Surface Registration to Estimate Motion in CBCT

Bastian Bier, Mathias Unberath, Nishant Ravikumar, Jennifer Maier, Ali Gooya, Zeike A. Taylor, Alejandro F. Frangi, Garry Gold, Rebecca Fahrig, and Andreas Maier

***Abstract*—C-arm cone-beam CT (CBCT) has been used recently to acquire images of a patient's knee under weight-bearing conditions. The resulting reconstructions allow to assess knee joint health under more realistic conditions than in supine configuration. However, motion during the scan severely corrupts image quality of the reconstructions. Recently, a method to compensate for this motion has been proposed that, in contrast to its predecessors, uses a depth camera in addition to the CBCT. The method is based on the popular Iterative Closest Point (ICP) algorithm, that requires the selection of a reference and, therefore, induces bias.**

In this work, we investigate a group-wise registration approach to estimate motion, which is more robust to noise and outliers. Further, the group-wise character can estimate the mean shape more accurately. We compare our results to a state-of-the-art marker-based method, as well as to the previously proposed ICP-based method.

Image quality improves compared to the ICP-based method with an improvement of the Structural Similarity (SSIM) from 0.96 to 0.98. Streaks in the images could be reduced slightly.

The preliminary results presented here are promising. Dense surface information together with the stochastic formulation of the proposed method allows for the incorporation of more complex, e.g. compound motion, that better reflects true joint motion. Further investigations in this direction are subject to future work.

I. Introduction

IMAGE acquisition under weight-bearing conditions has been enabled lately due to advances in the flexibility of C-arm cone-beam CT (CBCT) systems [1] or dedicated imaging systems [2]. These systems acquire projection images from different views of the standing patient [3]. Relatively long scan times lead to inevitable patient motion that results in motion artifacts in the reconstructed images. Diagnostic assessment of these volumes requires motion estimation and compensation of the data prior to reconstruction in order to achieve a clinically useful image quality.

Currently, one state-of-the-art method relies on fiducial markers placed on the patient's skin that are then tracked in projection images [1]. Marker placement, however, is time consuming and tedious and only yields sparse surface data. Range imaging, in contrast, yields dense 3D surface information, does not require markers, and can be used to compensate for motion when acquired simultaneously to the X-ray sequence. Fotouhi et al. [4] mounted a range camera on the detector to acquire the patient's surface. The surface is reconstructed using a SLAM-based object reconstruction and used as prior for iterative reconstruction to enhance reconstruction quality. In contrast to [4], Bier et al. [5] assume a static depth camera and estimate object motion directly using an Iterative Closest Point (ICP) algorithm by registering all frames to the first motion state. Unfortunately, this method has several limitations: estimation in vertical direction proved unreliable, and, more importantly, registration is carried out in a pair-wise manner, suggesting the need for reference frame selection that induces bias. Further, ICP-based techniques are known to be highly sensitive to both the initialization and the presence of outliers. An elegant solution is to cast the registration problem into a probabilistic framework; registration then becomes probability density estimation.

In this study, a multi-resolution, group-wise, rigid registration approach (with 6 degrees of freedom in 3D, namely, translation and rotation) based on Student's t-mixture model (TMM) [6], [7] is employed to further improve direct motion estimation. Group-wise registration methods are in general preferable to pair-wise approaches as they ensure unbiased estimation of the desired spatial transformations and correspondences, whereas the latter are biased towards the chosen reference surface. Furthermore, the choice of TMM-based registration confers a greater degree of automatic robustness to outliers in the data, than conventional Gaussian mixture model based methods. Consequently, such an approach is well suited to clinical applications, where noise and other sources of artifacts are to be expected. We applied this method on simulated data and estimate 3D as well as 6D motion. For evaluation, we compare with the ground truth, the uncorrected, the marker-based, and the ICP-based reconstructions, comparing image quality using the Structural Similarity (SSIM) [8].

II. Materials and Methods

Group-Wise Registration: Probabilistic group-wise registration approaches consider the group of shapes to be aligned, as noisy transformed observations of a central mixture model. This central model, which represents the mean shape of the group, is fit to each sample shape in the group using expectation-maximization (EM) [9]. EM alternates between two steps: the E-step, where the expectations of the latent variables (or posterior probabilities) are evaluated for the initial/current estimate of the model and transformation parameters; and the M-step, where the estimated posterior probabilities are in turn employed to maximize the complete data likelihood, with respect to the model and transformation parameters. Consequently, the desired spatial transformation that maps the mean shape to each sample in the group

B. Bier, M. Unberath, N. Ravikumar, J. Maier, and A. Maier are with the Department of Computer Science, Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany.

G. Gold and R. Fahrig* are with the School of Medicine, Stanford University. *now with Siemens Healthcare GmbH

N. Ravikumar, A. Goya, Z. Taylor, and A. Frangi are with CISTIB Center for Computational Imaging and Simulation Technologies, The University of Sheffield, UK.

is estimated as parameters of the model in the M-step of the algorithm. Additionally, at each M-step the mean shape itself (represented by the centroids of the mixture model) is iteratively refined along with the spatial transformations [6], [7]. The multi-resolution variant of the TMM-registration algorithm used in this study increases the density of the mean model in an adaptive fashion at each successive resolution. This approach reduces the influence of local minima during the registration process, as discussed in [7].

Data: X-ray projection images and 3D point clouds are simulated on a segmented knee extracted from a clinical high resolution supine CBCT reconstruction. To simulate realistic motion states, the high resolution scan is transformed according to real patient motion acquired with a motion capture system [5], [1]. X-ray projections are created using a short scan trajectory (200°) with 248 projection images (1240×960 pixels, isotropic pixel spacing 0.308 mm). Reconstructions have a size of 512^3 and an isotropic voxel size of 0.5 mm. The depth camera is modeled to be similar to the Microsoft Kinect One v2 that acquires points on a grid with a depth resolution of 1 mm. The depth camera is statically located in front of the knee observing the same scene for all time points.

Experiments: Two experiments are conducted: (a) 3D translational motion is simulated and then estimated. We compare our proposed approach with the marker-based [1] and the ICP-based method [5]. (b) real 6D affine motion is applied and the results are compared to the marker-based result only. Note, that the ICP-based method is currently restricted to 3D translational motion, whereas the marker-based method and the proposed approach are fit to estimate 6D rigid motion.

In order to compare the image quality of the respective motion-compensated reconstructions, the SSIM is computed. To this end, all reconstructions are registered to the motion free reference reconstruction. The estimated motion for each time point is incorporated into the projection matrices of the imaging geometry prior to reconstruction.

III. Results

Axial and sagittal slices of the reconstructions of experiment (a) are shown in Figure 1. Compared to the ICP-based method [5], streak artifacts are reduced, which is supported by the improvement of the SSIM from 0.96 to 0.98, see Table I.

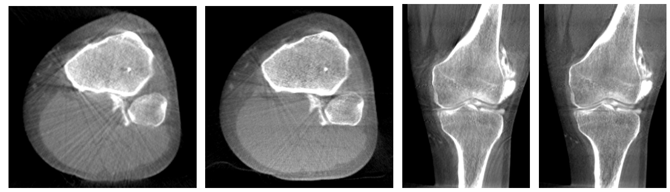
The results of the 6D rigid body motion (b) is shown in Figure 2. We show the ground truth, the uncompensated, the marker-based, and the proposed approach. Both the marker-based and the proposed approach are able to substantially reduce the motion artifacts, comparable SSIM values are obtained with 0.98 and 0.97, respectively.

Figure 3 shows the estimated translations and the error of the estimated rotations. Note, that a constant offset in the estimation only corresponds to a shift of the reconstruction in the volume. The rotation error is calculated and compared to the Identity rotation using the Frobenius norm.

IV. Discussion and Conclusion

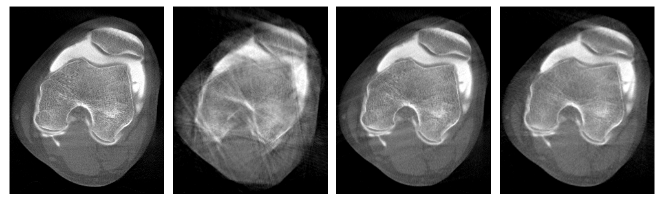
In this work, we applied group-wise point cloud registration to solve a motion estimation problem in CBCT acquisitions of knees under weight-bearing conditions, improving the previously presented pair-wise ICP-based approach. Point clouds at different motion states are registered in a probabilistic framework resulting in 3D or 6D motion estimates, that are incorporated into the tomographic reconstruction.

Future work will evaluate the algorithm on a broader database with different motion patterns. A noise study has to be conducted to evaluate the robustness and accuracy of the estimation under more realistic conditions [10]. However, results are promising and encourage future work that could leverage the availability of dense surface information to estimate more realistic joint motion such as compound rigid or non-rigid motion.



(a) ICP-based. (b) Group-wise. (c) ICP. (d) Group-wise.

Fig. 1: Results (a): axial and sagittal slice of the knee joint.



(a) Ground truth. (b) Corrupted. (c) Marker-based. (d) Group-wise.

Fig. 2: Results (b): shows axial slices.

References

[1] J.-H. Choi, R. Fahrig, A. Keil, T. F. Besier, S. Pal, E. J. McWalter, G. S. Beaupré, and A. Maier, “Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization,” *Med. Phys.*, vol. 41, no. 6, p. 061902, 2014.

[2] A. Sisniega, J. Stayman, J. Yorkston, J. Siewerdsen, and W. Zbijewski, “Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion.” *Phys. Med. Biol.*, vol. 62, no. 9, 2017.

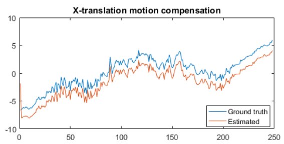
[3] A. Maier, J.-H. Choi, A. Keil, C. Niebler, M. Sarmiento, A. Fieselmann, G. Gold, S. Delp, and R. Fahrig, “Analysis of vertical and horizontal circular C-arm trajectories,” *SPIE Medical Imaging*, vol. 7961, pp. 796 123–1 – –796 123–8, 2011.

[4] J. Fotouhi, B. Fuerst, W. Wein, and N. Navab, “Can real-time RGBD enhance intraoperative Cone-Beam CT?” *IJCARS*, 2017.

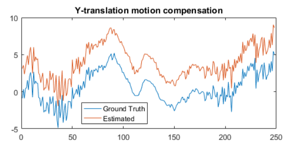
[5] B. Bier, M. Unberath, T. Geimer, J. Maier, G. Gold, M. Levenston, R. Fahrig, and A. Maier, “Motion Compensation using Range Imaging in C-arm Cone-Beam CT,” in *MIUA (accepted)*, 2017.

TABLE I: SSIM of the reconstructed images.

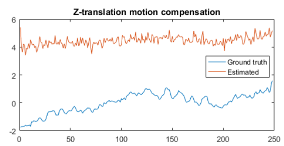
Method	3D estimation	6D estimation
Uncorrected	0.90	0.91
Marker-based [1]	0.98	0.98
ICP-based [5]	0.96	-
Group-wise registration [6]	0.98	0.97



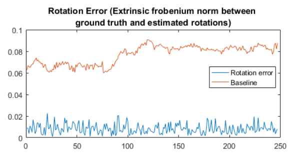
(a) X estimate.



(b) Y estimate.



(c) Z estimate.



(d) Rotation error.

Fig. 3: Translation estimates of the proposed method as well as the calculated error of the rotation.

[6] N. Ravikumar, A. Gooya, A. F. Frangi, and Z. A. Taylor, "Robust groupwise rigid registration of point sets using t-mixture model," in *SPIE Medical Imaging 2016*, vol. 9784, 2016, p. 97840S.

[7] N. Ravikumar, A. Gooya, S. Cimen, A. F. Frangi, and Z. A. Taylor, "A Multi-resolution T-Mixture Model Approach to Robust Group-Wise Alignment of Shapes," in *MICCAI 2016)*, vol. 9900, 2016, p. 697.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE TMI*, vol. 13, no. 4, pp. 600–612, 2004.

[9] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, Nov 1996.

[10] J. Rausch, A. Maier, R. Fahrig, J. H. Choi, W. Hinshaw, F. Schebesch, S. Haase, J. Wasza, J. Hornegger, and C. Riess, "Kinect-based correction of overexposure artifacts in knee imaging with C-Arm CT systems," *International Journal of Biomedical Imaging*, vol. 2016, pp. 1–9, 2016.