2D/3D Registration for Motion Compensated Reconstruction in Cone-Beam CT of Knees Under Weight-Bearing Condition

M. Berger[1,2,3], K. Müller[2], J.-H. Choi[2], A. Aichert[1], A. Maier[1,3,4] and R. Fahrig[2]

[1] Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany
[2] Radiological Sciences Lab, Stanford University, Stanford, USA
[3] Research Training Group 1773 "Heterogeneous Image Systems", Erlangen, Germany
[4] Erlangen Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany

Abstract— **Over the last decade, increased effort has been made to acquire three dimensional images of knee joints under weight-bearing condition. Cone-beam CT systems are popular because of their high flexibility with respect to patient position and scan trajectory. However, scans in a standing or squatting patient position are affected by involuntary patient motion during the acquisition, which results in streaking and blurring artifacts in the reconstructed volumes. Previous work suggested the use of fiducial markers to estimate and compensate for motion artifacts. However, marker placement on the skin might not accurately reflect the motion at the center of the joint. In this work, we propose a marker-free motion compensation method that is based on 2D/3D rigid registrations of individual projection images to a segmentation of the bones from a prior, motion-free scan. The estimated motion of the individual bones is then combined to a global motion field to allow for a motion compensated reconstruction. Qualitative and quantitative results show substantial improvement compared to uncorrected images. Incorporating smoothness constraints into the estimated motion parameters during the registration further improved the results.**

Keywords— **Image Reconstruction, Knee-Joint Imaging, Motion Compensation, 2D/3D Registration**

I. Introduction

Computed Tomography (CT) is a common diagnostic tool to assess the anatomical morphology of the knee joint. Compared to diagnostic CT, cone-beam CT (CBCT) provides more flexibility in terms of patient position and positioning of the system. It has been shown that modern angiography CBCT systems are able to acquire 3D volumetric images of the knee joint under physiologic weight-bearing conditions, e.g. in a standing or squatting position [1]. However, the high acquisition time ($>$ 5s) of CBCT devices can lead to severe streaking and blurring artifacts caused by patient motion.

External fiducials are popular for motion estimation because they can be tracked in the 2D projection images. Choi et al. have used externally attached fiducial markers to correct for patient motion during standing and squatting acquisitions of knee joints [1]. Even though the markers can be tracked accurately, they may not reflect the internal motion of the bones or cartilage. Additionally, attaching markers might lead to increased examination times and patient discomfort. Purely data-driven motion estimation methods could help to overcome these limitations.

One approach of data-driven motion correction is data consistency conditions (DCC). DCCs build a mathematical description of the redundant information that must be fulfilled by projection images acquired in an ideal setup. These methods typically estimate a motion model by optimizing the underlying DCC and have been successfully applied to simulated data [2, 3, 4]. Yet, little work has been done on real acquisitions, leaving their practical applicability in question.

In previous work we introduced a data-driven motion correction based on 2D/2D registrations between maximum-intensity projections of the motion corrupted reconstruction and the acquired projections [5]. No additional information was necessary, but the improvement in image quality was limited. In this paper we propose a novel data-driven motion compensation method that is based on 2D/3D registrations of segmented bone volumes from a prior, motion-free scan.

II. Methods

The pipeline of our motion estimation and correction approach is depicted in Fig. 1. First, the bones that best represent the joint motion are segmented. The segmentations are roughly aligned in 3D with respect to an initial motion corrupted reconstruction. Motion parameters are estimated by 2D/3D registration of each bone to each projection image. After combining the individual bone motions to a global deformation, we can generate a corrected reconstruction.

A. Initialization

As the prior data might be in a substantially different coordinate system, e.g. supine vs. standing, an initialization step

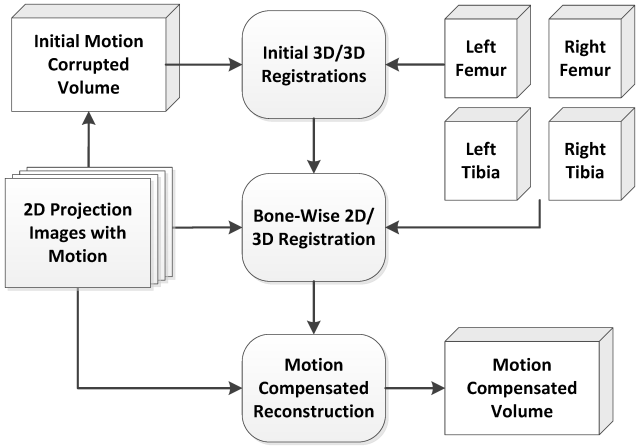


Fig. 1: Pipeline of our motion correction method.

is necessary. Thus, an initial, motion corrupted reconstruction is built and used to manually align the individual bone volumes. The result is a set of affine transformation matrices $\mathbf{S}_j \in \mathbb{R}^{4\times4}$, where j denotes the j-th bone volume. These transformations will roughly align the bones' coordinate systems with those of the current scan.

B. 2D/3D Registration

We applied the "projection-method" described in Markelj et al. [6], which is a commonly used 2D/3D registration technique. The basic idea is to vary the 3D position of a bone, simulate the X-ray acquisition by a digitally reconstructed radiograph (DRR), and compare the DRR result with the acquired projection using a similarity measure. The similarity measure is then optimized with respect to the position.

Let $\mathbf{P}_i \in \mathbb{R}^{3\times4}$ be the i-th of N projection matrices that describe the scanner's geometry and define how a 3D point in space is mapped to the 2D detector. We further introduce an affine mapping $\mathbf{M}_{ij} \in \mathbb{R}^{4\times4}$ that encodes the estimated rigid motion for the j-th bone at the i-th projection image. Then a 3D point $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^\mathrm{T}$, belonging to a bone segmentation, can be mapped to a homogeneous 2D point by

$$\mathbf{u}_i = \begin{bmatrix} u_i \\ v_i \\ k_i \end{bmatrix} = \mathbf{P}_i \, \mathbf{M}_{ij} \, \mathbf{S}_j \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \quad .$$

The actual detector coordinate can be computed by the mapping $h : \mathbb{R}^3 \mapsto \mathbb{R}^2$ with $h(\mathbf{u}_i) = \begin{bmatrix} u_i/k_i & v_i/k_i \end{bmatrix}^\mathrm{T}$. With f_j being the intensity function of the j-th bone, we can now formalize the DRR with help of the Dirac function as follows:

$$d_{ij}(u,v) = \int_{-\infty}^{\infty} f_j(\mathbf{x}) \, \delta \left(\left\| h\left(\mathbf{P}_i \, \mathbf{M}_{ij} \, \mathbf{S}_j \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \right) - \begin{bmatrix} u \\ v \end{bmatrix} \right\|_2 \right) \quad (1)$$

Let $c(.)$ be the similarity measure that compares a projection image $p_i(u,v)$ with the DRR image $d_{ij}(u,v)$, then the optimization function for a single bone can be denoted as

$$\Phi(\mathbf{M}_{1j}, \cdots, \mathbf{M}_{Nj}) = \sum_{i=1}^{N} c(d_{ij}(u,v), p_i(u,v)) \quad .$$

As similarity measure we used the gradient correlation measure as described in Penney et al. [7], but with a row and column-wise Laplacian instead of the 1st order derivatives. This was helpful to further reduce the influence of surrounding tissue and keep the focus on the bones' outline. We limited the optimization to 3D translations only, which assumes that the rotational contributions have been estimated sufficiently well in the initialization step. Thus, the optimization function reduced to $\Phi(\mathbf{t}_{1j}, \cdots, \mathbf{t}_{Nj})$, where $\mathbf{t}_{ij} \in \mathbb{R}^3$ is the estimated 3D translation of the j-th bone for the i-th projection.

To increase robustness we also imposed a smoothness constraint to the translations in temporal direction. Our regularization aims to minimize the energy of the difference between original and low-pass filtered translations. Let k_i^σ be Gaussian filter coefficients based on a standard deviation σ, then we can reformulate the cost function to

$$\widehat{\Phi}(\mathbf{t}_{1j}, \cdots, \mathbf{t}_{Nj}) = \Phi(\mathbf{t}_{1j}, \cdots, \mathbf{t}_{Nj}) + \lambda \sum_{i=1}^{N} \| \mathbf{t}_{ij} - (\mathbf{t} * k^\sigma)_{ij} \|_2^2 \quad ,$$

where $(\mathbf{t} * k^\sigma)_{ij}$ denotes the low-pass filtering over i.

C. From Local to Global Motion

Three planes that separate 1) left and right leg, 2) left femur and left tibia and 3) right femur and right tibia, have been defined manually in the motion corrupted reconstruction. Thus, the volume is divided into four compartments, where each has its own translation $\mathbf{t}_{ij}$. To avoid artifacts in proximity to a plane we introduce a weighting function that ensures a smooth transition at plane boundaries. The resulting translation $\mathbf{t}_{ij}(\mathbf{x})$ is given by,

$$w(\mathbf{x}) = 1/(1 + \exp(-c\,\varphi(\mathbf{x})))$$

$$\mathbf{t}_{ij}(\mathbf{x}) = w(\mathbf{x})\mathbf{t}_{ij}^1 + (1 - w(\mathbf{x}))\mathbf{t}_{ij}^2 \quad ,$$

where $\varphi(\mathbf{x})$ is the distance to the plane, $\mathbf{t}_{ij}^1$ and $\mathbf{t}_{ij}^2$ are the translations assigned to each side of the plane and c is a scaling factor that was determined heuristically.

III. Results

A. Experimental Setup

To verify our approach we used a simulated XCAT-based dataset from Choi et al. [8]. XCAT is a numeric phantom

based on segmentations of real bones, organs and tissue [9]. Real motion parameters of healthy volunteers were incorporated into the XCAT model using an optical tracking system. For further details, refer to Choi et al. [8].

The bone volumes were generated by volume rendering the XCAT bones in a supine coordinate system using an isotropic voxel spacing of 0.5 mm. The initial alignment of all bones to the motion corrupted reconstruction was done manually using the software 3D Slicer [10]. For the 2D/3D registration, we implemented the DRR generation and the gradient correlation using CONRAD [11]. Overall $248 \times 3 = 744$ parameters were estimated for each bone using an interior-point optimization algorithm.

B. Qualitative and Quantitative Assessment

Transversal slices of the reference, motion corrupted, and motion corrected reconstructions are shown in Fig. 2. Top and bottom rows show the femur and tibia, respectively. The motion induced streaking and blurring artifacts are clearly visible in the non-corrected reconstruction. All correction methods substantially improve the image quality. We can see a reduction of streak artifacts and clearer bone edges with $\sigma = 1$. Increasing the regularization to $\sigma = 2$ led to slightly increased streaking but still improved certain locations.

We conducted an image based quantitative comparison of the reference and corrected reconstructions by computing the relative root-mean-square-error (rRMSE) and the universal image quality index (UQI) [12]. For the reference image, all projections were fixed at the first time-frame. This means that the global alignment and therefore also our measures is heavily dependent on the accuracy of only 3 of the 744 parameters. To avoid this dependency, we applied a global 3D/3D registration for each bone region to the reference reconstruction using 3D Slicer [10]. The pipeline was as follows:

1. Corrected reconstruction with respect to the first frame
2. Generation of ROI for each bone, including soft tissue
3. For each ROI:
   (a) 3D/3D registration of ROI to reference volume
   (b) New reconstruction including estimated global transformation (to avoid additional interpolation)
   (c) Computation of rRMSE + UQI for each ROI
4. Construction of mean value of bone-wise measures as shown in Table 1

All correction methods improved the image quality compared to the case without correction (cf. Tab. 1). Regularization helped to improve values for rRMSE and UQI. The best result was obtained by $\sigma = 1$. The increased streaking at $\sigma = 2$ did not substantially influence the quantitative measures.

Table 1: Mean quantitative measurements over all registered bone regions.

	rRMSE (%)	SSIM ($\times 10^{-2}$)
With Motion	12.80	38.37
Corrected (No Reg.)	5.78	86.02
Corrected ($\sigma = 1$)	**5.02**	**88.69**
Corrected ($\sigma = 2$)	5.18	87.95

IV. Discussion

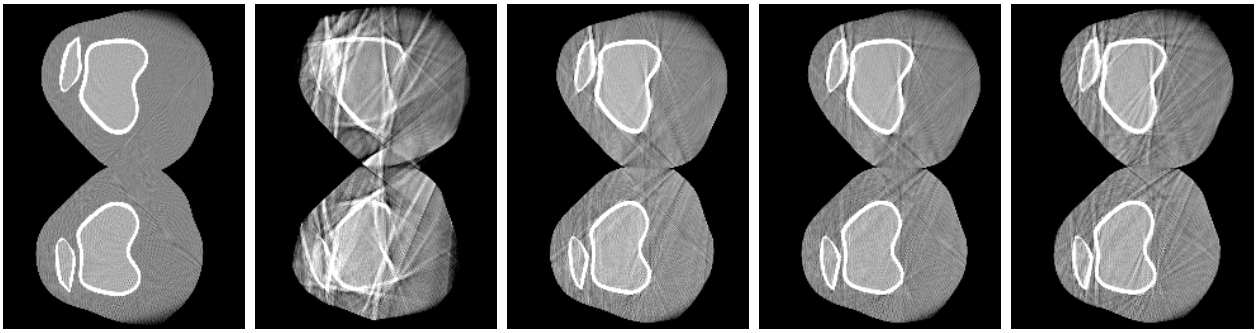
Motion compensation in CBCT is a challenging task and still an active field of research. In most applications reported in the literature, an auxiliary signal is needed to reduce complexity. One example is the ECG signal used in cardiovascular imaging or fiducial markers used in previous weight-bearing knee imaging studies. In this work, we present a data-driven approach that uses segmented bones of a prior, motion-free dataset. The method enables marker-free motion compensation and does not require an auxiliary signal.

Qualitative and quantitative results show great improvement over a case with no motion correction. In particular the method accurately restores the bones' outlines. This is important because most relevant structures in the knee joint are located close to the bones, e.g., the ligaments and cartilage. Regularization of the objective function by requiring smooth translations over time further improve the results. However, care must be taken in adjusting the amount of regularization because overly-smooth transformation parameters might not be able to fully recover the underlying motion (cf. Fig. 2(e)).

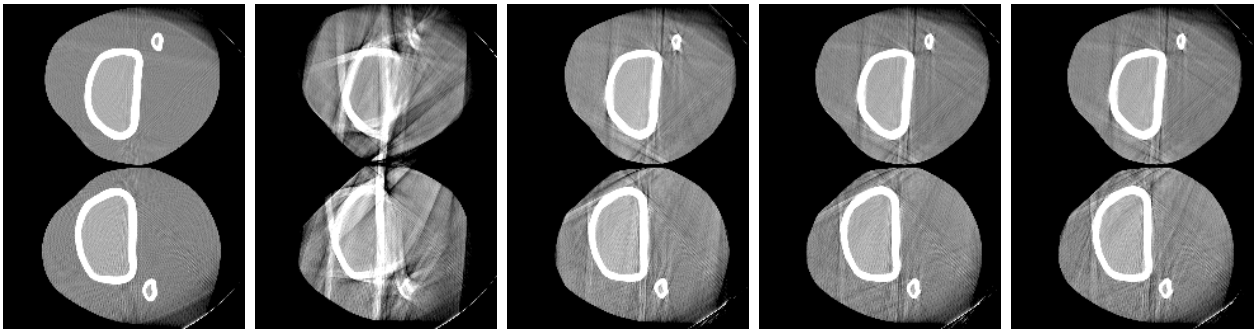
Our simulated data are based on a real motion pattern and the anthropomorphic nature of the XCAT phantom suggests that our method's effectiveness on real data is promising. To confirm our results, more evaluation on simulated and real data is necessary and is a goal for future work. To improve results, we plan to extend the registration by optimizing not only the bones' translations but also their rotations. Additionally, we plan to incorporate information given by the physiological model of the joint to further improve the registration.

V. Conclusion

A novel data-driven approach for motion compensation in C-arm CT acquisition of the knee joint is presented. The method applies 2D/3D registration of bones segmented from a prior, motion free scan. The estimated transformations are then incorporated into a global motion field to allow for a motion-corrected reconstruction. Our qualitative and quantitative results show a substantial improvement in image quality compared with an uncorrected reconstruction.



(a) Reference (b) With Motion (c) Corrected (No Reg.) (d) Corrected ($\sigma = 1$) (e) Corrected ($\sigma = 2$)



(f) Reference (g) With Motion (h) Corrected (No Reg.) (i) Corrected ($\sigma = 1$) (j) Corrected ($\sigma = 2$)

Fig. 2: Qualitative comparison of the reconstructed volumes. Top row: Transverse slice of femur and patella. Bottom row: Transverse slice of Tibia and Fibula. The visualization window was chosen at a density of [0.7, 2.1].

ACKNOWLEDGMENT

The authors gratefully acknowledge funding support from the NIH Shared Instrument Grant S10 RR026714 supporting the zeego@StanfordLab, and Siemens AX. The authors also gratefully acknowledge funding of the Research Training Group 1773 Heterogeneous Image Systems and the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG).

REFERENCES

1. Choi J.H., Maier A., Keil A., et al. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. II. Experiment *Medical Physics.* 2014;41:061902.
2. Clackdoyle R., Desbat L.. Data consistency conditions for truncated fanbeam and parallel projections *Medical Physics.* 2015;42:831–845.
3. Berger Martin, Maier Andreas, Xia Yan, et al. Motion Compensated Fan-Beam CT by Enforcing Fourier Properties of the Sinogram in *Proceedings of the third international conference on image formation in x-ray computed tomography*:329–332 2014.
4. Aichert A., Maass N., Deuerling-Zheng Y., et al. Redundancies in X-ray images due to the epipolar geometry for transmission imaging in *Proceedings of the third international conference on image formation in x-ray computed tomography*:333–337 2014.
5. Unberath M., Choi J.H., Berger M., et al. Image-based Compensation for Involuntary Motion in Weight-bearing C-arm Cone-beam CT Scanning of Knees in *SPIE Medical Imaging* 2015. (In press).
6. Markelj P., Tomaževič D., Likar B., et al. A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis.* 2012;16:642–61.
7. Penney G.P., Weese J., Little J.A., et al. A comparison of similarity measures for use in 2-D-3-D medical image registration *Medical Imaging, IEEE Transactions on.* 1998;17:586-595.
8. Choi J.H., Fahrig R., Keil A., et al. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization. *Medical physics.* 2013;40:091905.
9. Segars W.P., Sturgeon G.M., Mendonca S., et al. 4D XCAT phantom for multimodality imaging research *Medical Physics.* 2010;37:4902.
10. Pieper S., Halle M., Kikinis R.. 3D Slicer in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*:632-635 Vol. 1 2004.
11. Maier A., Hofmann H.G., Berger M., et al. CONRAD–a software framework for cone-beam imaging in radiology. *Medical physics.* 2013;40:111914.
12. Wang Zhou, Bovik A.C.. A universal image quality index *Signal Processing Letters, IEEE.* 2002;9:81-84.

Author:	Martin Berger
Institute:	Pattern Recognition Lab,
	Friedrich-Alexander-University Erlangen-Nuremberg
Street:	Martensstrasse 3
City:	Erlangen
Country:	Germany
Email:	martin.berger@cs.fau.de