD, the configurations resulting in the lowest mean SRE_{GS} for 2000 (best performance) and 500 points (best performance while being

² <https://sites.google.com/site/myronenko/research/cpd>.

³ <https://github.com/bing-jian/gmmreg>.

⁴ <http://hciweb2.iwr.uni-heidelberg.de/opengm/>.

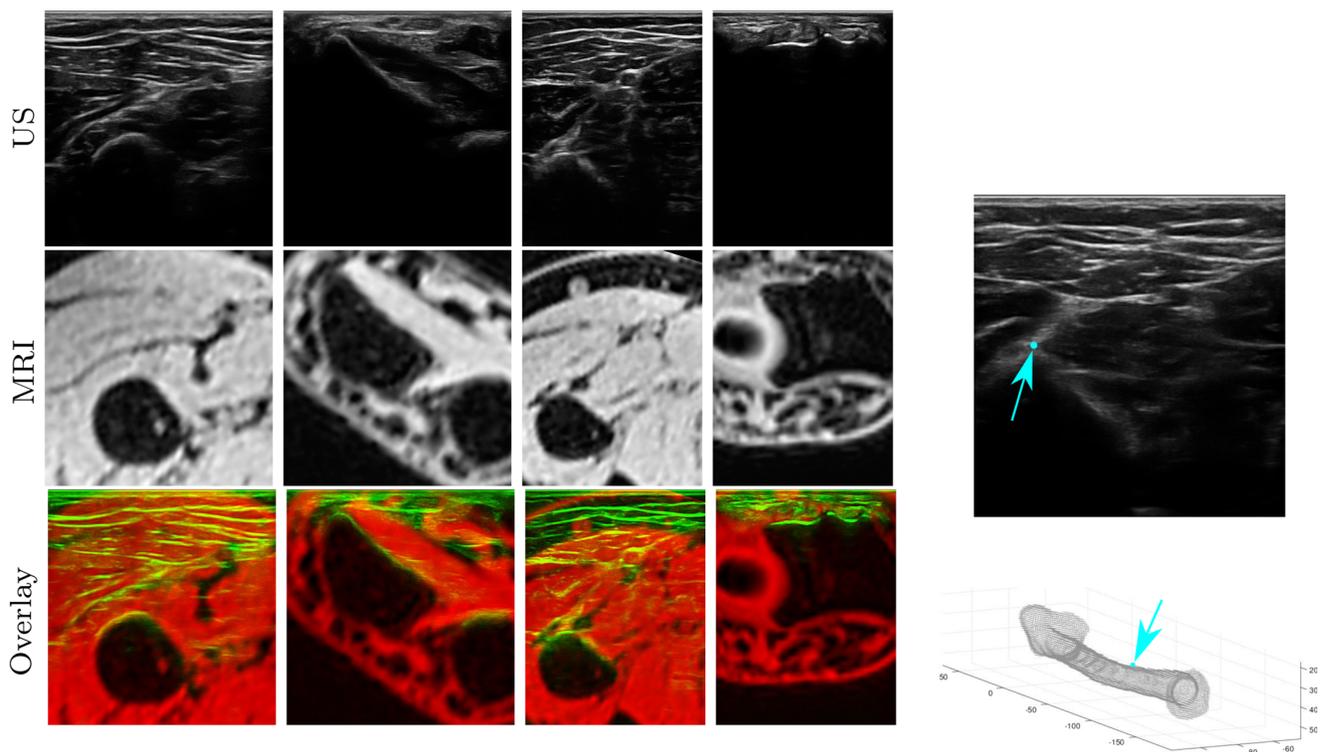


Fig. 7 (left) Example registration results with US, MRI, and overlaid slices of corresponding locations with the proposed ICP-based alignment. (right) A tendon insertion (pronator teres) visible in US (top) is

projected onto the 3D model (below) using the proposed ICP-based alignment, e.g., to facilitate preoperative planning

faster than ICP) are shown, while for GMM, the *one* configuration resulting in the lowest mean SRE_{GS} is shown. As can be seen, the best GMM result is roughly 20% worse than both, ICP and CPD. While ICP and CPD yielded similar registration accuracies, it was necessary to manually reduce the 3D bone model to the aspect imaged in US to get good alignments using CPD. Also, CPD approaches the ICP registration accuracy only with less downsampling, which drastically increases run-time. The best CPD performance with a registration time *below* that of ICP was obtained with downsampling to 500 points, which leads to a roughly 20% reduction in mean accuracy. More striking differences in terms of registration accuracy can be seen in Fig. 6b, which shows that $d_{GS,max}$ for GMM and CPD2 ($N = 2000$, $w = 0.03$) can go as high as 17 and 8 mm, respectively, while the worst $d_{GS,max}$ for ICP is 3.6 mm. Given registration run-times in Fig. 6c, only ICP and CPD1 ($N = 500$, $w = 0.01$) are found suitable for near real-time performance, where their median $d_{GS,max}$ are 1.7 and 2.9 mm, cf. Fig. 6b, indicating substantial improvements of ICP over the other two registration methods. Indeed, a registration error of 2 mm was considered sufficiently accurate in [10]. Thus, from all the investigated registration algorithms, ICP was the only one providing good registration accuracy within a reasonable time frame [32].

Additional initializations with axial translations of the US bone surface points along the long axis of the model surface may improve the results. However, we found such improvement to be minor and dependent on parametrization, and thus we excluded this computationally-costly step from our method. Results were not sensitive to the manual annotations, as mean SRE_{GS} varied by less than 0.01 mm when results from two independent annotators were compared on a subset of 15 datasets.

We are confident that the obtained accuracies would have been even better, if we had access to forearm CT scans as opposed to MRI, as CT has typically a higher resolution and allows for a more reliable, automated bone surface extraction.

Example qualitative results obtained using ICP are shown in Fig. 7. It can be observed that lower parts of the registered MRI and US slices exhibit excellent correspondence, while doing less so in upper parts. This is due to pressure applied by the US transducer on the imaged forearm, which mainly deforms soft tissue structures near the transducer surface. Fig. 8 shows a qualitative comparison between ICP- and GMM- or CPD-based registrations. The results are seen to differ substantially, where large misalignments of anatomical structures are visible for GMM and CPD.

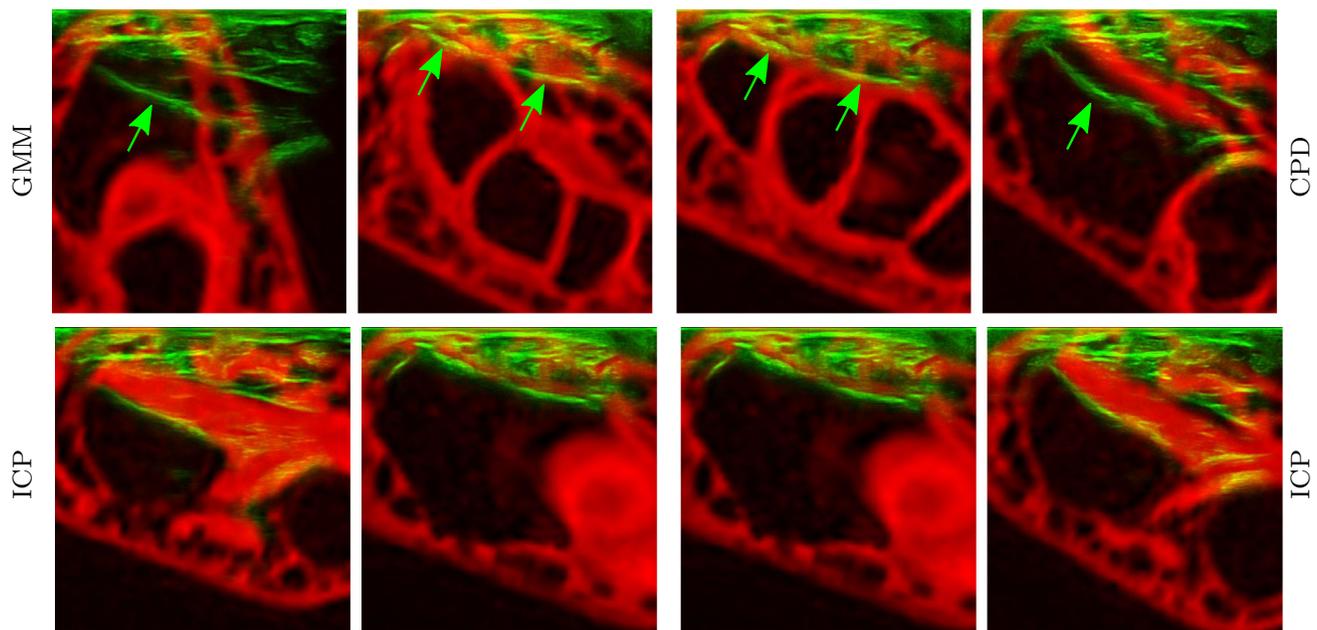


Fig. 8 Illustrations of differences between methods via overlays of registered images. Top: GMM and CPD ($N = 2000$, $w = 0.03$). Bottom: ICP for the corresponding US slices. Green arrows indicate the US response of the radius

Conclusions

We have proposed a novel, robust registration pipeline from 3D freehand US to a 3D bone model. To extract large parts of the imaged bone surface from US, we use a factor graph-based approach that utilizes the expected shape of the imaged bone, along with few manual initializations by the user. We then register the two surfaces by means of a multiple-initialization approach, that utilizes the physics of US to find a good initial alignment between them, thereby avoiding local minima registration algorithms typically suffer from. We evaluated our pipeline on 36 US sequences of in-vivo data of the radius from two volunteers, comparing three registration algorithms: point-to-plane ICP, CPD and GMM. We obtained a gold standard SRE (mean \pm standard deviation) of 0.57 ± 0.08 mm using ICP, which outperforms registration using CPD or GMM, particularly with respect to robustness. Furthermore, a qualitative evaluation of our registration confirmed excellent results. We therefore believe that the proposed robust registration framework can be used successfully in the clinical setting. The alignment of US images can help augment 3D bone models with information available only from US [e.g., the tendon insertion shown in Fig.7 (right)], e.g., for preoperative planning. Additionally, the fast runtime would enable near real-time application of the method, such as for surgical guidance of osteotomies and excision of soft tissue sarcomas.

Acknowledgements This work was funded by the Swiss National Science Foundation (SNSF) and a Highly Specialized Medicine (HSM2) grant of the Canton of Zurich.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the provincial ethics committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Court-Brown CM, Caesar B (2006) Epidemiology of adult fractures: a review. *Injury* 37(8):691–697
2. Nellans KW, Kowalski E, Chung KC (2012) The epidemiology of distal radius fractures. *Hand Clin* 28(2):113–125
3. Chung KC, Spilson SV (2001) The frequency and epidemiology of hand and forearm fractures in the United States. *J Hand Surg* 26(5):908–915
4. Rohling R, Gee A, Berman L (1999) A comparison of free-hand three-dimensional ultrasound reconstruction techniques. *Med Image Anal* 3(4):339–359
5. Flach B, Makhinya M, Goksel O (2016) PURE: panoramic ultrasound reconstruction by seamless stitching of volumes. In: MICCAI W SASHIMI, pp 75–84

6. Prager RW, Gee AH, Treece GM, Cash CJ, Berman LH (2003) Sensorless freehand 3-D ultrasound using regression of the echo intensity. *Ultrasound Med Biol* 29(3):437–446
7. Penney GP, Barratt DC, Chan CSK, Slomczykowski M, Carter TJ, Edwards PJ, Hawkes DJ (2005) Cadaver validation of intensity-based ultrasound to CT registration. In: MICCAI, pp 1000–1007
8. Gill S, Mousavi P, Fichtinger G, Chen E, Boisvert J, Pichora D, Abolmaesumi P (2009) Biomechanically constrained groupwise US to CT registration of the lumbar spine. In: MICCAI, pp 803–810
9. Wein W, Khamene A, Clevert D, Kutter O, Navab N (2007) Simulation & fully auto-matic multimodal registration of medical ultrasound. In: MICCAI, pp 136–143
10. Hacıhaliloglu I, Rasoulian A, Rohling RN, Abolmaesumi P (2014) Local phase tensor features for 3-D ultrasound to statistical shape + pose spine model registration. *IEEE TMI* 33:2167–2179
11. Taqee F, Goksel O, Mahdavi SS, Keyes M, Morris WJ, Spadinger I, Salcudean S (2012) Deformable prostate registration from MR and TRUS images using surface error driven FEM models. In: SPIE medical imaging, vol. 8316
12. Khallaghi S, Sánchez CA, Sun J, Imani F, Khale AKG, Goksel O, Rasoulian A, Romagnoli C, Abdi H, Chang S, Mousavi P, Fenster A, Ward A, Fels S, Abolmaesumi P (2015) Biomechanically constrained surface registration: application to MR-TRUS fusion for prostate interventions. *IEEE TMI* 34(11):2404–14
13. Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. *IEEE TPAMI* 14(2):239–256
14. Beitzel J, Ahmadi S, Karamalis A, Wein W, Navab N (2012) Ultrasound bone detection using patient-specific CT prior. 2664–2667
15. Kowal J, Amstutz C, Langlotz F, Talib H, Ballester MG (2007) Automated bone contour detection in ultrasound B-mode images for minimally invasive registration in computer-assisted surgery in vitro evaluation. *IJCARS* 3(4):341–348
16. Brounstein A, Hacıhaliloglu I, Guy P, Hodgson A, Abugharbieh R (2011) Towards real-time 3D US to CT bone image registration using phase and curvature feature based GMM matching. In: MICCAI, pp 235–242
17. Jian B, Vemuri BC (2005) A robust algorithm for point set registration using mixture of Gaussians. In: ICCV, pp 1246–1251
18. Moghari MH, Abolmaesumi P (2007) Point-based rigid-body registration using an unscented Kalman filter. *IEEE TMI* 26(12):1708–1728
19. Echeverría R, Cortes C, Bertelsen A, Macia I, Ruiz Ó E, Flórez J (2016) Robust CT to US 3D–3D registration by using principal component analysis and Kalman filtering. In: Computational methods and clinical applications for spine imaging, pp 52–63
20. Behnami D, Sedghi A, Anas EMA, Rasoulian A, Seitel A, Lessoway V, Ungi T, Yen D, Osborn J, Mousavi P, Rohling R, Abolmaesumi P (2017) Model-based registration of preprocedure MR and intra-procedure US of the lumbar spine. *IJCARS* 12(6):973–982
21. Nagpal S, Abolmaesumi P, Rasoulian A, Ungi T, Hacıhaliloglu I, Osborn J, Borschneck DP, Lessoway VA, Rohling RN, Mousavi P (2014) CT to US registration of the lumbar spine: a clinical feasibility study. In: IPCAI, pp 108–117
22. Johnson H, Harris G, Williams K (2007) BRAINSFit: mutual information registrations of whole-brain 3D images, using the insight toolkit. *Insight J* 57(1)
23. Song X, Myronenko A (2010) Point set registration: coherent point drift. *IEEE TPAMI* 32:2262–2275
24. Foroughi P, Boctor E, Swartz MJ, Taylor RH, Fichtinger G (2007) P6D-2 ultrasound bone segmentation using dynamic programming. In: IEEE ultrasonics symposium, pp 2523–2526
25. Wein W, Karamalis A, Baumgartner A, Navab N (2015) Automatic bone detection and soft tissue aware ultrasound-CT registration for computer-aided orthopedic surgery. *IJCARS* 10(6):971–979
26. Karamalis A, Wein W, Klein T, Navab N (2012) Ultrasound confidence maps using random walks. *Med Image Anal* 16(6):1101–1112
27. Hacıhaliloglu I, Abugharbieh R, Hodgson A, Rohling RN (2009) Bone surface localization in ultrasound using image phase-based features. *Ultrasound Med Biol* 35(9):1475–1487
28. Kovessi P (1999) Image features from phase congruency. *Videre J Comput Vis Res* 1(3):1–26
29. Ozdemir F, Ozkan E, Goksel O (2016) Graphical modeling of ultrasound propagation in tissue for automatic bone segmentation. In: MICCAI, pp 256–264
30. Chen Y, Medioni G (1992) Object modelling by registration of multiple range images. *Image Vis Comput* 10(3):145–155
31. Lasso A, Heffter T, Rankin A, Pinter C, Ungi T, Fichtinger G (2014) Plus: open-source toolkit for ultrasound-guided intervention systems. *IEEE TBME* 61(10):2527–2537
32. Schumann S (2016) State of the art of ultrasound-based registration in computer assisted orthopedic interventions. In: Computational radiology for orthopaedic interventions. Springer, pp 271–297