Inertial Measurements for Motion Compensation in Weight-bearing Cone-beam CT of the Knee

Jennifer Maier\textsuperscript{1,2}, Marlies Nitschke\textsuperscript{2}, Jang-Hwan Choi\textsuperscript{3}, Garry Gold\textsuperscript{4}, Rebecca Fahrig\textsuperscript{5}, Bjoern M. Eskofier\textsuperscript{2}, and Andreas Maier\textsuperscript{1}

\textsuperscript{1} Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
\textsuperscript{2} Machine Learning and Data Analytics Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
\textsuperscript{3} College of Engineering, Ewha Womans University, Seoul, Korea
\textsuperscript{4} Stanford University, Stanford, California, USA
\textsuperscript{5} Siemens Healthcare GmbH, Forchheim, Germany
jennifer.maier@fau.de

Abstract. Involuntary motion during weight-bearing cone-beam computed tomography (CT) scans of the knee causes artifacts in the reconstructed volumes making them unusable for clinical diagnosis. Currently, image-based or marker-based methods are applied to correct for this motion, but often require long execution or preparation times. We propose to attach an inertial measurement unit (IMU) containing an accelerometer and a gyroscope to the leg of the subject in order to measure the motion during the scan and correct for it. To validate this approach, we present a simulation study using real motion measured with an optical 3D tracking system. With this motion, an XCAT numerical knee phantom is non-rigidly deformed during a simulated CT scan creating motion corrupted projections. A biomechanical model is animated with the same tracked motion in order to generate measurements of an IMU placed below the knee. In our proposed multi-stage algorithm, these signals are transformed to the global coordinate system of the CT scan and applied for motion compensation during reconstruction. Our proposed approach can effectively reduce motion artifacts in the reconstructed volumes. Compared to the motion corrupted case, the average structural similarity index and root mean squared error with respect to the no-motion case improved by 13-21\% and 68-70\%, respectively. These results are qualitatively and quantitatively on par with a state-of-the-art marker-based method we compared our approach to. The presented study shows the feasibility of this novel approach, and yields promising results towards a purely IMU-based motion compensation in C-arm CT.

Keywords: Motion Compensation · Inertial Measurements · CT Reconstruction.

1 Introduction

Osteoarthritis is a disease affecting articular cartilage in the joints, leading to a higher porosity and eventually to loss of tissue \cite{2}. The structural change of
cartilage also has an influence on its mechanical properties, i.e. its behavior when put under stress [18]. To analyze how diseased cartilage in the knee joint changes under stress compared to healthy tissue, it can be imaged under load conditions. This can be realized by scanning the knee joint in a weight-bearing standing position using a flexible C-arm cone-beam Computed Tomography (CT) system rotating on a horizontal trajectory, as depicted in Figure 1a [15]. However, standing subjects will show more involuntary motion due to body sway when naturally standing compared to the conventional supine scanning position [22]. Since standard CT reconstruction assumes stationary objects, this motion leads to streaking artifacts, double contours, and blurring in the reconstructed volumes making them unsuitable for clinical diagnosis.

Multiple approaches to correct for this motion have been proposed in literature. There exist purely image-based methods, like 2D/3D registration [2] or the use of a penalized image sharpness criterion [21], that show very good motion compensation results but are computationally expensive. It is also possible to use epipolar consistency conditions for motion compensation, but for knee imaging this has so far only been applied for estimating translation and not rotation [4]. Recently, also deep learning methods were applied in CT motion correction. Bier et al. proposed a neural network to detect anatomical landmarks in projection images, but the approach was not robust when other objects were present and was only evaluated for tracking motion [5]. Its feasibility for compensating motion was not investigated. An approach requiring external hardware is using range cameras to track motion during the scan, which so far worked well on purely simulated data [6]. The gold standard method for knee motion compensation is based on small metallic markers attached to the scanned leg, and was
Inertial Measurements for CT Motion Compensation

proposed in [8,9]. The markers tracked in the projections can be used to iteratively estimate motion. However, the placement of the markers is tedious and the metal produces artifacts in the reconstructions.

Inertial measurement units (IMUs) containing an accelerometer and a gyroscope have found use in C-arm CT for navigation [13] and calibration [14] purposes. We propose to use these small and lightweight sensors for motion compensation in C-arm CT. For this purpose, an IMU is attached to the leg of the subject to measure motion during the scan. To show the feasibility of this approach, we present a simulation study using real 3D human swaying motion recorded with an optical tracking system. These measurements are used to animate an OpenSim biomechanical model by inverse kinematics computation [11,20]. The model’s movement is on the one hand used to deform an XCAT numerical phantom for the generation of motion corrupted CT projections [19]. On the other hand, it is used to simulate inertial measurements of a sensor placed on the leg. The simulated measurements are processed in a multi-stage motion correction pipeline consisting of gravity removal, local velocity computation, global transformation, and projection geometry correction.

2 Materials and Methods

In order to generate realistic X-ray projections and IMU measurements, the XCAT and OpenSim models are animated with real human swaying motion. This motion is recorded with a Vicon optical motion capture system (Vicon, Oxford, UK) tracking seven reflective markers attached to the subject’s body at a sampling rate of 120 Hz. The markers are placed on the sacrum, and on the right and left anterior superior iliac spine, lateral epicondyle of the knee and malleolus lateralis. Seven healthy subjects are recorded holding a squat at 30 and 60 degrees of knee flexion. Afterwards, a biomechanical model of the human lower body based on the model presented in [12] is scaled to each subject’s anthropometry using the software OpenSim, see Fig. 1b [11,20]. With this model and the measured 3D marker positions, the inverse kinematics are computed in order to find the generalized coordinates (i.e. global position and orientation of the pelvis and the joint angles) that best represent the measured motion. Before further processing, the generalized coordinates are filtered with a moving average filter with a span of 60 in order to remove system noise from the actual movement. The 3D positions of the sacrum, and the left and right hip joint center, knee joint center and ankle joint center over time are extracted from the animated model and used for the XCAT CT projection generation (Section 2.1). A virtual sensor is placed on the animated model’s shank in order to simulate IMU measurements (Section 2.2), which are used for the motion compensated reconstruction (Section 2.3).

2.1 Generation of motion corrupted projections

The XCAT numerical phantom is a model of the human body with its legs consisting of the bones tibia, fibula, femur and patella including bone marrow
and surrounding body soft tissue \cite{19}. The shapes of these structures are defined by non-uniform rational B-Splines (NURBS) and by the positions of their control points. By shifting these control points for each time step of the simulated CT scan based on the hip, knee and ankle joint centers of the OpenSim model, the upper leg and lower leg of the XCAT model are animated individually leading to a non-rigid motion. With the deformed model for each time step, X-ray projection images of a simulated CT scan are generated. The detector of size $620 \times 480$ pixels with isotropic pixel resolution of $0.616$ mm rotates on a virtual circular trajectory with a source detector distance of $1198$ mm and a source isocenter distance of $780$ mm. The angular increment between projections is $0.8$ degrees and in total $248$ projections are generated, corresponding to a sampling rate of $31$ Hz. Forward projections of the deformed model are created as described in \cite{17}. Since a healthy human's knees in a natural standing position are too far apart to both fit on the detector, the rotation center of the scan is placed in the center of the left leg.

## 2.2 Simulation of inertial measurements

An IMU is a small lightweight device that measures its acceleration and angular velocity on three perpendicular axes. Additionally, the accelerometer always also senses the earth’s gravitational field distributed on its three axes depending on the current orientation. Such a sensor is virtually placed on the shank $14$ cm below the left knee joint aligned with the shank segment. The simulated acceleration $a_i$ and angular velocity $\omega_i$ at time point $i$ are computed as follows \cite{7,10}:

\begin{equation}
\begin{aligned}
a_i &= R_i^\top (\ddot{r}_{\text{Seg},i} + \ddot{R}_i p_{\text{Sen},i} - g) \\
\omega_i &= (\omega_{x,i}, \omega_{y,i}, \omega_{z,i})^\top
\end{aligned}
\end{equation}

\begin{equation}
[\omega_i]_x = R_i^\top \dot{R}_i = \begin{pmatrix}
0 & -\omega_{z,i} & \omega_{y,i} \\
\omega_{z,i} & 0 & -\omega_{x,i} \\
-\omega_{y,i} & \omega_{x,i} & 0
\end{pmatrix}
\end{equation}

The $3 \times 3$ rotation matrix $R_i$ describes the orientation of the sensor at time point $i$ in the global coordinate system, $\dot{R}_i$ and $\ddot{R}_i$ are its first and second order derivatives with respect to time. The position of the segment the sensor was mounted on in the global coordinate system at time point $i$ is described by $r_{\text{Seg},i}$, with $\ddot{r}_{\text{Seg},i}$ being its second order derivative. $p_{\text{Sen},i}$ is the position of the sensor in the local coordinate system of the segment the sensor was mounted on. All required parameters are obtained by computing the forward kinematics of the biomechanical model. $g = (0, -9.80665, 0)^\top$ is the global gravity vector.

## 2.3 Motion compensated reconstruction

The simulated IMU measurements are used to estimate a rigid motion describing the 3D change of orientation and position from each time step to the next in the global coordinate frame. The sensor's coordinate system $S_i$ at each time step
Inertial Measurements for CT Motion Compensation

$i$, i.e. its orientation and position in the global frame is described by the affine matrix

\[ S_i = \begin{pmatrix} \hat{R}_i & \hat{t}_i \\ 0 & 1 \end{pmatrix}, \]  

(4)

where $\hat{R}_i$ is a $3 \times 3$ rotation matrix, and $\hat{t}_i$ is a $3 \times 1$ translation vector. The initial pose $S_0$ is assumed to be known.

The gyroscope measures the angular velocity, which is the change of orientation over time on the three axes of the sensor’s local coordinate system. Therefore, this measurement can directly be used to rotate the sensor from each time step to the next. The measured accelerometer signal, however, needs to be processed to obtain the positional change over time. First, the gravity measured on the sensor’s three axes is removed based on its global orientation. If $G_i$ is a 3D rotation matrix containing the rotation change measured by the gyroscope at time step $i$, the global orientation of the sensor is described by

\[ \hat{R}_{i+1} = \hat{R}_i G_i. \]  

(5)

The global gravity vector $g$ is transformed to the sensor’s local coordinate system at each time step $i$ using

\[ g_i = \hat{R}_i^\top g. \]  

(6)

To obtain the gravity-free acceleration $\bar{a}_i$, the gravity component then is removed from $a_i$ by adding the local gravity vector $g_i$ at each time step $i$. To obtain the sensor’s local velocity $v_i$, i.e. its position change over time, the integral of the gravity-free acceleration is computed. The integration must be performed considering the sensor’s orientation changes.

\[ v_{i+1} = G_i^\top (\bar{a}_i + v_i). \]  

(7)

In this study, we assume that the sensor’s initial velocity $v_0$ is known.

The local rotational change $\omega_i$ and positional change $v_i$ for each time step $i$ are linearly resampled to the CT scan’s sampling frequency and rewritten to an affine matrix

\[ \Delta_{l,i} = \begin{pmatrix} G_i & v_i \\ 0 & 1 \end{pmatrix}. \]  

(8)

To obtain the change in the global coordinate system, $\Delta_{l,i}$ is transformed for each time step using its pose $S_i$:

\[ \Delta_{g,i} = S_i \Delta_{l,i} S_i^{-1} \]  

(9)

\[ S_{i+1} = \Delta_{g,i} S_i \]  

(10)

For the motion compensated reconstruction, the first projection is used as reference without motion, so the affine matrix containing rotation and translation is the identity matrix $M_0 = I$. For each subsequent time step, the motion matrix is computed as

\[ M_{i+1} = M_i \Delta_{g,i}. \]  

(11)

These motion matrices are applied to the projection matrices of the system’s geometry to correct for motion.
2.4 Evaluation

We reconstruct the projections as volumes of size $512^3$ with isotropic spacing of 0.5 mm using GPU accelerated filtered back-projection in the software framework CONRAD [16].

To evaluate the performance of the proposed method, the resulting reconstructions are compared to a reconstruction without motion correction, and to a ground truth reconstruction from projections where the initial pose of the subject was kept static throughout the scan. Furthermore, we compare to reconstructions from the gold standard marker-based approach. For this purpose, small highly attenuating circular markers on the skin are simulated and tracked as proposed in [9]. All volumes are scaled from 0 to 1 and registered to the reference reconstruction without motion for comparison. As metrics for image quality evaluation we compute the root mean squared error (RMSE) and the structural similarity (SSIM). The SSIM index considers differences in luminance, contrast and structure for comparison and ranges from 0 (no similarity) to 1 (identical images) [24].

3 Results

Exemplary slices of the resulting reconstructions are shown in Fig. 2. A visual comparison of the results shows a similar reduction in streaking and blurring in the reconstructions of our proposed approach and the reference marker-based approach compared to the uncorrected case. The average RMSE and SSIM values of the proposed approach and the marker-based reference approach compared to ground truth excluding background voxels are similar, see Table 1. Both methods resulted in a high average SSIM value of 0.99 (proposed) and 0.98 (reference) for 30 degrees squats and 0.98 (proposed) and 0.97 (reference) for 60 degrees squats. The average RMSE of 0.02 for the proposed method was slightly lower than for the marker-based approach with an average 0.03. Table 2 shows the average improvement in percent compared to the uncorrected case. While the SSIM values show an improvement of 12-21% for both proposed and reference approach, the RMSE improved on average by 68-70% for the proposed method, and on average by 57% for the marker-based approach.

4 Discussion and Conclusion

The results presented in Fig. 2 and Table 1 and 2 show that the proposed method is able to estimate and correct for involuntary subject motion during a standing acquisition. Compared to the case without motion, some artifacts are still visible and some double edges could not be restored. A reason for this is that the motion applied for projection generation is a non-rigid motion, where the upper and lower leg can move individually. The inertial sensor placed on the shank, however, is only able to estimate a rigid motion consisting of a 3D rotation and translation. Therefore it is not possible to entirely restore image quality with
Fig. 2: Exemplary slices of a reconstructed volume of a 30 degrees squat. First row: axial slice through shank, second row: axial slice through thigh, third row: sagittal slice. (a) scan without motion, (b) uncorrected case, (c) proposed method, (d) marker-based reference method. The motion artifacts clearly visible in the uncorrected case can be reduced by both the proposed and the reference method.

one sensor, even though a clear improvement compared to the uncorrected case is observable. To overcome this limitation, a second sensor could be placed on the thigh and both measurements could be combined to account for non-rigid motion.

The reference approach tracking small metallic markers in the projection images for motion compensation also estimates a 3D rigid motion, thereby allowing for a fair comparison to our proposed approach. The presented evaluation even shows slightly better results for the proposed approach in a qualitative as well as in a quantitative comparison. Compared to the marker-based approach, where a sufficient number of markers has to be carefully placed around the knee, for our approach, only one or two sensors per leg are necessary to track the subject’s motion. This can help to facilitate and speed up the clinical process of weight-bearing imaging.

In this initial study, an optimal sensor in a well-controlled setting is assumed, however, in a real setting, some challenges will arise. One potential issue will be noise in the sensor signals. Furthermore, the gravitational signal is considerably
Table 1: Average structural similarity (SSIM) index and root mean squared error (RMSE) over all subjects for 30° and 60° squats. Best values are printed bold.

| SSIM | RMSE |
|------|------|
| 30°  | 60°  | 30°  | 60°  |
| uncorrected | 0.881 ± 0.053 | 0.821 ± 0.098 | 0.070 ± 0.015 | 0.081 ± 0.020 |
| proposed  | **0.989 ± 0.005** | **0.984 ± 0.013** | **0.022 ± 0.004** | **0.024 ± 0.009** |
| marker-based | 0.980 ± 0.008 | 0.969 ± 0.017 | 0.029 ± 0.004 | 0.034 ± 0.008 |

Table 2: Average improvement in percent of structural similarity (SSIM) index and root mean squared error (RMSE) compared to the uncorrected case over all subjects for 30° and 60° squats. Best values are printed bold.

| SSIM | RMSE |
|------|------|
| 30°  | 60°  | 30°  | 60°  |
| proposed  | **12.6 ± 6.7** | **21.2 ± 14.3** | **67.6 ± 7.1** | **70.4 ± 5.5** |
| marker-based | 11.6 ± 6.3 | 19.3 ± 14.0 | 56.9 ± 7.0 | 57.1 ± 7.4 |

larger than the motion to be estimated, which could pose a problem for an accurate estimation. To evaluate the influence of measurement errors on the proposed method, in a subsequent study, errors like noise or gyroscope bias will be included in the simulation.

For this study, we assumed that the initial pose and velocity of the sensor at the beginning of the C-arm scan are known, while in a real setting, these values have to be estimated. An approach for estimating the initial pose from the X-ray projection images was presented in [23]. If the sensor is placed close enough to the knee joint, it is also visible in the projection images without affecting image quality in the area of interest, and its metal components can be tracked. Using the system geometry, an average position over the scan can be obtained, which then can be refined by the first projection images. The initial velocity of the sensor could be estimated by assuming a zero mean velocity over the whole scan, since it has been shown that standing persons sway around their center of mass [1].

We presented a novel approach for motion compensation in weight-bearing C-arm CT using inertial measurements. This initial simulation is a first step towards a purely IMU-based compensation of motion in C-arm CT. We showed the feasibility of our approach which was able to improve image quality and achieve results similar to a state-of-the-art marker-based motion compensation.

**Acknowledgment** Bjoern Eskofier gratefully acknowledges the support of the German Research Foundation (DFG) within the framework of the Heisenberg professorship programme (grant number ES 434/8-1). The authors acknowledge funding support from NIH 5R01AR065248-03 and NIH Shared Instrument Grant No. S10 RR026714 supporting the zeego@StanfordLab.
References

1. Abrahamova, D., Hlavačka, F.: Age-related changes of human balance during quiet stance. Physiological Research 57(6) (2008)
2. Arden, N., Nevitt, M.C.: Osteoarthritis: Epidemiology. Best Practice & Research Clinical Rheumatology 20(1), 3 – 25 (2006)
3. Berger, M., Müller, K., Aichert, A., Unberath, M., Thies, J., Choi, J.H., Fahrig, R., Maier, A.: Marker-free motion correction in weight-bearing cone-beam CT of the knee joint. Medical physics 43(3), 1235-48 (2016)
4. Bier, B., Aichert, A., Felsner, L., Unberath, M., Levenston, M., Gold, G., Fahrig, R., Maier, A.: Epipolar Consistency Conditions for Motion Correction in Weight-Bearing Imaging. In: Maier-Hein, T.M.D.K. (ed.) Bildverarbeitung für die Medizin 2017 Algorithmen Systeme Anwendungen. pp. 209–214. Berlin (2017)
5. Bier, B., Katharina, A., Syben, C., Unberath, M., Levenston, M., Gold, G., Fahrig, R., Maier, A.: Detecting Anatomical Landmarks for Motion Estimation in Weight-bearing Imaging of Knees. In: Maier, A., Knoll, F. (eds.) Machine Learning for Medical Imaging Reconstruction. pp. 83–90 (2018)
6. Bier, B., Ravikumar, N., Unberath, M., Levenston, M., Gold, G., Fahrig, R., Maier, A.: Range Imaging for Motion Compensation in C-Arm Cone-Beam CT of Knees under Weight-Bearing Conditions. Journal of Imaging 4(1), 1–16 (2018)
7. van den Bogert, A.J., Read, L., Nigg, B.M.: A method for inverse dynamic analysis using accelerometry. Journal of Biomechanics 29(7), 949 – 954 (1996)
8. Choi, J.H., Fahrig, R., Keil, A., Besier, T.F., Pal, S., McWalter, E.J., Beaurp, G.S., Maier, A.: Fiducial marker-based correction for involuntary motion in weight-bearing c-arm ct scanning of knees. part i. numerical model-based optimization. Medical Physics 40(9), 091905–1 – 091905–12 (2013)
9. Choi, J.H., Maier, A., Keil, A., Pal, S., McWalter, E.J., Beaurp, G.S., Gold, G.E., Fahrig, R.: Fiducial marker-based correction for involuntary motion in weight-bearing c-arm ct scanning of knees. ii. experiment. Medical Physics 41(6), 061902–1 – 061902–16 (2014)
10. De Sapio, V.: Advanced Analytical Dynamics: Theory and Applications. Cambridge University Press (2017)
11. Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G.: Opensim: Open-source software to create and analyze dynamic simulations of movement. IEEE Transactions on Biomedical Engineering 54(11), 1940–1950 (2007)
12. Hamner, S.R., Seth, A., Delp, S.L.: Muscle contributions to propulsion and support during running. Journal of Biomechanics 43(14), 2709 – 2716 (2010)
13. Jost, G., Walti, J., Mariani, L., Cattin, P.: A novel approach to navigated implantation of s-2 alar iliac screws using inertial measurement units. Journal of Neurosurgery: Spine 24(3), 447–453 (2016)
14. Lemnammer, I., Michel, O., Ayasso, H., Zozor, S., Bernard, G.: Online mobile c-arm calibration using inertial sensors: a preliminary study in order to achieve cbct. International Journal of Computer Assisted Radiology and Surgery 15, 213224 (2019)
15. Maier, A., Choi, J.H., Keil, A., Niebler, C., Sarmiento, M., Fieselmann, A., Gold, G., Delp, S., Fahrig, R.: Analysis of vertical and horizontal circular C-arm trajectories. SPIE Medical Imaging 7961, 796123–1 – 796123–8 (2011)
16. Maier, A., Hofmann, H., Berger, M., Fischer, P., Schwemmer, C., Wu, H., Müller, K., Hornegger, J., Choi, J.H., Riess, C., Keil, A., Fahrig, R.: CONRAD - A software
framework for cone-beam imaging in radiology. Medical Physics 40(11), 111914 (2013)
17. Maier, A., Hofmann, H., Schwemmer, C., Hornegger, J., Keil, A., Fahrig, R.: Fast Simulation of X-ray Projections of Spline-based Surfaces using an Append Buffer. Physics in Medicine and Biology 57(19), 61936210 (2012)
18. Powers, C.M., Ward, S.R., Fredericson, M., Guillet, M., Shellock, F.G.: Patellofemoral Kinematics During Weight-Bearing and Non-Weight-Bearing Knee Extension in Persons With Lateral Subluxation of the Patella: A Preliminary Study. Journal of Orthopaedic & Sports Physical Therapy 33(11), 677–685 (2003)
19. Segars, W.P., Sturgeon, G., Mendonca, S., Grimes, J., Tsui, B.M.W.: 4d xcat phantom for multimodality imaging research. Medical Physics 37(9), 4902–4915 (2010)
20. Seth, A., Hicks, J.L., Uchida, T.K., Habib, A., Dembia, C.L., Dunne, J.J., Ong, C.F., DeMers, M.S., Rajagopal, A., Millard, M., Hammer, S.R., Arnold, E.M., Yong, J.R., Lakshmikanth, S.K., Sherman, M.A., Ku, J.P., Delp, S.L.: Opensim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. PLOS Computational Biology 14(7), 1–20 (2018)
21. Sisniega, A., Stayman, J.W., Yorkston, J., Siewerdsen, J.H., Zbijewski, W.: Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion. Physics in Medicine and Biology 62(9), 3712–3734 (2017)
22. Sisniega, A., Stayman, J., Cao, Q., Yorkston, J., Siewerdsen, J., Zbijewski, W.: Image-based motion compensation for high-resolution extremities cone-beam ct. Proceedings of SPIE–the International Society for Optical Engineering 9783, 97830K (2016)
23. Thies, M., Maier, J., Eskofier, B., Maier, A., Levenston, M., Gold, G., Fahrig, R.: Automatic orientation estimation of inertial sensors in C-Arm CT Projections. Current Directions in Biomedical Engineering 5, 195–198 (2019)
24. Zhou Wang, Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)