Automatic Annotation of Hip Anatomy in Fluoroscopy for Robust and Efficient 2D/3D Registration

Robert B. Grupp¹ · Mathias Unberath¹ · Cong Gao¹ · Rachel A. Hegeman² · Ryan J. Murphy³ · Clayton P. Alexander⁴ · Yoshito Otake⁵ · Benjamin A. McArthur⁶,⁷ · Mehran Armand²,⁴ · Russell H. Taylor¹

Received: date / Accepted: date

Abstract

Purpose Fluoroscopy is the standard imaging modality used to guide hip surgery and is therefore a natural sensor for computer-assisted navigation. In order to efficiently solve the complex registration problems presented during navigation, human-assisted annotations of the intraoperative image are typically required. This manual initialization interferes with the surgical workflow and diminishes any advantages gained from navigation. In this paper we propose a method for fully automatic registration using anatomical annotations produced by a neural network.

Methods Neural networks are trained to simultaneously segment anatomy and identify landmarks in fluoroscopy. Training data is obtained using a computationally-intensive, intraoperatively incompatible, 2D/3D registration of the pelvis and each femur. Ground truth 2D segmentation labels and anatomical landmark locations are established using projected 3D annotations. Intraoperative registration couples a traditional intensity-based strategy with annotations inferred by the network and requires no human assistance.

Results Ground truth segmentation labels and anatomical landmarks were obtained in 366 fluoroscopic images across 6 cadaveric specimens. In a leave-one-subject-out experiment, networks trained on this data obtained mean dice coefficients for left and right hemipelves, left and right femurs of 0.86, 0.87, 0.90, and 0.84, respectively. The mean 2D landmark localization error was 5.0 mm. The pelvis was registered within 1° for 86% of the images when using the proposed intraoperative approach with an average runtime of 7 seconds. In comparison, an intensity-only approach without manual initialization, registered the pelvis to 1° in 18% of images.

Conclusions We have created the first accurately annotated, non-synthetic, dataset of hip fluoroscopy. By using these annotations as training data for neural networks, state of the art performance in fluoroscopic segmentation and landmark localization was achieved. Integrating these annotations allows for a robust, fully automatic, and efficient intraoperative registration during fluoroscopic navigation of the hip.

Keywords Landmark Detection · Semantic Segmentation · 2D/3D Registration · X-ray Navigation · Orthopaedics

1 Introduction

Minimally invasive surgical interventions of the hip manipulate, modify, or augment anatomical structures which are hidden or not reliably visible [1]. Clinicians commonly use intraoperative fluoroscopy in order to overcome this occlusion and ascertain the poses of anatomy, surgical instruments, or artificial implants. However, mental interpretation of these images is a difficult task and subject to an extensive learning curve [2,3]. Computer-assisted navigation systems ease this burden by track-
ing relevant objects and reporting their poses within the context of a surgical plan or scenario. Systems leveraging fluoroscopy have been developed for total hip arthroplasty [4], hip resurfacing [5], cement injection [6], and osteotomies of the acetabulum [7] or proximal femur [8].

In order to report object poses accurately, fluoroscopic navigation systems rely on 2D/3D registrations between intraoperative 2D images and the appropriate 3D models [9]. Large errors in pose estimates may occur when a registration is not initialized sufficiently close to the actual pose of an object. Quality initializations are derived from some manual human input, often through annotated landmark locations in the fluoroscopic image. Although these systems report favorable navigation-related results, the manual initialization of processing may interrupt surgical workflows and negatively affect patient outcomes due to increased operating time or blood loss.

Convolutional neural networks (CNNs) have excelled at detecting landmarks and performing semantic segmentation when sufficiently large annotated datasets are available for training [10, 11]. However, since existing large-scale hip datasets have focused on 3D image modalities and pre and postoperative radiography, rather than intraoperative fluoroscopy [12], applications of CNNs to fluoroscopy have been mostly limited to recognizing surgical instruments and tools [13–16].

Several authors have coupled image segmentation with landmark estimation using multi-task networks and achieved favorable results. By reusing segmentation features from an encoder-decoder style network for the computation of landmark heatmaps, Laina was able to automatically annotate segmentation labels and landmark locations of tools used in laparoscopy and retinal microsurgery [17]. Gao also leveraged this approach for the localization of a dextrous continuum manipulator in fluoroscopy [16]. Kordon demonstrated that a CNN, trained from 149 manually annotated preoperative radiographs, could successfully segment the four bones of the knee joint, and locate two anatomical landmarks and a surgically relevant line [18].

Using a large collection of simulated fluoroscopy, Bier trained CNNs to annotate anatomical landmarks of the pelvis [19]. When evaluated on five sequences of actual fluoroscopy across two cadaveric specimens, mean annotation errors of 12-24 mm in the detector plane were reported. Pelvis poses were estimated using these annotations, yielding reprojection errors of 14-34 mm for other landmarks not learned by the network. Their work was extended in [20], whereby each network was fine-tuned on simulated fluoroscopy for a specific patient of interest. The approach was evaluated by estimating landmark locations in previously unseen simulated images, and using these estimates to produce quality initializations for 2D/3D registration. No analysis on actual fluoroscopy was conducted in [20].

In this paper, we propose a method for 2D/3D registration of hip anatomy that simultaneously combines image intensities with higher-level landmark and segmentation features, making it robust against large initial offsets from actual object poses. Segmentation labels and landmark annotations are produced by a CNN similar in architecture to those found in [16] and [17]. Contrary to [16, 19, 20], we train our networks using smaller datasets of actual fluoroscopy and achieve state of the art results on clinically relevant data. Annotated fluoroscopy for training is semi-automatically obtained using a computationally expensive 2D/3D registration, with runtimes on the order of several minutes per image. The intraoperative registration is fully automatic, requires no initialization from a user, and completes in an order of seconds.

2 Methods

We now describe the details of the data preprocessing, the methods for creating an annotated, training, dataset of hip fluoroscopy, the CNN architecture, and the intraoperative registration strategy. The reader is referred to the supplementary document for details regarding the lower-level parameters used for the registration pipelines and network training.

2.1 Data Preprocessing

Using the procedure described in [7], lower torso 3D CT scans are resampled to have 1 mm isostropic spacing. Segmentations of the pelvis, femurs, and vertebrae are obtained semi-automatically. A total of 14 landmarks are manually annotated in 3D: the left and right (L./R.) centers of the femoral head (FH), L./R. greater sciatic notches (GSN), L./R. inferior obturator foramen (IOF), L./R. medial obturator foramen (MOF), L./R. superior pubis symphysis (SPS), L./R. inferior pubis symphysis (IPS), and the L./R. anterior superior iliac spine (ASIS). These landmarks were previously identified as being useful for obtaining initial registration estimates of the pelvis [7]. The anterior pelvic plane (APP) coordinate system for each specimen is defined using the L./R. ASIS and L./R. SPS landmarks [21], and is later used to estimate nominal anterior/posterior (AP) poses and as a reference coordinate frame during registration. Segmentations of each hemipelvis and sacrum are separated from the full pelvis segmentation,
Fig. 1 Three views of the 6 anatomical structures and 14 landmarks to be annotated in 2D fluoroscopy. All landmarks are bilateral with left (L.) and right (R.) denoted. The L. hemipelvis is shown in green, the R. hemipelvis in red, L. femur in cyan, R. femur in orange, vertebrae in blue, and upper sacrum in yellow. Each landmark is overlaid as a purple sphere.

and any sacrum labels inferior to the sacroiliac joint are discarded. Fig. 1 shows an example 3D visualization of the individual bone surfaces and the anatomical landmarks.

Fluoroscopy is collected with a Siemens CIOS Fusion mobile C-arm with 30 cm detector. Images are 1536 × 1536 pixels with 0.194 mm pixel spacings. Each image is cropped by 50 pixels along each border to remove collimator artifacts and intensity values are log-corrected (“bone is bright”).

2.2 2D/3D Registration

Our approach to 2D/3D registration of single-view fluoroscopy and CT uses the multiple-resolution, multiple-component, 2D/3D, intensity-based registration pipeline introduced in [7]. The registration problem of finding the rigid poses of the pelvis \((\theta_P)\), left femur \((\theta_{LF})\), and right femur \((\theta_{RF})\) with respect to a single fluoroscopic view, \(I\), is defined by the optimization problem (1), where \(\mathcal{P}\) indicates a projection operator creating digitally reconstructed radiographs (DRRs), \(\mathcal{S}\) indicates a similarity measure between DRRs and fluoroscopy, \(\mathcal{R}\) is a regularization penalizing implausible poses, and \(\lambda \in [0,1]\) is a tuning parameter.

\[
\min_{\theta_P, \theta_{LF}, \theta_{RF}} \lambda \mathcal{S} (\mathcal{P}(\theta_P, \theta_{LF}, \theta_{RF}), I) + (1 - \lambda) \mathcal{R}(\theta_P, \theta_{LF}, \theta_{RF})
\]

In this paper, \(\mathcal{S}\) is defined as the weighted sum of normalized cross-correlations of 2D image gradients computed over image patches [22]. For all registrations using regularization, \(\lambda = 0.9\).

2.3 Training Dataset Creation

The training dataset of annotated fluoroscopy images is constructed using an automatic 2D/3D registration of the pelvis and both femurs. Once anatomy is registered to each image, the 3D segmentation labels and landmarks are propagated to 2D. Since this registration is performed “offline,” we use a computationally expensive combination of global search strategies, followed by several local strategies. Manual inspection is performed so that images corresponding to failed registrations are pruned from the dataset. It should be emphasized that, although this registration is automatic and global, the amount of computation precludes it from intraoperative application.

An attempt is first made to register the pelvis using a mixture of the Differential Evolution [23], exhaustive grid search, Particle Swarm [24], Covariance Matrix Adaptation: Evolutionary Search (CMA-ES) [25], and Bounded Optimization by Quadratic Approximation (BOBYQA) [26] optimization strategies at multiple resolutions. Using a combination of the CMA-ES and BOBYQA strategies, the left and right femurs are registered once the pelvis is registered. The rotation components of the left femur and right femurs are independently estimated, keeping the pelvis fixed at its current pose estimate. Next, simultaneous optimization over the rigid poses of the pelvis and both femurs is performed. For each of the preceding registrations uniform patch weightings were applied for \(\mathcal{S}\).

The 2D location of each landmark is obtained by projecting the corresponding 3D landmark onto the detector. When a landmark projects outside the detector region, it is identified as not visible in the image.

Each 2D pixel is labeled as the anatomy for which the corresponding source-to-detector ray intersects. Discrete labels are used to assign a single class of anatomy
to each pixel. Femurs are given highest precedence in labeling: any ray/femur intersection yields a label of the corresponding femur. Hemipelvis have the next highest precedence, with any rays intersecting both hemipelvis assigned a label corresponding to the hemipelvis closer to the X-ray source. Vertabrae intersections are given next precedence, followed by the upper sacrum. All remaining pixels are assigned to background.

The 2D labels and landmarks for each projection are manually inspected and verified.

2.4 Network Architecture and Training

In constructing our network, we follow the approach described by [16] and [17], appending segmentation and landmark heatmap network paths after an encoder-decoder structure. Supp. Figs. S-2, S-3, and S-4 describe the network architecture used in this work. For the encoder-decoder in this paper, we adopt a 6 level U-Net [27] design with 32 features at the top level and 1024 features at the bottom. Our implementation is fully-convolutional with learned 2x2 convolutions of stride 2 for downsampling, and transposed convolutions for upsampling.

The segmentation path follows directly from the original U-Net design. The differentiable dice score [28] is computed for each class and then averaged. This value is bounded, taking on values in [0, 1], with larger values indicating a higher quality segmentation.

Segmentation features prior to soft-max are concatenated with the features output from the encoder-decoder, and passed through two 1x1 convolutions to obtain a feature map where each channel estimates the heatmap of a landmark.

Ground truth heatmaps for each landmark location are defined by a symmetric 2D normal distribution with mean value equal to the landmark location and standard deviations of \( \sigma = 3.88 \) mm in each direction. The value of \( \sigma \) was subjectively chosen to approximate the variance found in manual landmark annotation. Each heatmap is set to be identically zero when the landmark is not visible. Examples of ground truth heatmaps are shown in the third row of Fig. 2. Heatmap loss is computed using the average normalized-cross correlation (NCC) between each ground truth heatmap and the corresponding estimate. This term is bounded, taking on values in \([-1, 1]\), with larger positive values indicating stronger correlation between ground truth heatmaps and estimated heatmaps.

By scaling and shifting the average NCC value into the range of \([0, 1]\), the heatmap loss may be weighted equally to the dice term without any additional hyper-parameter tuning. Finally, the combined dice and heatmap terms are negated (for minimization).

Networks are trained using stochastic gradient descent, with an initial learning rate of 0.1, Nesterov momentum of 0.9, weight decay of 0.0001, and a batch size of five images. Training and validation data sets are obtained by applying a random 85%/15% split to the data not used for testing. Test data sets are comprised of images collected from a single specimen, and no images derived from this specimen are present in the training and validation data. Extensive online data augmentation is applied to each image with probability 0.5. If an image is to be augmented, the intensities are randomly inverted, random noise is added to the image, the contrast is randomly adjusted, a random 2D affine warp is applied, and a random number of regions are corrupted with very large amounts of noise. Each image is normalized to have zero mean and standard deviation one before input into the network. Training is run for a maximum of 500 epochs and the learning rate is multiplied by 0.1 after validation loss plateaus. The network expects images of size \(192 \times 192\) pixels, and fluoroscopy data is downsampled \(8 \times\) in each dimension, accordingly. PyTorch 1.2 was used to implement, train, and test the networks.

2.5 Extracting Landmark Locations

Both, segmentations and heatmaps, are used to estimate anatomical landmark locations. Candidate locations of the FH landmarks are restricted to pixels labeled as the corresponding femur, and all remaining landmarks are restricted to locations labeled as the corresponding hemipelvis. Restricting candidate locations in this way avoids possible false alarms when the ipsilateral landmark is not in the view and a large heatmap intensity is located about the contralateral landmark. The final proposed location of each landmark is defined as the candidate location with maximal heatmap intensity. In order to distinguish between the cases of landmark detection, no detection, and spurious heatmap values, a \(25^2\) pixel region of the estimated heatmap, centered around the proposed location, is matched against the 2D symmetric normal distribution template of a detection at the center of the region. A detection is reported when NCC between the two regions is greater than 0.9, and no detection is reported otherwise.

2.6 Intraoperative Registration

The intraoperative registration strategies in this paper attempt to solve (1) in a similar fashion as the method used for construction of the training data set: the pelvis
Fig. 2 Example annotations of four specimens. The top row shows the ground truth segmentation labels for each object overlaid onto the fluoroscopic images, along with the landmark locations as yellow circles. The colors of each object correspond to those from Fig. 1. CNN estimates are shown in the second row, with ground truth landmark locations shown as yellow circles and estimated locations shown as yellow crosshairs (+). Missed detections are indicated by a circle without a corresponding cross. Ground truth heatmaps for the R. MOF, L. ASIS, L. GSN, and L. IOF, in (a), (b), (c), and (d), respectively, are overlaid and shown in the third row. Estimated heatmaps for these landmarks are shown in the bottom row. The heatmap shown in (b) highlights a successful no detection report for L. ASIS.

is registered first, followed by optimizations of each femur’s rotation, followed by a simultaneous optimization over the rigid poses of all objects.

Method 1: A naive approach for efficient automatic registration uses only intensity information, with uniform patch weightings and no regularization applied. The single landmark initialization described in [7] is used to calculate an initial AP pose of the pelvis, aligning the 3D centroid of the L. ASIS, R. ASIS, L. SPS, and R. SPS with the center of the image.

Method 2: However, a great deal of information about the 2D image is known, courtesy of the segmentation and landmark localizations produced by the CNN. A less naive approach uses detected landmarks to solve the PnP problem [29] and automatically initializes an intensity-based registration. The segmentation is used to apply non-uniform patch weightings in $S$, and soft-bounds are applied through a regularization on pose parameters.

Method 3: Instead of treating landmark features and intensity features separately, the detected landmark locations may be incorporated into a robust reprojection regularizer for intensity-based registration. The regularizer is defined in (2), with the $l^{th}$ landmark location in 3D is denoted by $p_{3D}^{(l)}$, and corresponding estimated 2D location, $p_{2D}^{(l)}$.

$$
\mathcal{R}(\theta_P) = \frac{1}{2\sigma^2} \sum_{i=1}^{N_L} \left\| P\left(p_{3D}^{(l)}; \theta_P\right) - p_{2D}^{(l)} \right\|_2^2
$$

As with the PnP approach, non-uniform patch weightings are applied using the segmentation. Using one of the estimated 2D landmark locations, the single landmark initialization is used to calculate an initial AP
pose of the pelvis. Contrary to most other registration strategies, this method never places a penalty on the magnitudes of the pose estimates offset from their initialization. The registration is free to minimize the image similarity term so long as known 3D landmarks project to approximately the correct location in 2D. Conversely, the registration is free to minimize landmark reprojection error, so long as the candidate poses produce images that approximately match the observed image.

Pelvis registrations first use a CMA-ES optimization, followed by the BOBYQA strategy at a finer resolution without patch weightings or regularization. When using patch weightings, patches centered at pixels labeled as either hemipelvis are given uniform weight and the remaining patches are weighted zero. Next, femur registration proceeds identically to that used during construction of the training data set, except for the case of patch weighting. When registering the individual femurs and using patch weightings, patches centered at pixels labeled as either hemipelvis or either femur are given uniform weight and all remaining patches are weighted zero.

3 Experiments and Results

3.1 Data Collection and Training Dataset Creation

Lower torso CT scans were obtained for three male and three female cadaveric specimens with median age 88 and ranging 57-94 years. Each CT scan was semi-automatically segmented and 3D landmarks were manually digitized. A total of 399 fluoroscopic images were collected at various C-arm poses. The “offline” training data set registration pipeline was run on each image and a total of 366 images were verified to have registered successfully. For each successfully registered image, manual annotations were made for each femur indicating whether enough of the bone was visible to use for future registration evaluation. These counts are also broken down across each specimen in Supp. Table S-6. Poses recovered from successful registrations in this phase were treated as ground truth during intraoperative registrations. Examples of generated 2D ground truth annotations are shown in the top row of Fig. 2. The mean total registration time per image was 4 minutes using a NVIDIA Tesla P100 (PCIe) GPU.

3.2 Segmentation and Landmark Localization

A total of six networks were trained in a leave-one-specimen-out experiment. For each network, the training and validation data consisted of all labeled images from five specimens and all labeled images from the remaining specimen were used as test data. Across all test images, mean dice coefficients of 0.86±0.20, 0.87±0.18, 0.90±0.24, 0.84±0.31, 0.74±0.19, and 0.63±0.13 were obtained for the left hemipelvis, right hemipelvis, left femur, right femur, vertebra, and upper sacrum, respectively. A listing of dice coefficients for each object of each specimen is shown in Supp. Table S-7. The average landmark 2D localization error was 5.0±5.2 mm in the detector plane. Table 1 lists the average landmark errors, false negative rates, and false positive rates for each landmark. Fig. 3 shows a plot of localization error thresholds and corresponding correct detection rates. The mean time for segmentation and landmark detection per image was 24.0±0.4 milliseconds using a NVIDIA Tesla P100 (PCIe) GPU.

3.3 Intraoperative Registration

Using estimated segmentations and landmark locations of the 366 test images, each intraoperative registration strategy was run and compared to ground truth. Registrations with pelvis rotation error less than 1° were defined as successful. This error was computed in the projective frame with center of rotation at the ground truth midpoint between the two femoral heads. Femur errors were computed using relative poses of the femur with respect to the pelvis in the APP with center of rotation at the ipsilateral femoral head. Table 2 lists the mean errors computed over all specimens and Supp. Table 8 shows the mean errors per specimen. The mean runtime for method 3 was 7.2±0.7 seconds using a
methods 2 and 3 were also successful in 62 and 64 of method 1’s successes, each with a mean femur rotation error of $0.8^\circ \pm 0.5^\circ$. This indicates that the three methods perform comparably on images for which the naive approach succeeds. Moreover, the larger errors of methods 2 and 3 in the remaining cases are in part caused by the more challenging pelvis registration problems presented in these images, for which the naive method was unable to solve.

The performance of our approach degrades as images collected with C-arm poses not found in the training data set are processed. This is highlighted in Supp. Table S-8, showing the poor performance of specimen 6 compared to all other specimens. When testing on spec. 6, average landmark localization error was 10.0 mm, with a false negative rate of 30%, and successful pelvis registration rate of 29% for both methods 2 and 3. Supp. Fig. S-6 shows a visualization of all 366 ground truth projection geometries. The geometries associated with spec. 6 are clearly collected at different C-arm poses than those used for training the network tested on spec. 6. Three examples of spec. 6 are shown in Fig. 4 (b)-(d). This limitation may be overcome by collecting more fluoroscopy data for training. However, it is conceivable that some C-arm poses encountered during testing will still be absent during training. By augmenting actual fluoroscopy with realistic synthetic fluoroscopy [30]

Table 1 Landmark detection errors across all trained networks for each landmark.

| Landmark | Error |
|----------|-------|
|          | Pixels | mm      | False Negative Rate | False Positive Rate |
| L. FH    | 1.9 ± 0.9 | 3.0 ± 1.5 | 0.02 | 0.02 |
| R. FH    | 3.2 ± 1.9 | 5.0 ± 3.0 | 0.04 | 0.01 |
| L. GSN   | 4.4 ± 2.0 | 6.8 ± 3.1 | 0.20 | 0.00 |
| R. GSN   | 4.7 ± 2.3 | 7.3 ± 3.6 | 0.14 | 0.01 |
| L. IOF   | 2.8 ± 4.0 | 4.3 ± 6.2 | 0.23 | 0.01 |
| R. IOF   | 2.3 ± 3.3 | 3.5 ± 5.1 | 0.16 | 0.02 |
| L. MOF   | 3.7 ± 3.2 | 5.8 ± 5.0 | 0.17 | 0.04 |
| R. MOF   | 3.4 ± 1.9 | 5.2 ± 3.0 | 0.17 | 0.02 |
| L. SPS   | 3.1 ± 2.4 | 4.7 ± 3.7 | 0.27 | 0.02 |
| R. SPS   | 3.7 ± 2.9 | 5.8 ± 4.5 | 0.22 | 0.01 |
| L. IPS   | 1.9 ± 1.6 | 3.0 ± 2.4 | 0.17 | 0.02 |
| R. IPS   | 1.5 ± 1.0 | 2.3 ± 1.6 | 0.15 | 0.01 |
| L. ASIS  | 9.0 ± 9.6 | 14.0 ± 14.9 | 0.29 | 0.01 |
| R. ASIS  | 3.9 ± 3.7 | 6.0 ± 5.7 | 0.14 | 0.01 |
| All      | 3.2 ± 3.4 | 5.0 ± 5.2 | 0.17 | 0.01 |

Table 2 Pelvis and femur registration errors from successful pelvis registrations using the three intraoperative approaches. Femur registrations errors are reported for all successful pelvis registrations which have sufficient visibility of a femur.

| Regi. Method | Pelvis Errors | Femur Errors |
|--------------|---------------|--------------|
| # Success | Rot. $(^\circ)$ | Trans. (mm) | # | Rot. $(^\circ)$ | Trans. (mm) |
| 1: Naive | 66 (18%) | 0.1 ± 0.1 | 0.4 ± 0.9 | 22 | 0.4 ± 0.3 | 0.4 ± 0.3 |
| 2: PnP Init. | 299 (82%) | 0.1 ± 0.2 | 1.0 ± 1.5 | 183 | 1.4 ± 3.4 | 0.5 ± 0.5 |
| 3: Combined | 313 (86%) | 0.2 ± 0.2 | 1.4 ± 2.0 | 192 | 1.5 ± 3.3 | 0.6 ± 0.7 |

NVIDIA Tesla P100 (PCIe) GPU. Examples of automatic annotation and registration with method 3 are shown in the supplementary video.[1]

4 Discussion and Conclusion

The naive intraoperative registration performed poorly, succeeding in only 18% of trials, while the methods leveraging CNN annotations succeeded over 4 times as often. Despite the fairly large false-negative detection rate of 17%, an average of 7 landmarks per image were detected, allowing methods 2 and 3 to perform well. Method 3’s performance was robust when only 2, 3, and 4 landmarks were detected; reporting success in 2, 15, and 30 cases, respectively. Fig. 4 (a) shows an image with 2 detected landmarks and was registered successfully. Highlighting the robustness gained from mixing intensity-features with landmark features, method 2 was only successful with these number of detections in 0, 7, and 26 cases, respectively. The low false positive detection rate ensured that inconsistent features would not confound the registration.

Although the naive registration only succeeded in 66 cases, the mean femur rotation errors were about 1° smaller than those of methods 2 and 3. However, a false negative rate of inconsistent features would not confound the registration.

[1] https://youtu.be/5AwGlNkcp9o
during training, we believe that quality performance at these “missing” poses may be achieved.

In conclusion, this work presents a new 2D/3D registration strategy coupling image intensities and higher-level features, making it robust against poor initializations. This robustness removes the need for manual initialization and allows navigation to be more naturally integrated into existing workflows. Finally, we believe the ground truth labeling method will translate to fluoroscopy of other anatomy and enable machine learning applications in other specialties.

Acknowledgements We thank Mr. Demetres Boston for assisting with the cadaveric data acquisition. This research was supported by NIH/NIBIB grants R01EB006839, R21EB020113, Johns Hopkins University Internal Funds, and a Johns Hopkins University Applied Physics Laboratory Graduate Student Fellowship. Part of this research project was conducted using computational resources at the Maryland Advanced Research Computing Center (MARCC).

Compliance with Ethical Standards
Conflict of Interest: The authors declare that they have no conflict of interest.
Ethics Approval: This article does not contain any studies with human participants performed by any of the authors.
Informed Consent: This article does not contain patient data.

References
1. Woerner, M., Sendtner, E., Springorum, R., Craiovian, B., Worlcek, M., Renkawitz, T., Grilka, J., Weber, M.: Visual intraoperative estimation of cup and stem position is not reliable in minimally invasive hip arthroplasty. Acta Orthop. 87(3), 225–230 (2016)
2. Slotkin, E.M., Patel, P.D., Suarez, J.C.: Accuracy of fluoroscopic guided acetabular component positioning during direct anterior total hip arthroplasty. J. Arthroplasty 30(9), 102–106 (2015)
3. Troehlen, A.: Surgical advances in periacetabular osteotomy for treatment of hip dysplasia in adults. Acta Orthop. 80(suppl32), 1–33 (2009)
4. Kelley, T.C., Swank, M.L.: Role of navigation in total hip arthroplasty. J. Bone Joint Surg.-Am Vol. 91(Supplement_1), 153–158 (2009)
5. Belei, P., Skwara, A., Puente, M.D.L., Schkommodau, E., Fuchs, S., Wirtz, D.C., Kämper, C., Radermacher, K.: Fluoroscopic navigation system for hip surface replacement. Comput. Aided Surg. 12(3), 160–167 (2007)
6. Malan, D.F., van der Walt, S.J., Raidou, R.G., van den Berg, B., Stoel, B.C., Botha, C.P., Nelissen, R.G., Valsstar, E.R.: A fluoroscopy-based planning and guidance software tool for minimally invasive hip refixation by cement injection. Int. J. Comput. Assist. Radiol. Surg. 11(2), 281–296 (2016)
7. Grupp, R.B., Hegeman, R., Murphy, R., Alexander, C., Otake, Y., McArthur, B., Armmond, M., Taylor, R.H.: Pose estimation of periacetabular osteotomy fragments with intraoperative X-ray navigation. IEEE Trans. Biomed. Eng. (2019)
8. Gottschling, H., Roth, M., Schweikard, A., Burgkart, R.: Intraoperative, fluoroscopy-based planning for complex osteotomies of the proximal femur. Int. J. Med. Robot. Comput. Assis. Surg. 1(4), 67–73 (2005)
9. Markelj, P., Tomažević, D., Likar, B., Perušič, F.: A review of 3D/2D registration methods for image-guided interventions. Med. Image Anal. 16(3), 642–661 (2012)
10. Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: Proc. Eur. Conf. Comput. Vis., pp. 483–499. Springer (2016)
11. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. (4), 640–651 (2017)
12. Otake, Y., Takao, M., Yokota, F., Fukuda, N., Uemura, K., Sugano, N., Sato, Y.: Construction and application of large-scale image database in orthopedic surgery. In: Computer Assisted Orthopaedic Surgery for Hip and Knee, pp. 191–197. Springer (2018)
13. Miao, S., Wang, Z.J., Liao, R.: A CNN regression approach for real-time 2D/3D registration. IEEE Trans. Med. Imag. 35(5), 1352–1363 (2016)
14. Ambrosini, P., Ruijters, D., Niessen, W.J., Moeler, A., van Walsum, T.: Fully automatic and real-time catheter segmentation in X-ray fluoroscopy. In: Proc. Med. Image Comput. Comput.-Assist. Interv., pp. 577–585. Springer (2017)
15. Breininger, K., Albarqouni, S., Kurzendorfer, T., Pfister, M., Kowarschik, M., Maier, A.: Intraoperative stent segmentation in X-ray fluoroscopy for endovascular aortic repair. Int. J. Comput. Assist. Radiol. Surg. 13(8), 1221–1231 (2018)
A Supplementary Methods

A.1 2D/3D Registration

The \( \mathfrak{s}(3) \) Lie algebra parameterization of \( SE(3) \) with reference point at the initial pose estimate of the object, is used during the optimization of rigid poses. The \( \mathfrak{s}(3) \) Lie algebra parameterization of \( SO(3) \) is used when only optimizing over the rotation component. Differential Evolution (DE) [23], exhaustive search, Covariance Matrix Adaptation: Evolutionary Search (CMA-ES) [25], and Bounded Optimization by Quadratic Approximation (BOBYQA) [26].

The DE optimization uses a regularizer designed to penalize poses which: do not project at least one femoral head center within the 2D image bounds, project inferior landmarks above superior landmarks in the image, or place either femoral head center behind the detector or too close to the X-ray source. This regularization is defined by \( R_{DE} \) in (3).

\[
R_{DE}(\theta_P) = 2 \left[ R_{\text{visible}}(p_{\text{left}}^{\text{FH}}; \theta_P) R_{\text{visible}}(p_{\text{right}}^{\text{FH}}; \theta_P) \right] + 2 \left[ R_{\text{depth}}(p_{\text{left}}^{\text{FH}}; \theta_P) + R_{\text{depth}}(p_{\text{right}}^{\text{FH}}; \theta_P) \right] + R_{\text{up}}(p_{\text{ASIS}}^{\text{left}}; p_{\text{IOF}}^{\text{left}}; \theta_P) + R_{\text{up}}(p_{\text{ASIS}}^{\text{right}}; p_{\text{IOF}}^{\text{right}}; \theta_P)
\]

(3)

The individual penalty applied for projecting a point outside of the field of view is defined in (4). The number of pixels, in the row direction, by which the point is projected “out-of-bounds” is indicated by \( r \) and \( c \) is the corresponding value in the column direction. Both \( r \) and \( c \) are zero-valued for projected locations within the image bounds.

\[
R_{\text{visible}}(p; \theta_P) = r^2(p; \theta_P) + c^2(p; \theta_P)
\]

(4)

The individual penalty applied for points at unexpected depths is shown in (5). The depth of a point, as a ratio of source-to-detector depth, is denoted by \( d \). Zero indicates a depth equal to the X-ray source and one indicates the depth of the X-ray detector.

A.2 Training Dataset Creation

Fig. S-1 shows a high-level workflow of the registrations used during creation of the training data set. The amounts of downsampling used for each method are listed in Table S-1. Any box constraints used by the following methods are listed in Table S-2.

A.2.1 Computationally Expensive Automatic Pelvis Registration

Two attempts are made to solve for the pose of the entire pelvis. The first attempt sequentially applies the following optimization strategies: Differential Evolution (DE) [23], exhaustive search, Covariance Matrix Adaptation: Evolutionary Search (CMA-ES) [25], and Bounded Optimization by Quadratic Approximation (BOBYQA) [26].

The DE optimization uses a regularizer designed to penalize poses which: do not project at least one femoral head center within the 2D image bounds, project inferior landmarks above superior landmarks in the image, or place either femoral head center behind the detector or too close to the X-ray source. This regularization is defined by \( R_{DE} \) in (3).
The individual penalty applied for projecting a certain point “above” another is defined in (6). For image visualization in this paper, smaller row values are located above larger values. Each image is assumed to be oriented “patient-up,” so that regions. Therefore, $\mathcal{R}_{up}(p, q; \theta_P)$ applies a penalty when the, relatively inferior, IOF landmark is projected above the, relatively superior, ASIS landmark.

\[
\mathcal{R}_{up}(p, q; \theta_P) = \begin{cases} 
(P(q; \theta_P)_{row} - P(p, \theta_P)_{row})^2 & \text{if } P(q; \theta_P)_{row} < P(p, \theta_P)_{row} \\
0 & \text{otherwise}
\end{cases}
\]  

If the first pelvis registration attempt is not successful, then another attempt is made using the following sequence of optimizations: exhaustive grid search, Particle Swarm Optimization (PSO) [24], and two runs of BOBYQA at increasing resolutions levels. No regularization is used during this attempt. The grid search used during this attempt is performed at coarser increments, but over a larger region, compared to the first attempt’s grid search. PSO was run for 50 iterations, with 21,000 particles, momentum $\omega = 0.7298$, local weight upper bound $\varphi_p = 1.4961$, and global weight upper bound $\varphi_p = 1.4961$.

If the first pelvis registration attempt is also unsuccessful, then no further attempt is made to annotate the current fluoroscopy image if this attempt is also unsuccessful.

### A.3.2 Registration of the Femurs

If the pelvis registration is successful, then an attempt is made to register the left and right femurs. This strategy first registers the left femur only, keeping the pelvis fixed at its current pose estimate. Next, the right femur is registered, again keeping the pelvis fixed. Both of these registrations use CMA-ES. Contrary to the previous registrations, these only search the 3D space of rotations, with the center of rotation fixed at the ipsilateral femoral head center. Regularization is applied to the total rotation magnitude using a folded normal distribution with $\mu = 45^\circ$ and $\sigma = 45^\circ$. Table S-4 lists the CMA-ES parameters. Once again, successful registrations of each object are manually verified.

| Object         | Strategy | Factor |
|----------------|----------|--------|
| Pelvis Attempt | DE       | 32×    |
|                | Grid     | 32×    |
|                | CMA-ES   | 8×     |
|                | BOBYQA   | 4×     |
| Pelvis Attempt | Grid     | 32×    |
|                | PSO      | 32×    |
|                | BOBYQA 1 | 8×     |
|                | BOBYQA 2 | 4×     |
| Femurs         | CMA-ES   | 8×     |
| All Objects    | BOBYQA   | 4×     |

### Table S-1 Amount of downsampling along each 2D image dimension applied during each optimization.

A.3.2 Registration of the Femurs

If the pelvis registration attempt is successful, then an attempt is made to register the left and right femurs. This strategy first registers the left femur only, keeping the pelvis fixed at its current pose estimate. Next, the right femur is registered, again keeping the pelvis fixed. Both of these registrations use CMA-ES. Contrary to the previous registrations, these only search the 3D space of rotations, with the center of rotation fixed at the ipsilateral femoral head center. Regularization is applied to the total rotation magnitude using a folded normal distribution with $\mu = 45^\circ$ and $\sigma = 45^\circ$. Table S-4 lists the CMA-ES parameters. Once again, successful registrations of each object are manually verified.

### A.3.2 Registration of the Femurs

If the pelvis registration attempt is successful, then an attempt is made to register the left and right femurs. This strategy first registers the left femur only, keeping the pelvis fixed at its current pose estimate. Next, the right femur is registered, again keeping the pelvis fixed. Both of these registrations use CMA-ES. Contrary to the previous registrations, these only search the 3D space of rotations, with the center of rotation fixed at the ipsilateral femoral head center. Regularization is applied to the total rotation magnitude using a folded normal distribution with $\mu = 45^\circ$ and $\sigma = 45^\circ$. Table S-4 lists the CMA-ES parameters. Once again, successful registrations of each object are manually verified.

### A.3.3 Loss Functions

For the segmentation branch of the network, the differentiable dice score [28] is computed for each class and then averaged as shown in (7). $N_C$ is equal to the number of classes including background (7 in this paper), $w$ are the network weights, $\hat{M}^{(k)}$ is the ground truth segmentation mask for class $k$, and $\tilde{M}^{(k)}$ is the estimated segmentation mask for class $k$.

\[
D(w) = \frac{1}{N_C} \sum_{k=1}^{N_C} \frac{2\sum_{x,y} M^{(k)}(x, y; w) \tilde{M}^{(k)}(x, y)}{\sum_{x,y} M^{(k)}(x, y; w)^2 + \sum_{x,y} \tilde{M}^{(k)}(x, y)^2}
\]  

Ground truth heatmaps for each landmark location, $(\tilde{z}^{(l)}, \tilde{y}^{(l)})$, are defined by (8), which is a symmetric 2D normal distribution with mean $(\tilde{z}^{(l)}, \tilde{y}^{(l)})$ and $\sigma = 3.88$ mm (2.5 pixels).

\[
\tilde{h}^{(l)}(x, y) = \begin{cases} 
(2\pi\sigma^2)^{-1} \exp \left\{ -\frac{(x-\tilde{z}^{(l)})^2 + (y-\tilde{y}^{(l)})^2}{2\sigma^2} \right\} & \text{if } (\tilde{z}^{(l)}, \tilde{y}^{(l)}) \text{ is visible} \\
0 & \text{otherwise}
\end{cases}
\]  

For two equal sized images $A$ and $B$, NCC is defined in (9). Each image has $P$ pixels, means $\mu_A$ and $\mu_B$, and
Fig. S-1 High-level workflow of the registrations used for ground truth labeling of fluoroscopy.

Table S-2 The $\mathfrak{se}(3)$ box constraints used for the registrations used to obtain ground truth annotations. For the all objects case, the box constraints are repeated for the three objects.

| Object | Strategy | Dimension 1 | Dimension 2 | Dimension 3 | Dimension 4 | Dimension 5 | Dimension 6 |
|--------|----------|-------------|-------------|-------------|-------------|-------------|-------------|
| Pelvis Attempt 1 | DE | ±60° | ±40° | ±10° | ±200 | ±200 | ±250 |
| | Grid | ±5° | ±5° | ±5° | ±10 | ±10 | ±50 |
| | BOBYQA | ±2.5° | ±2.5° | ±2.5° | ±5 | ±5 | ±5 |
| Pelvis Attempt 2 | Grid | ±60° | ±40° | ±10° | ±200 | ±200 | ±250 |
| | PSO | ±7.5° | ±10° | ±10° | ±20 | ±20 | ±25 |
| | BOBYQA 1 | ±5° | ±5° | ±5° | ±10 | ±10 | ±20 |
| | BOBYQA 2 | ±2.5° | ±2.5° | ±2.5° | ±5 | ±5 | ±10 |
| All Objects | BOBYQA | ±2.5° | ±2.5° | ±2.5° | ±2.5 | ±2.5 | ±2.5 |

The dice and heatmap terms are combined into the final loss shown in (11). In order to weight the dice and heatmap terms equally, $H(w)$ is scaled and shifted to the range of $[0, 1]$. Since the optimization during training seeks to find a
### A.3.3 Data Augmentation

Table S-5 lists the operations performed when an image is randomly selected to be augmented during training. Images are padded to 384 using reflection prior to warping, and in order to avoid possible intensity discontinuities. Fig. S-5 shows data before, and after, augmentation.

![Fig. S-2](image)

**Table S-5** Operations performed during data augmentation.

| Method               | Description                                                                |
|----------------------|-----------------------------------------------------------------------------|
| Intensity Inversion  | With probability 0.5                                                         |
| Additive Random Noise| $N(0, \sigma)$, $\sigma \sim U(0.005, 0.01)$                                |
| Noise                | $\gamma \sim U(0.7, 1.3)$                                                   |
| Alike Correction     | Translation direction uniformly sampled                                      |
| Affine Warp          | Translation magnitude from $U(0, 20)$ pixels                                 |
|                      | Rotation angle from $U(-5^\circ, +5^\circ)$                                |
|                      | Shear angle from $U(-2^\circ, +2^\circ)$                                    |
|                      | Scale from $U(0.9, 1.1)$                                                    |
| Local Corruption     | With probability 0.25                                                       |
|                      | Number of rectangular regions from $U\{(1, 2, 3, 4, 5)\}$                  |
|                      | Region dimensions from $N(d, d)$, $d = 0.15 \times$ image width            |
|                      | Location uniformly sampled, rejection sampling to ensure region is within image|
|                      | Additive noise from $N(0, 0.2m)$, $m$ is the range of intensities in a region|

During CMA-ES registration of the pelvis, 8x downsampling is used along with the parameters listed in Table S-4. For BOBYQA registration of the pelvis, 4x downsampling is used along with the BOBYQA box constraints for “Pelvis Attempt 1” in Table S-2.

### B Supplementary Results

#### B.1 Annotated Dataset Creation

Table S-6 lists the counts of the total number of images initially collected, the number of images with successful ground truth annotations, and the number of images with sufficient fields to view to perform femur registration. Fig. S-6 shows the C-arm poses from the 366 successful registrations used to build the training data set. Using a NVIDIA Tesla P100 (PCI-e), mean runtimes of 60.3 ± 13.3, 142.5 ± 35.8, and 2.5 ± 0.3 seconds were measured for attempt 1 of pelvis registration, attempt 2 of pelvis registration, and femur registration, respectively.

#### B.2 Segmentation and Landmark Localization

The mean training time for each network was 0.8 ± 0.1 hours using a NVIDIA Tesla P100 (PCI-e) GPU.

A listing of mean dice coefficients for each object of each “left-out” specimen is shown in Table S-7.

#### B.3 Intraoperative Registration

Counts of successful registrations and mean registration errors for each “left-out” specimen is shown in Table S-8. Full
Fig. S-3 The architecture of the U-Net encoder-decoder used in this work.

Fig. S-4 High-level network structure used in this paper. After an image is processed through a U-Net encoder-decoder module, the segmentation is computed using the standard approach. Segmentation features prior to soft-max are concatenated with the features output from the encoder-decoder and two 1x1 convolutions are used to estimate the landmark heatmaps.

Decompositions about each axis for pelvis pose errors are given Table S-9 and highlight that nearly all of the pelvis translation error is in the projective depth direction.

Table S-6 The number of fluoroscopy images identified for potential use and the number of images used for network training. Only images which were successfully registered with the ground truth labeling method were used for training. Of the images used for training, counts of the images with sufficient visibility of the left and right femurs for registration purposes are also listed. All specimens except one are used when training a specific network; the images for the left-out specimen are used as the test dataset.

| Specimen | Total Images | # Images Used For Training | # Training Images for L. Femur | # Training Images for R. Femur |
|----------|--------------|----------------------------|------------------------------|-------------------------------|
| 1        | 119          | 111                        | 52                           | 27                            |
| 2        | 108          | 104                        | 39                           | 24                            |
| 3        | 30           | 24                         | 0                            | 2                             |
| 4        | 53           | 48                         | 17                           | 18                            |
| 5        | 63           | 55                         | 13                           | 16                            |
| 6        | 26           | 24                         | 0                            | 12                            |
| All      | 399          | 366                        | 121                          | 99                            |
Fig. S-5  Example data augmentation of the projections from Fig. 2. The original projections are shown in the top row, and shown again in the second row with the original annotations overlaid. Projections after augmentation are shown in the third row, and the augmented annotations are overlaid in the bottom row.

Fig. S-6  A visualization of all ground truth projection geometries using the APP as the world coordinate frame. Each sphere represents a position of the X-ray source, each square represents the position of the X-ray detector, and each line connects the X-ray source to the principal point on the detector. Red arrows highlight difficult to see geometries of specimen 6.
Table S-7  Average dice coefficients obtained from each trained network from the leave-one-specimen-out experiment. Actual dice coefficient is reported, not dice loss defined by (7).

| Specimen | L. Hemipelvis | R. Hemipelvis | L. Femur | R. Femur | Vertebrae | Sacrum |
|----------|---------------|---------------|----------|----------|-----------|--------|
| 1        | 0.89 ± 0.15   | 0.89 ± 0.13   | 0.93 ± 0.17 | 0.78 ± 0.37 | 0.72 ± 0.21 | 0.63 ± 0.09 |
| 2        | 0.86 ± 0.23   | 0.85 ± 0.22   | 0.91 ± 0.23 | 0.94 ± 0.18 | 0.81 ± 0.09 | 0.66 ± 0.12 |
| 3        | 0.89 ± 0.07   | 0.91 ± 0.06   | 0.85 ± 0.33 | 0.56 ± 0.48 | 0.71 ± 0.18 | 0.59 ± 0.17 |
| 4        | 0.82 ± 0.24   | 0.81 ± 0.21   | 0.95 ± 0.08 | 0.83 ± 0.25 | 0.76 ± 0.14 | 0.53 ± 0.09 |
| 5        | 0.85 ± 0.18   | 0.88 ± 0.18   | 0.87 ± 0.28 | 0.85 ± 0.28 | 0.76 ± 0.21 | 0.68 ± 0.17 |
| 6        | 0.71 ± 0.25   | 0.89 ± 0.07   | 0.67 ± 0.41 | 0.97 ± 0.01 | 0.51 ± 0.30 | 0.56 ± 0.15 |
| All      | 0.86 ± 0.20   | 0.87 ± 0.18   | 0.90 ± 0.24 | 0.84 ± 0.31 | 0.74 ± 0.19 | 0.63 ± 0.13 |

Table S-8  Pelvis and femur registration errors from successful pelvis registrations using the three intraoperative approaches and broken down by cadaver specimen. Femur registrations errors are reported for all successful pelvis registrations which have sufficient visibility of a femur.

| Regi. Method | Spec. Spec. | Pelvis Errors | Femur Errors |
|--------------|-------------|---------------|--------------|
|              | # Success   | Rot. (°) | Trans. (mm) | # | Rot. (°) | Trans. (mm) |
| 1: Native    | 1           | 32 (29%)   | 0.1 ± 0.1   | 0.3 ± 0.2 | 13 | 0.4 ± 0.2 | 0.3 ± 0.3 |
|              | 2           | 15 (14%)   | 0.1 ± 0.2   | 0.8 ± 1.9 | 5  | 0.7 ± 0.4 | 0.4 ± 0.5 |
|              | 3           | 1 (4%)     | < 0.1       | 0.2      | 0  | —         | —         |
|              | 4           | 4 (8%)     | 0.1 ± 0.1   | 0.4 ± 0.4 | 0  | —         | —         |
|              | 5           | 13 (24%)   | 0.1 ± 0.1   | 0.4 ± 0.3 | 3  | 0.4 ± 0.2 | 0.7 ± 0.4 |
|              | 6           | 1 (4%)     | 0.1         | 0.3      | 1  | 0.6       | 0.2       |
|              | All         | 66 (18%)   | 0.1 ± 0.1   | 0.4 ± 0.9 | 22 | 0.4 ± 0.3 | 0.4 ± 0.3 |
| 2: PreP. Init. | 1          | 99 (89%)   | 0.1 ± 0.1   | 0.8 ± 1.1 | 73 | 1.7 ± 5.2 | 0.6 ± 0.5 |
|              | 2           | 96 (92%)   | 0.1 ± 0.2   | 1.0 ± 1.4 | 59 | 1.2 ± 1.0 | 0.5 ± 0.4 |
|              | 3           | 19 (79%)   | 0.2 ± 0.2   | 1.6 ± 3.1 | 2  | 0.9 ± 0.8 | 0.4 ± 1.0 |
|              | 4           | 38 (79%)   | 0.2 ± 0.2   | 1.4 ± 1.8 | 27 | 1.3 ± 1.2 | 0.4 ± 0.4 |
|              | 5           | 40 (73%)   | 0.1 ± 0.1   | 0.7 ± 0.9 | 20 | 0.8 ± 0.8 | 0.6 ± 0.7 |
|              | 6           | 7 (29%)    | 0.1 ± 0.1   | 0.8 ± 1.0 | 2  | 1.3 ± 0.6 | 0.4 ± 0.1 |
|              | All         | 299 (82%)  | 0.1 ± 0.2   | 1.0 ± 1.5 | 183| 1.4 ± 3.4 | 0.5 ± 0.5 |
| 3: Combined | 1           | 101 (91%)  | 0.1 ± 0.1   | 1.0 ± 1.5 | 73 | 1.8 ± 5.2 | 0.6 ± 0.5 |
|              | 2           | 99 (95%)   | 0.2 ± 0.2   | 1.4 ± 1.7 | 61 | 1.3 ± 1.0 | 0.7 ± 0.8 |
|              | 3           | 18 (75%)   | 0.2 ± 0.2   | 2.8 ± 3.4 | 2  | 1.1 ± 1.1 | 1.0 ± 1.3 |
|              | 4           | 41 (85%)   | 0.2 ± 0.2   | 2.1 ± 2.9 | 29 | 1.6 ± 1.3 | 0.6 ± 1.0 |
|              | 5           | 47 (85%)   | 0.1 ± 0.1   | 0.9 ± 1.2 | 24 | 0.8 ± 0.8 | 0.5 ± 0.6 |
|              | 6           | 7 (29%)    | 0.3 ± 0.3   | 3.0 ± 3.2 | 3  | 1.0 ± 0.7 | 0.3 ± 0.2 |
|              | All         | 313 (86%)  | 0.2 ± 0.2   | 1.4 ± 2.0 | 192| 1.5 ± 3.3 | 0.6 ± 0.7 |

Table S-9  Mean absolute values of each decomposed component of pelvis pose errors for which intraoperative registration was successful using method 3. The axes are aligned with the projective coordinate frame with Z corresponding to depth.

| Specimen | Rotation (°) | Translation (mm) |
|----------|--------------|------------------|
|          | X  Y  Z      | X  Y  Z          |
| 1        | 0.1 ± 0.1   | 0.1 ± 0.1       | < 0.1 0.1 ± 0.1 0.1 ± 0.1 | 1.0 ± 1.5 |
| 2        | 0.1 ± 0.1   | 0.1 ± 0.1       | < 0.1 0.1 ± 0.1 0.1 ± 0.1 | 1.4 ± 1.7 |
| 3        | 0.1 ± 0.1   | 0.2 ± 0.2       | 0.1 ± 0.1 0.1 ± 0.1 0.1 ± 0.1 | 2.7 ± 3.4 |
| 4        | 0.1 ± 0.1   | 0.2 ± 0.2       | 0.1 ± 0.1 0.1 ± 0.1 0.1 ± 0.1 | 2.0 ± 2.9 |
| 5        | 0.1 ± 0.1   | 0.1 ± 0.1       | < 0.1 0.1 ± 0.1 < 0.1 0.1 ± 0.1 | 0.9 ± 1.2 |
| 6        | 0.2 ± 0.3   | 0.2 ± 0.1       | < 0.1 0.1 ± 0.1 0.1 ± 0.1 | 3.0 ± 3.2 |
| All      | 0.1 ± 0.1   | 0.1 ± 0.1       | < 0.1 0.1 ± 0.1 0.1 ± 0.1 | 1.4 ± 2.0 |