

Marker-free Motion Correction in Weight-Bearing Cone-Beam CT of the Knee Joint

M. Berger,^{1,*} K. Müller,² A. Aichert,¹ M. Unberath,^{1,†}

J. Thies,^{3,*} J.-H. Choi,² R. Fahrig,² and A. Maier^{1,†}

¹*Pattern Recognition Lab, Friedrich-Alexander-University*

Erlangen-Nuremberg, Erlangen, Germany

²*Radiological Sciences Lab, Stanford University, Stanford, CA, USA*

³*Computer Graphics Lab, Friedrich-Alexander-University*

Erlangen-Nuremberg, Erlangen, Germany

(Dated: February 22, 2016)

Abstract

Purpose: To allow for a purely image-based motion estimation and compensation in weight-bearing cone-beam computed tomography of the knee joint.

Methods: Weight-bearing imaging of the knee joint in a standing position poses additional requirements for the image reconstruction algorithm. In contrast to supine scans, patient motion needs to be estimated and compensated. We propose a method that is based on 2D/3D registration of left and right femur and tibia segmented from a prior, motion-free reconstruction acquired in supine position. Each segmented bone is first roughly aligned to the motion-corrupted reconstruction of a scan in standing or squatting position. Subsequently, a rigid 2D/3D registration is performed for each bone to each of K projection images, estimating $6 \times 4 \times K$ motion parameters. The motion of individual bones is combined into global motion fields using thin-plate-spline extrapolation. These can be incorporated into a motion-compensated reconstruction in the backprojection step. We performed visual and quantitative comparisons between a state-of-the-art marker-based (MB) method and two variants of the proposed method using gradient correlation (GC) and normalized gradient information (NGI) as similarity measure for the 2D/3D registration.

Results: We evaluated our method on four acquisitions under different squatting positions of the same patient. All methods showed substantial improvement in image quality compared to the uncorrected reconstructions. Compared to NGI and MB, the GC method showed increased streaking artifacts due to misregistrations in lateral projection images. NGI and MB showed comparable image quality at the bone regions. Because the markers are attached to the skin, the MB method performed better at the surface of the legs where we observed slight streaking of the NGI and GC methods. For a quantitative evaluation we computed the universal quality index (UQI) for all bone regions with respect to the motion-free reconstruction. Our quantitative evaluation over regions around the bones yielded a mean UQI of 18.4 for no correction, 53.3 and 56.1 for the proposed method using GC and NGI, respectively, and 53.7 for the MB reference approach. In contrast to our registration-based corrections, the MB reference method caused slight non-rigid deformations at bone outlines when compared to a motion-free reference scan.

Conclusion: We showed that our method based on the NGI similarity measure yields reconstruction quality close to the MB reference method. In contrast to the MB method, the proposed method does not require any preparation prior to the examination which will improve the clinical workflow and patient comfort. Further, we found that the MB method causes small, non-rigid deformations

at the bone outline which indicates that markers may not accurately reflect the internal motion close to the knee joint. Therefore, we believe that the proposed method is a promising alternative to MB motion management.

* Also with the Graduate School 1773 “Heterogeneous Image Systems”, Erlangen, Germany.

† Also with the Erlangen Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany

I. INTRODUCTION

Recently we proposed a method that allows for weight-bearing imaging of the knee-joint using a C-arm cone-beam CT (CBCT) that is usually operated in the interventional suite [1, 2]. The CBCT is mounted on a robotic arm and acquires volumetric images with high-spatial resolution and a relatively large field-of-view (FOV), using a horizontal trajectory around a standing patient [3]. Such weight-bearing scans pose a challenging reconstruction problem, as the patients' unsupported standing or squatting position induces involuntary motion of the knee joint. Initially, we used externally attached metallic markers that can be tracked in the projection images to estimate the motion [1, 4]. The markers yielded accurate reconstruction results and were able to remove most of the motion artifacts. However, in clinical routine the attachment of the markers will be bothersome as they need to be placed carefully to avoid overlaps in the projection images. Also, the internal motion of the joint might not be accurately reflected as the markers are attached to the skin. Finally, the markers cause metallic artifacts in the reconstruction which degrade image quality [5]. Therefore, a motion correction method which performs similarly well as the marker-based (MB) method but works directly on the acquired projection images is desirable. This could substantially reduce the preparation time for weight-bearing CT and increase patient comfort.

One method to estimate motion from projection images is 2D/2D registration. Here, the acquired projections are usually registered in 2D to digitally rendered radiographs (DRR). The DRR can be computed from an initial gated reconstruction [6, 7] or alternatively from previously acquired images. The estimated 2D deformations can be directly incorporated into a motion-compensated reconstruction.

Alternatively, the motion can be estimated directly in the volume domain using 2D/3D registration with a similarity measure that is defined over the projection images. Prior work for the knee-joint was done by Tsai et al. where they introduced a new similarity measure called weighted edge matching score (WEMS) [8]. WEMS matches edges extracted by a Canny edge detector incorporating increased weights for long edges. WEMS has also been used by Lin et al. to register a prior 3D MRI to a real-time 2D MRI slice during exercise using a custom weight-bearing apparatus [9]. A slightly different method for 2D/3D registration of bones in the knee- and shoulder-joint was presented by Zhu et al. [10, 11]. They first computed the outline of the forward projected vertices of a 3D mesh originating from a

42 prior segmentation. The outline is then registered to the 2D outline of the bone which was
43 segmented in the 2D projection image.

44 In skeletal 2D/3D registration, Wang et al. proposed a differential approach for regis-
45 tration of a prior thorax CT to fluoroscopic images [12]. Their method is also based on
46 2D contour segmentation, where they report increased accuracy in estimating longitudinal
47 off-plane motion. Bifulco et al. used the normalized cross-correlation (NCC) to register
48 prior 3D CT volumes of vertebrae to 2D fluoroscopic images [13, 14]. Otake et al. used
49 normalized gradient information (NGI) for registration of vertebra to a single projection
50 image [15]. NGI compares the 2D gradient directions and weighs them with the minimum
51 gradient magnitude. Yet, in more recent work [16] they decided to use the well known gra-
52 dient correlation (GC) measure as described by Penney et al. [17]. Only very little work has
53 been done in 2D/3D registration using a full set (> 100) of projection images from a CBCT
54 scan. Recently, Ouadah et al. [18] presented a method for image-based geometric calibration
55 of mobile C-arm systems using the NGI similarity measure and a statistical optimization.
56 Overall, 9 parameters were estimated for each of the 360 projection images yielding a total
57 of 3240 parameters.

58 In weight-bearing knee-joint imaging we neither have access to surrogate signals such as
59 ECG in cardiac imaging nor can we rely on motion periodicity. Thus, many 2D/2D and
60 2D/3D registration methods known in literature are not applicable. Furthermore, methods
61 based on implicit segmentations in the projection domain, such as WEMS or the work by
62 Zhu et al. [10, 11], work only on single- or pairs of projection images which can usually be
63 positioned such that there is only little overlap with background structures. We scan both
64 knees using a horizontal trajectory, i.e., overlap of tibia, femur, fibula, patella and the skin
65 boundary is inevitable which make these methods hardly applicable.

66 In a first attempt of image-based motion estimation we used a simple 2D/2D registration
67 for marker-free weight-bearing CBCT projection images of the knee-joint [19]. The DRR
68 was calculated by maximum intensity projections (MIP) of an initial, motion-corrupted
69 reconstruction. In the context of this work, we estimated 2D translations for each pair
70 of projection images using a similarity measure based on mutual information (MI) and a
71 gradient-descent optimizer. The approach did not require any additional or prior data and
72 improved the image quality compared to a reconstruction without correction. However, we
73 still observed large discrepancies from the image quality achieved with the MB approach.

To further improve our marker-free motion compensation we now make use of an already existent, motion-free CBCT reconstruction of the knee-joints acquired in supine position. This enables rigid 2D/3D registration of the segmented left and right tibia and femur from the motion-free scan. All four segmentations are rigidly registered to each of the K motion-corrupted projection images, for example for $K = 248$ yielding a total of $4 \times 6 \times 248 = 5952$ parameters. The estimated motion between the acquired projection images and the static 3D reference is then incorporated into a motion-compensated FDK-type reconstruction to obtain a corrected 3D volume. A first version of the proposed method has already shown promising results when evaluated on a numerically simulated dataset as shown in Berger et al. [20].

II. METHODS AND MATERIALS

A. Reference Method Using Metallic Fiducial Markers

To thoroughly evaluate the proposed method we compare it with the motion-compensated reconstruction that is based on externally attached fiducial markers. In the following, we briefly summarize the MB motion estimation method. For details, we refer to the work of Müller et al. [4], which combines the MB motion estimation presented by Choi et al. [1] with the automatic marker detection and removal presented by Berger et al. [5]. The optimization problem for K projection images and M markers is given by

$$\arg \min_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}) = \arg \min_{\boldsymbol{\alpha}} \frac{1}{2} \sum_{k=1}^K \sum_{i=1}^M \|h(\mathbf{n}) - \mathbf{u}_{ik}\|_2^2 \quad (1)$$

$$\mathbf{n} = \begin{pmatrix} n_1 & n_2 & n_3 \end{pmatrix}^T = \mathbf{P}_k \cdot \mathbf{T}_k(\boldsymbol{\alpha}) \cdot \begin{pmatrix} \mathbf{x}_i & 1 \end{pmatrix}^T ,$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{6K}$ is a vector containing three rotation and translation parameters per projection, $\mathbf{P}_k \in \mathbb{R}^{3 \times 4}$ is the k -th projection matrix as given by the system's calibration, $\mathbf{T}_k(\boldsymbol{\alpha}) \in \mathbb{R}^{4 \times 4}$ applies the rigid motion for projection k given the parameters in $\boldsymbol{\alpha}$, $\mathbf{x}_i \in \mathbb{R}^3$ is the 3D reference position of the i -th marker and \mathbf{u}_{ik} the i -th marker's measured 2D position in projection k . Further, \mathbf{n} is the homogeneous representation of the motion-compensated and projected 3D reference position \mathbf{x}_i . The function $h : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ describes the mapping of \mathbf{n} to 2D pixel coordinates. It is defined by $h(\mathbf{n}) = \begin{pmatrix} n_1 & n_2 \\ n_3 & n_3 \end{pmatrix}^T$. Thus, by adjusting the motion parameters we minimize the distance between forward projected 3D reference positions and

measured 2D marker positions over all detected markers and all projections. Marker detection in the projection-domain is done by applying the Fast-Radial-Symmetry-Transform (FRST) [21] with subsequent thresholding and center-point detection. The 3D reference positions for each marker are obtained automatically. First, we apply a Gaussian filter to the 2D FRST result yielding blob-like structures at marker locations. These images are then backprojected to 3D, resulting in high-intensity 3D blobs where the backprojected 2D blobs overlap. These blobs are then segmented using a maximum entropy thresholding. Finally, the 3D reference positions are extracted from the blobs' centroids using a 3D connected-components analysis. The assignment of 2D detections to 3D reference positions is given by the smallest Euclidean distance of forward projected 3D reference and 2D detection. For more details on 2D marker detection and 3D reference point extraction we refer to our previous work [5].

As a new feature, we introduce an analytic gradient computation of the cost-function which reduced the algorithm's run-time drastically compared to a forward-differences type gradient estimation used in Müller et al. [4] and Choi et al. [1]. Using the chain rule for multivariate functions the partial derivatives of Eq. 2 with respect to the individual parameters is given by

$$\frac{\partial f(\boldsymbol{\alpha})}{\partial \alpha_j} = \sum_k \sum_i (h(\mathbf{n}) - \mathbf{u}_{ik})^T \cdot J_h(\mathbf{n}) \cdot \mathbf{P}_k \frac{\partial \mathbf{T}_k}{\partial \alpha_j} \begin{pmatrix} \mathbf{x}_i \\ 1 \end{pmatrix}, \quad (2)$$

where $J_h(\mathbf{n})$ denotes the Jacobian of function h

$$J_h(\mathbf{n}) = \begin{pmatrix} \frac{1}{n_3} & 0 & -\frac{n_1}{n_3^2} \\ 0 & \frac{1}{n_3} & -\frac{n_2}{n_3^2} \end{pmatrix}. \quad (3)$$

For outlier detection and removal we applied an iterative removal of worst contributions. After optimization we find those \mathbf{u}_{ik} that belong to the 0.5% highest 2D distances with respect to their forward-projected reference point. They were then removed from the measurements using the following rules: 1) only remove one detection per projection and 2) only remove if at least M_{min} detections are left for this projection. This process is repeated iteratively for J times.

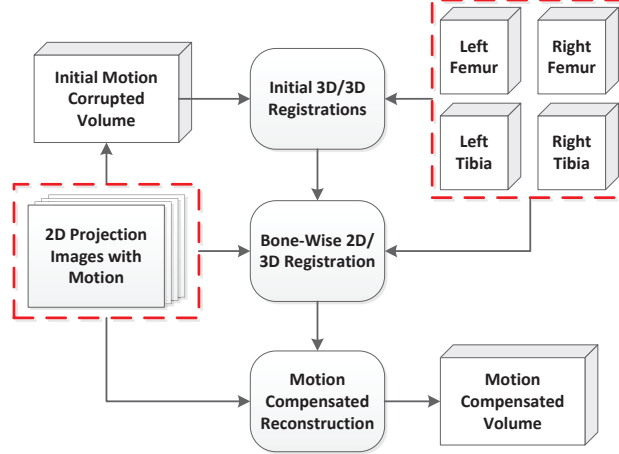


Figure 1. Overview of the proposed motion compensation approach. The inputs are the femur and tibia volumes and the 2D projection images, both marked by a dashed frame.

B. Motion Compensation Using 2D/3D Registration

Our method is based on 2D/3D registration of segmentations from a prior, motion-free reconstruction acquired in supine position. To limit complexity of the optimization problem, we focus on four bones that represent both knee-joints, i.e., left and right femur and tibia. An overview of the proposed method is given in Figure 1. As input we have a stack of projection images acquired under weight-bearing conditions along with the segmented femur and tibia volumes, both emphasized by a dashed box. First we perform a motion-corrupted reconstruction of the acquired projections. Then, a 3D/3D registration of each segmented bone volume to the motion-corrupted reconstruction to align the standing and supine coordinate system, and to account for different positions of the bones among one another. Note, that this global 3D/3D alignment does not include any estimation for intra-scan patient motion. Subsequently, the initial 3D/3D registration results are used as initialization for the 2D/3D registration. 2D/3D registration is performed between each bone and every acquired projection image using a rigid motion model. The result of the 2D/3D registration is the individual bone motion over the acquisition time. To perform a motion-compensated backprojection we also need to extrapolate the motion that occurs in between bones, e.g., at muscle or skin tissue. We use a thin-plate-spline (TPS) extrapolation method as explained in Müller et al. [22].

The accuracy of the 2D/3D registration is the important step for a successful motion

144 compensation. A crucial component is the similarity measure used to compare DRR images
 145 with the acquired projections. In this work, we assess the influence of two different similarity
 146 measures on the reconstructed image quality. First, we apply the GC that computes the
 147 NCC between vertical and horizontal gradient images. Further, we use the NGI measure
 148 that is also gradient-based but has been reported to be more stable against outlier intensities
 149 and thus more performant in presence of overlapping structures [15]. To avoid convolution-
 150 based gradient computation for every DRR we create gradient DRRs directly by ray-tracing
 151 through the precomputed 3D volume gradients [23, 24].

152 1. Gradient Correlation (GC)

153 GC is a state-of-the-art similarity measure and has been widely used to register bones
 154 to their projection images. For the initial formulation of GC we refer to Penney et al. [17].
 155 Let $\nabla p_k(\mathbf{u}; \boldsymbol{\gamma}) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the DRR's gradient and $\mathbf{u} \in \mathbb{R}^2$ a 2D pixel location. Further,
 156 $\nabla b_k(\mathbf{u}) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the gradient of the k -th acquired projection image computed using the
 157 Sobel operator. In contrast to the MB approach, the parameter vector is now $\boldsymbol{\gamma} \in \mathbb{R}^{6KL}$,
 158 containing 6 rigid parameters for each of the K projections and each of the L segmented
 159 bone volumes. The GC can be formulated as

$$GC(p, b; \boldsymbol{\gamma}) = \frac{1}{2} \sum_{k=1}^K \sum_{\mathbf{u} \in \Omega_k} (\nabla p_k(\mathbf{u}; \boldsymbol{\gamma})^T \mathbf{W}^{-1} \nabla b_k(\mathbf{u})) \quad . \quad (4)$$

160 In our formulation the normalization, i.e. a division by the standard deviations, is incor-
 161 porated into the weighting matrix $\mathbf{W} \in \mathbb{R}^{2 \times 2}$ given in Eq. 5.

$$\mathbf{W} = \begin{pmatrix} \sqrt{\sum_{\mathbf{u}} \left(\frac{\partial p_k(\mathbf{u}; \boldsymbol{\gamma})}{\partial u_1} \right)^2} & 0 \\ 0 & \sqrt{\sum_{\mathbf{u}} \left(\frac{\partial b_k(\mathbf{u})}{\partial u_1} \right)^2} \\ \sqrt{\sum_{\mathbf{u}} \left(\frac{\partial p_k(\mathbf{u}; \boldsymbol{\gamma})}{\partial u_2} \right)^2} & \sqrt{\sum_{\mathbf{u}} \left(\frac{\partial b_k(\mathbf{u})}{\partial u_2} \right)^2} \end{pmatrix} \quad (5)$$

162 The set Ω_k defines the image region used for the computation of the GC measure, which
 163 may vary for each projection image as indicated by the subscript k . The normalization \mathbf{W}
 164 is used to adjust intensity differences and depends on the region Ω_k . During our experi-
 165 ments we set Ω_k such that it contains every non-zero gradient value of the DRR image, i.e.,
 166 $\Omega_k = \{\mathbf{u} \mid \|\nabla p_k(\mathbf{u}; \boldsymbol{\gamma})\|_2 > 0\}$.

2. Normalized Gradient Information (NGI)

The idea behind NGI is to compare the similarity of gradient directions at each pixel position. This is done by computing the cosine of the angle between the gradient directions, followed by a weighting with the pixel's gradient magnitude. To improve robustness against outliers, at each pixel the minimum gradient magnitude of DRR and acquired image is used as a weighting factor for that pixel location. This scheme was described to be more robust against intensity outliers and thus overlapping structures [15]. In contrast to GC the NGI does not perform an intensity normalization and therefore intensities of the DRR image need to be adjusted heuristically. For more information we refer to Otake et al. [15]. The NGI can be formulated as

$$NGI(p, b; \gamma) = \frac{GI(p, b; \gamma)}{GI(b, b)} , \quad (6)$$

with the variable measure

$$GI(p, b; \gamma) = \sum_{k=1}^K \sum_{\mathbf{u} \in \Omega_k} \left(\frac{1}{2} \frac{\nabla p_k(\mathbf{u}; \gamma)^T \nabla b_k(\mathbf{u})}{\|\nabla p_k(\mathbf{u}; \gamma)\|_2 \|\nabla b_k(\mathbf{u})\|_2} + \frac{1}{2} \right) \times \min(\|\nabla p_k(\mathbf{u}; \gamma)\|_2, \|\nabla b_k(\mathbf{u})\|_2) \quad (7)$$

and the constant normalization

$$GI(b, b) = \sum_{k=1}^K \sum_{\mathbf{u} \in \Omega_k} \|\nabla b_k(\mathbf{u})\|_2 . \quad (8)$$

Because the gradient magnitude of the DRR equals zero outside the projected bone volume, we can set the region Ω_k such that it covers the full image domain for all k .

3. Regularization, Cost-Function and Optimization

We assume that the variation of all six motion parameters is physically limited given the knee-joint anatomy. Therefore, we add a temporal smoothness regularizer to our cost-function. We minimize the energy of the difference of the estimated parameters and their Gaussian filtered parameters. This can be understood as a minimization of energies present

186 in high temporal frequencies.

$$\gamma = \left(\zeta_{11}^T, \dots, \zeta_{1K}^T, \zeta_{21}^T, \dots, \zeta_{LK}^T \right)^T \quad (9)$$

$$\begin{aligned} \zeta_{lk} &= \left(\phi_x, \phi_y, \phi_z, t_x, t_y, t_z \right)_{lk}^T \\ r(\gamma) &= \sum_{l=1}^L \sum_{k=1}^K \|\zeta_{lk} - (\zeta * g^\sigma)_{lk}\|_2^2 \end{aligned} \quad (10)$$

187 In Eq. 9, we outline the structure of the parameter vector γ , where $\zeta_{lk} \in \mathbb{R}^6$ holds the Euler
188 angles ϕ_x, ϕ_y, ϕ_z and the translations t_x, t_y, t_z for the l -th bone and the k -th projection. Eq. 10
189 shows the smoothness regularizer, where g^σ are Gaussian filter coefficients for standard
190 deviation σ and $(\zeta * g^\sigma)$ denotes the convolution filtering over the temporal direction k .

191 The overall optimization problem is then given by

$$\arg \min_{\gamma} -c(p, b; \gamma) + \lambda r(\gamma), \quad (11)$$

192 where $c(p, b; \gamma)$ can be either $GC(p, b; \gamma)$ or $NGI(p, b; \gamma)$.

193 Both GC and NGI need to be maximized. This is achieved by minimization of the neg-
194 ative cost-function value. We use a non-constrained gradient-based minimization method
195 for optimization. The gradient is estimated by forward-differences and the Hessian is ap-
196 proximated using BFGS. The step-direction is then computed by attempting a Newton step
197 using the approximated Hessian. The step-size is calculated by a line-search method. We
198 optimize the rotational parameters in degrees instead of radians to ensure that rotation
199 and translation parameters are in a similar range. For more information we refer to the
200 optimizers documentation [25].

201 4. Noise Reduction in DRRs

202 We observed a high amount of noise in the forward projected gradient images which led
203 to unsatisfactory registration results. As the noise originates from the segmented volume
204 we applied a 3D edge-preserving bilateral filter as described by Lorch et al. [26] before
205 calculation of the 3D gradient volumes. Additionally, we observed that the trabecular bone
206 and the bone marrow show rather homogeneous intensities and hence, contain only little
207 structural information that is useful for 2D/3D registration. Therefore, the segmentation
208 masks were adjusted such that they focus on cortical bone, i.e. the outline of the bone. This

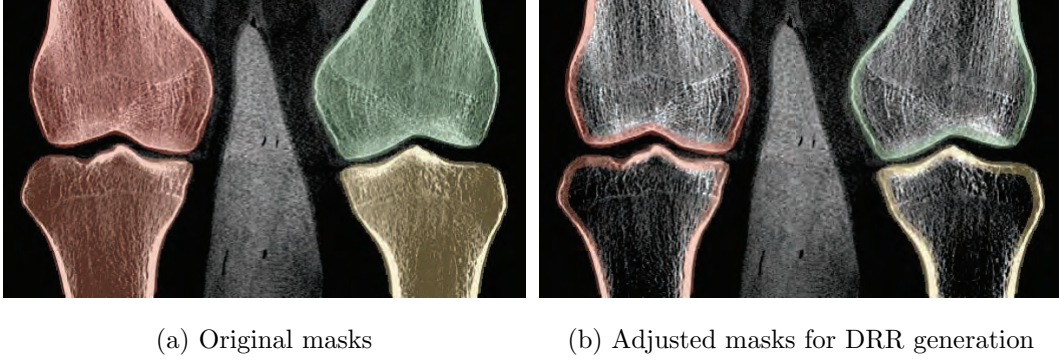


Figure 2. Adjustment of 3D segmentation masks to reduce the noise level in DRRs.

was done by first applying a 3D erosion to the segmentation masks and in another step a 3D dilation. Subtracting the eroded from the dilated mask results in a mask which contains the bone outline only. This process is illustrated in Figure 2.

5. Unified Coordinate System, TPS Estimation and Reconstruction

After 2D/3D registration, we know the individual bone motion over time. As a next step, we extrapolate a global non-rigid motion field $\mathbf{d}(\mathbf{x}, k) : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ based on the rigid bone motion using a TPS model [22]. We use the vertices of the segmented surface meshes as known TPS control points. The deformation field is estimated for each projection image separately. $\mathbf{I}_l \in \mathbb{R}^{4 \times 4}$ contains the rigid motion that was estimated by initial manual 3D/3D registration between supine and motion-corrupted scan. Further, $\hat{\mathbf{T}}_{lk} \in \mathbb{R}^{4 \times 4}$ are the final rigid motion matrices obtained from the 2D/3D registration. After selection of a reference time point \hat{k} we can express the rigid alignment to a common coordinate system by $\mathbf{R}_l = \hat{\mathbf{T}}_{lk}$. We choose \hat{k} to correspond to an anterior-posterior viewing direction, where no overlapping bones are present. First, we adjust all rigid transformations such that they operate in the reference coordinate system, i.e.

$$\mathbf{T}_{lk} = \hat{\mathbf{T}}_{lk} \mathbf{R}_l^{-1} . \quad (12)$$

Subsequently, all supine mesh vertices are propagated to the standing reference coordinate system. Let $\hat{\mathbf{v}}_l^n = \hat{\mathbf{v}}_l \in \mathbb{R}^3$ be the n -th vertex of the l -th bone in the supine coordinate system. In the following we omit the superscript n assigned to each vertex to improve clarity. Then, its static reference position $\mathbf{v}_l \in \mathbb{R}^3$ for the reference time point \hat{k} can be

calculated as is given by

$$\underline{\mathbf{v}}_l = (\mathbf{R}_l \mathbf{I}_l) \hat{\underline{\mathbf{v}}}_l , \quad (13)$$

where $\underline{\mathbf{v}} \in \mathbb{R}^4$ represents a 3D point in homogeneous coordinates. Finally, we apply the updated matrices \mathbf{T}_{lk} to the standing reference positions

$$\underline{\mathbf{v}}_{lk} = \mathbf{T}_{lk} \underline{\mathbf{v}}_l . \quad (14)$$

According to Davis et al. [27], the TPS deformation at a point $\mathbf{x} \in \mathbb{R}^3$ can be formulated by

$$\mathbf{d}(\mathbf{x}, k) = \sum_{n=1}^N \sum_{l=1}^L \mathbf{G}(\mathbf{x} - \mathbf{v}_{lk}) \mathbf{c}_{lk} + \mathbf{A}_k \mathbf{x} + \mathbf{b}_k , \quad (15)$$

where $\mathbf{c}_{lk} \in \mathbb{R}^3$ are the unknown spline coefficients and the matrix $\mathbf{A}_k \in \mathbb{R}^{3 \times 3}$ and vector $\mathbf{b}_k \in \mathbb{R}^3$ are additional rigid motion parameters. The kernel matrix $\mathbf{G}(\tilde{\mathbf{x}}): \mathbb{R}^3 \rightarrow \mathbb{R}^{3 \times 3}$ for a 3D deformation is given by:

$$\mathbf{G}(\tilde{\mathbf{x}}) = \|\tilde{\mathbf{x}}\|_2 \cdot \mathbf{I} , \quad (16)$$

where $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ denotes the identity matrix. To train the TPS model we need to determine the unknown coefficients \mathbf{c}_{lk} , \mathbf{A}_k and \mathbf{b}_k . As they have a linear relationship within Eq. 15, they can be estimated in a straightforward manner. Inserting the known 2D/3D motion vectors of the vertices, i.e. $\mathbf{u}_{lk} = \mathbf{v}_{lk} - \mathbf{v}_l$, for \mathbf{x} in Eq. 15 yields a system of linear equations which can be solved by singular value decomposition. To constrain the spline deformation at the periphery of the reconstruction we also added the 8 bounding box corners as control points for each volume. As displacement vector we assigned the motion generated by the geometrically closest bone, e.g., for the upper-left corners we applied the estimated left femur motion. For more details on solving the TPS equations we refer to Davis et al. [27].

The reconstruction pipeline includes the following steps: 1) a simplified beam-scatter-kernel scatter estimation [28] assuming that the object consists only of water and that the water-equivalent-thickness is uniform, 2) cosine weighting, 3) Parker redundancy weighting [29], 4) a simple truncation correction [30], ramp filtering with a smooth Shepp-Logan kernel [31] and a motion-compensated GPU backprojector [32]. The deformation field $\mathbf{d}(\mathbf{x}, k)$ was incorporated into the GPU-based backprojection step as described by Schäfer et al. [33]. That means, we evaluate the TPS model in Eq. 15 for each voxel coordinate and use the updated instead of the original coordinate to compute the 2D detector location. It should

be noted, that this type of reconstruction algorithm is approximate as it cannot guarantee a correct filtering and weighting of the projection images [33].

C. Evaluation Procedure

In this work, we compare the state-of-the-art MB approach and two versions of the proposed motion correction based on 2D/3D registration. We chose to focus on the reconstruction image quality, as this includes all possible steps and parameters of the individual approaches and also validates clinical applicability. To improve early diagnosis of osteoarthritis, we want to investigate the change in joint space under weight-bearing conditions. The joint space in the knee is defined between the femoral and tibial bone surfaces [34]. As the bones will be directly involved in the measurement process, we are especially interested in correcting motion at the distal femur and the proximal tibia. We tailor our evaluation pipeline accordingly and focus on the improvement of bone structure when applying our motion correction methods.

1. Data Acquisition and Parameter Selection

We evaluated our method on four acquisitions of the same patient, where large motion was present in two standing scans. The study included: 1) one motion-free scan in supine position with high angular resolution, 2) a standing scan with an upright stand, 3) a standing scan with 35° knee flexion and 4) a standing scan with 60° knee flexion. The motion severity increased with the flexion angle. Supine scanning took 20s acquiring 496 projection images over 200°, whereas the standing scans took 10s with 248 projection images over the same angular range. The detector size was 1240×960 pixels with an isotropic pixel size of 0.308mm. For the supine and initial motion-corrupted data we reconstructed a 512×512×256 volume with isotropic resolution of 0.5mm, using the same preprocessing steps as described for the motion-compensated reconstruction. A total of 12 metallic beads were attached to the skin at both knees. We set the minimum number of beads to be detected per projection to $M_{min} = 6$ and repeated the optimization for $J = 4$ times. The 2D/3D registration was applied using a two-fold multi-resolution, where only the resolution of the projection images was adjusted. For the first optimization, we used a projection image size of 310 × 240. In

the second step we initialized the parameters with the results from the first optimization and used a size of 620×480 . The weighting factor $\lambda = 5 \cdot 10^3$, as well as the standard deviation of the Gaussian smoothing $\sigma = 2$ have been determined heuristically and were kept constant for all experiments. Further, we ensured that the initial cost-function values for GC and NGI are within the same range by incorporating a normalization factor.

The 2D/3D registration and the marker-based approach have been implemented using CONRAD, a dedicated and open-source software platform for CBCT reconstruction [35]. Function evaluations for the 2D/3D registration are entirely done on the GPU using OpenCL. The segmentation was done using ITK-Snap [36], and initial 3D/3D registration was done manually using 3D Slicer [37]. To reduce artifacts due to detector saturation the patient’s legs were wrapped in a layer of plasticine which can be seen clearly in the reconstructed volumes (see Choi et al. [1] for more information).

2. Quantitative Evaluation

We conducted an image-based quantitative comparison between the supine, motion-free and the standing, motion-corrected reconstructions by computing the universal image quality index (UQI) [38]. A key problem for quantitative evaluation is that there is no unified coordinate system for the motion-corrected reconstruction. Depending on which reference projection index \hat{k} is chosen in Eq. 12, the final reconstruction will represent a different motion state. Assuming that the motion parameters for 2D/3D registration are estimated perfectly, the alignment between corrected and supine reconstruction should still work accurately. However, small errors in the 2D/3D registration for the reference projection \hat{k} lead to a piecewise rigid motion that is based on only $6 \times 4 = 24$ parameters of the $6 \times 4 \times 248 = 5952$ estimated parameters. This misalignment would dominate the image-based measures. Our main interest, however, is the improvement in the actual image quality. To become independent of this offset, we applied an automated 3D/3D rigid registration for each bone to the reference reconstruction using 3D Slicer [37] and evaluated the image quality for each bone region separately. For the rigid 3D/3D registration we used an MI-based similarity measure that ensured proper alignment even in presence of motion artifacts. For registration and UQI computation we used bone-wise region-of-interests (ROIs) as depicted in Figure 3. We made sure that the ROIs include a soft tissue margin around the bone. This was done by

312 dilating the segmentation masks in x and y direction with a circular structuring element of
 313 radius 2.5mm. We excluded the z-direction from the dilation as this would have caused an
 314 overlap of different bones in the ROIs. Details are given in overview 1.

Overview 1 Quantitative Evaluation Pipeline

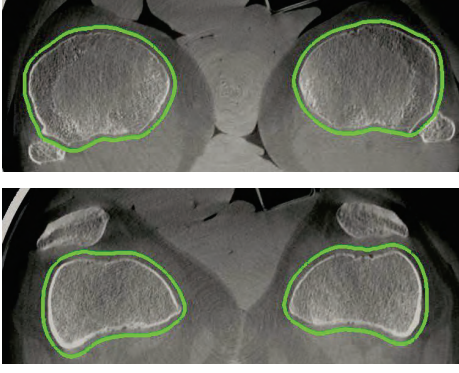
- 1: Corrected reconstruction with respect to reference projection \hat{k}
 - 2: Extraction of bone-wise ROI using dilated segmentation masks
 - 3: **for** each ROI **do**
 - 4: 3D/3D registration of ROI to supine volume
 - 5: New reconstruction including the registration result to avoid additional interpolation
 - 6: Computation of the UQI for each bone ROI
 - 7: **end for**
 - 8: Construction of mean and standard deviation of bone-wise measures
-

315 3. *Target Registration Error*

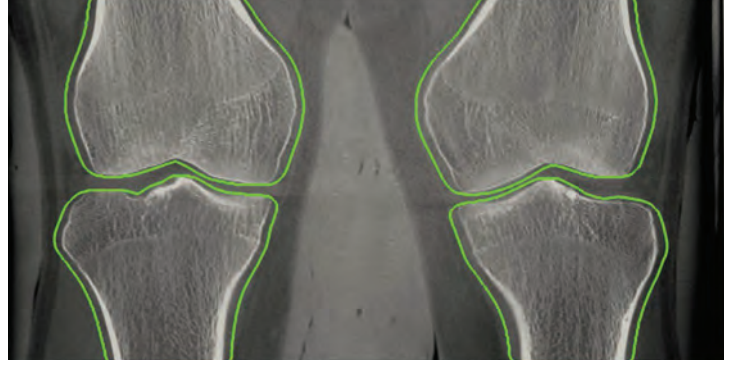
316 To evaluate how accurate the TPS extrapolation can model the motion in a certain dis-
 317 tance to the bone we calculated the Target Registration Error (TRE) using the attached
 318 metallic beads. To do so, we first detect the 3D bead locations in the corrected recon-
 319 structions using the same automatic detection approach as for the MB method, except that
 320 the backprojection now involves the estimated TPS motion fields. Then we apply the TPS
 321 motion field to the detected 3D point locations and project the resulting point to the indi-
 322 vidual projection images. The TRE can now be computed by the mean distance between
 323 the detected bead locations in 2D and the reprojections of the 3D bead locations. To avoid
 324 wrong assignments of 2D and 3D points we used the correspondences as determined after
 325 the outlier detection of the MB method.

326 4. *Relative Bone Motion*

327 To assess the amount of non-rigidity we measured the relative bone motion during a
 328 scan, i.e. how much the rigid motions of the bones deviate from each other. All estimated
 329 rigid motion matrices \mathbf{T}_{lk} of the 2D/3D registration are relative to their individual manual
 330 initialization \mathbf{I}_l , thus, a direct comparison of motion parameters will be difficult. As a first



(a) Supine - Axial



(b) Supine - Coronal

Figure 3. Axial and coronal slices of the motion-free supine data. The green line corresponds to the regions used for the numerical evaluation in section III B. The images show a clear reconstruction of the bones without any apparent motion artifacts.

step we remove the mean rigid transform $\overline{\mathbf{T}}_l$ over all time steps k , from all \mathbf{T}_{lk} , yielding the temporal, mean-free rigid transforms $\Psi_{lk} = \mathbf{T}_{lk} \overline{\mathbf{T}}_l^{-1}$. In case that all bones move with the same rigid transform their deviation to the mean rigid transform over the bones, i.e. $\overline{\Psi}_k$, is the identity matrix. To visualize the relative motion between the bones we decided to compute the differences to their mean rigid transform $\Delta\Psi_{lk} = \Psi_{lk} \overline{\Psi}_k^{-1}$. Finally, we extract all angles $(\Delta\phi_x \ \Delta\phi_y \ \Delta\phi_z)_{lk}$ and translations $(\Delta t_x \ \Delta t_y \ \Delta t_z)_{lk}$ from $\Delta\Psi_{lk}$.

III. RESULTS

A. Visual Comparison

For the 0° flexion angle, we observed only little motion artifacts. Slight streaking is present at the outline of femur and patella (Figure 4a) but also at tibia and fibula (Figure 5a). All three methods were able to restore the bone outlines, yielding similar visual results, shown in Figure 4(b)-(d) and Figure 5(b)-(d). Yet, the MB approach shows a slightly sharper correction of the fibula's interior compared to GC and NGI as indicated by the arrow in Figure 5d. Note the zoomed version of the reconstructed marker shown in the embedded box in the lower, right corner of the images. As expected the marker was accurately reconstructed using the MB method but also the NGI method did not substantially distort the marker's appearance, indicating a good estimation of the motion at the skin boundary. A little more

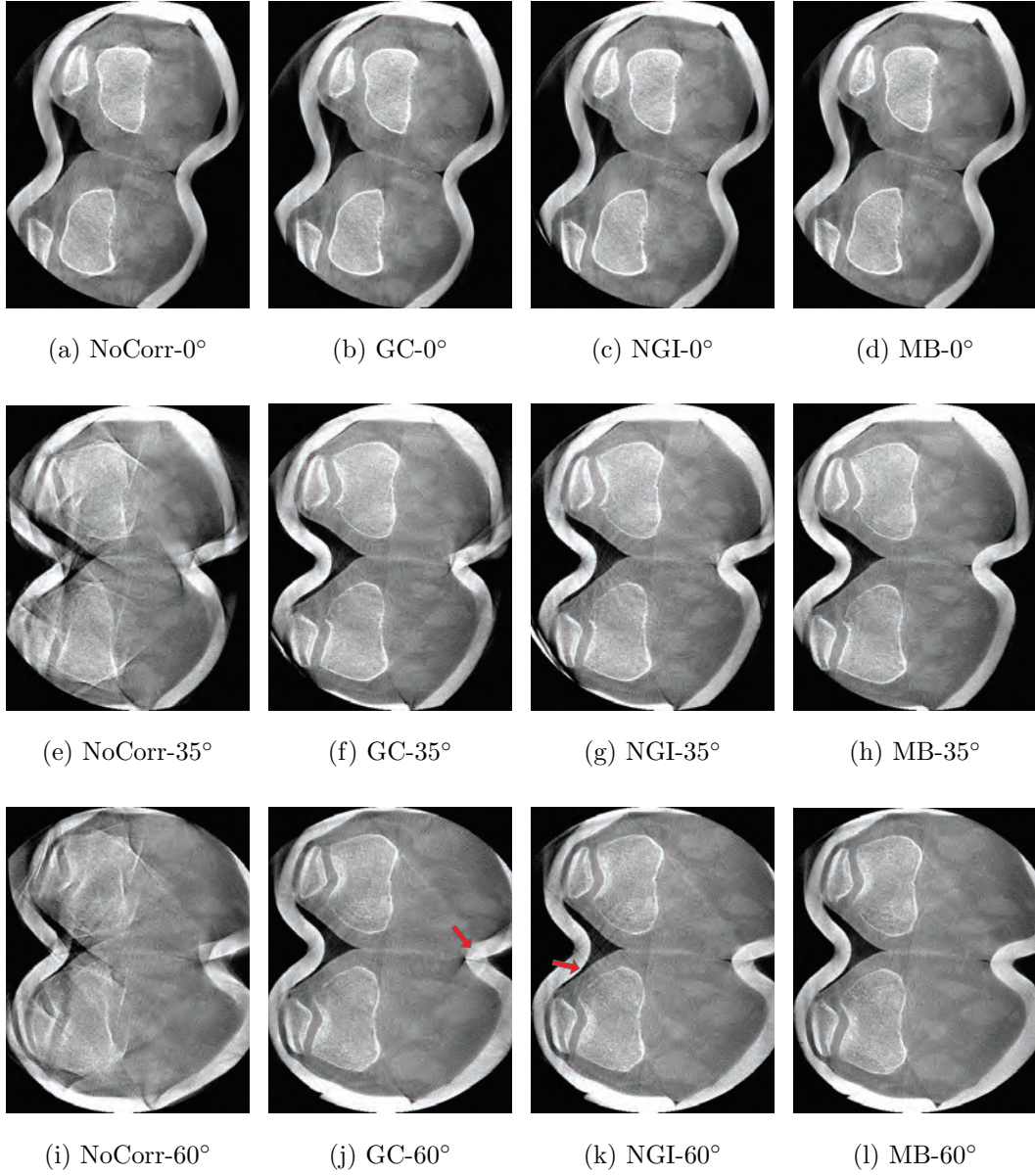


Figure 4. Axial slices through the femur. From left to right: Reconstructions without motion correction (NoCorr), the proposed method using GC, the proposed method using NGI and the MB reference method. The rows correspond to the three different weight-bearing scans from 0° flexion angle at the top to 60° flexion angle at the bottom (W: 2025 HU, C: 145 HU).

distortion (i.e., star-like appearance) can be seen at the marker for the GC case.

All methods could substantially reduce the large amount of motion artifacts for the 35° case. Not many differences are seen at bone outlines between the results for the femur slices in Figure 4(f)-(h). As expected, the MB method shows a better result at the skin boundary

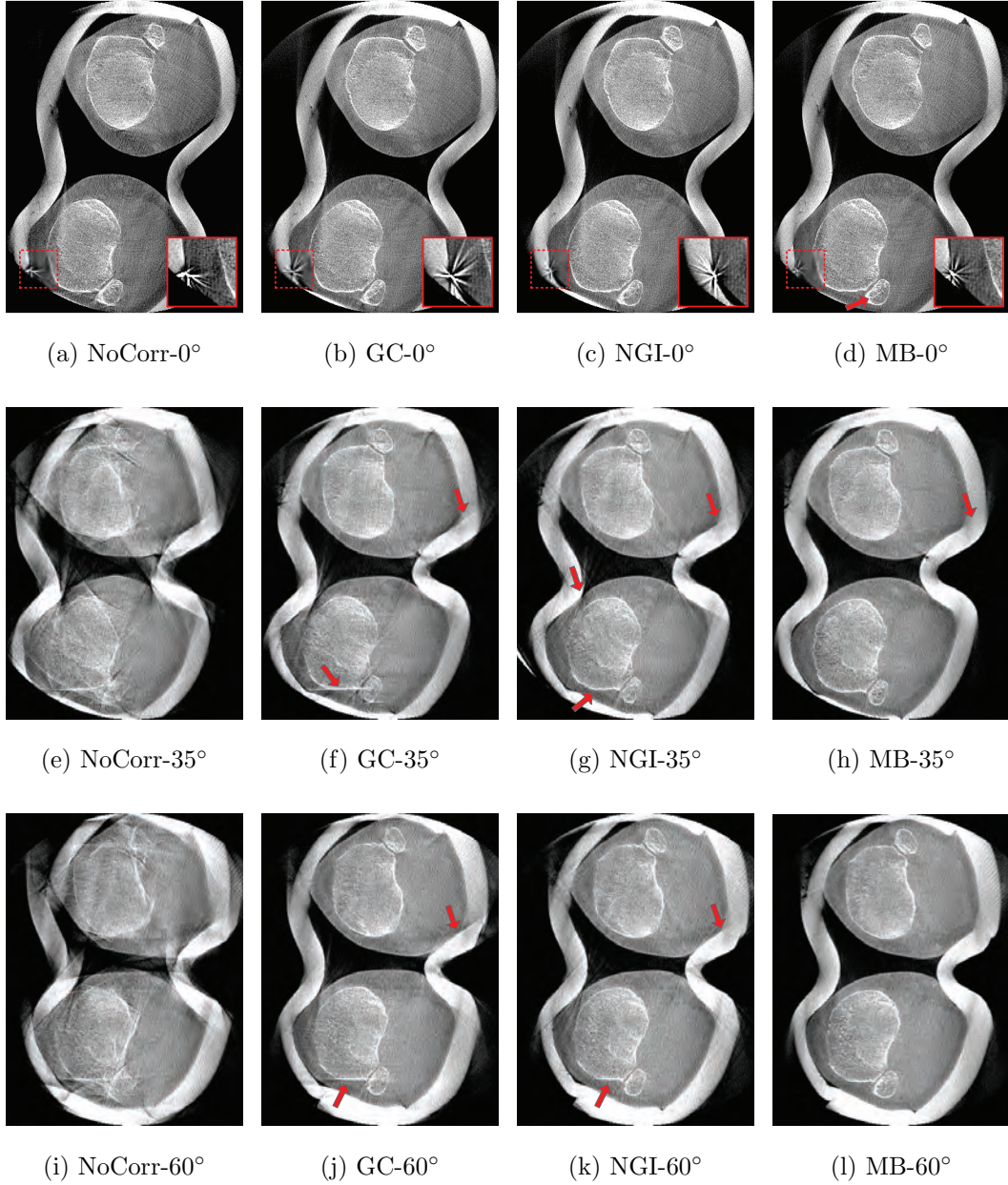


Figure 5. Axial slices through the tibia and fibula. From left to right: Reconstructions without motion correction (NoCorr), the proposed method using GC, the proposed method using NGI and the MB reference method. The rows correspond to the three different weight-bearing scans from 0° flexion angle at the top to 60° flexion angle at the bottom (W: 2025 HU, C: 145 HU).

352 and was able to restore the shape of the plasticine wrap. The slight streaking at the anterior
 353 skin boundary in NGI and GC originates from the plasticine wrap and is not related to the
 354 image quality of the bones. A clearer difference can be seen at the tibia. Again, all methods
 355 clearly improved image quality, yet, the GC could not fully correct the bones' outlines,

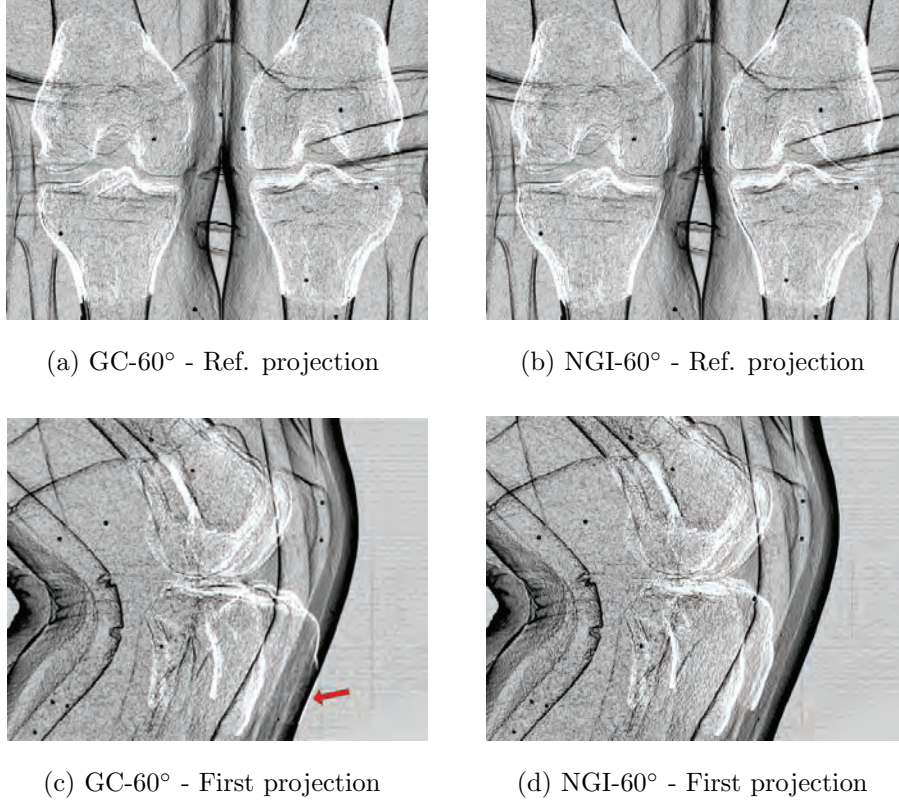


Figure 6. Difference of gradient magnitudes between DRR and acquired projections after registration. Top row: Projections used for extracting the reference coordinate system. Bottom row: Projections with large occlusions led to incorrect registration of the left tibia using the GC method.

356 especially in case of the left tibia (see Figure 5f). The MB and NGI images are of similar
 357 image quality with slightly more residual streaking in the NGI case (see Figure 5(g)-(h)).
 358 The reduced image quality of the GC case is due to misregistrations of the 2D/3D alignment
 359 as illustrated in Figure 6.

360 The highest amount of motion was observed in the scan with a 60° flexion angle, as can
 361 be seen in the uncorrected reconstructions in Figure 4i and Figure 5i. Both GC and NGI
 362 successfully estimated the patient’s femoral motion, yielding comparable reconstructions of
 363 the femur with substantially improved image quality compared to an uncorrected recon-
 364 struction. More streaking is present in the left tibia for GC, whereas NGI shows a good
 365 tibial reconstruction. Similar to the scan at 35° flexion angle, the skin boundaries are better
 366 corrected when using the MB approach with slight streaking in the 2D/3D registration-based
 367 approaches. Apart from that, the visual results are comparable.

UQI ($\times 10^2$)				
dataset	NoCorr	GC	NGI	MB
0°	34.9 \pm 2.3	62.6 \pm 3.6	63.5\pm4.3	57.2 \pm 6.4
35°	11.2 \pm 3.5	47.2 \pm 6.7	53.1\pm3.7	50.0 \pm 7.1
60°	9.0 \pm 4.1	49.9 \pm 5.7	51.7 \pm 5.0	52.9\pm7.4

Table I. Mean and standard deviation of the UQI over four bone regions. All correction methods lead to an increased UQI compared to an uncompensated reconstruction. The bold font emphasizes the method with the highest UQI for each dataset.

B. Image Quality Measures

The qualitative measures yielded an UQI value for each combination of bone, flexion angle and correction method. Table I shows the mean values over all four bone regions together with the standard deviation. Note that the UQI has been scaled by a factor of 100 throughout the paper for better visualization. Each weight-bearing scan showed moderate intensity variations due to different detector saturation and truncation artifacts. The UQI is known to be robust against intensity variations [38], which allows for a fair differentiation between the methods as well as the individual weight-bearing scans.

1. Inter-Scan Comparison

A reduction of the UQI value in case of non-corrected reconstructions from 34.9 to less than 9.0 is in line with the amount of motion observed visually. This is supported by the maximum achieved UQI values when applying correction methods. All methods showed the highest values for the 0° flexion with an average UQI of 61.1. The best achieved UQI values for flexion angles 35° and 60° have been substantially lower with a maximum UQI of 53.1. UQI values between 35° and 60° flexion did not show a substantial difference.

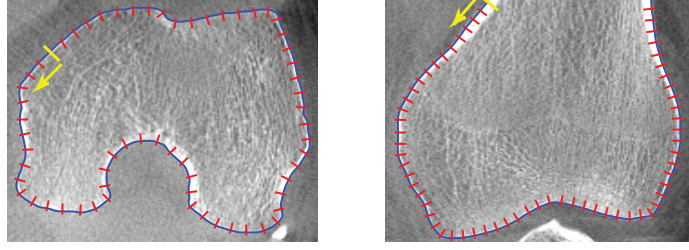
We notice a large improvement of the UQI values from no correction to any of the correction methods. The relative improvement for each dataset is well supported by the visual impression in Figure 4 and Figure 5. The GC showed the lowest improvement for all flexion angles and also high inter-bone variations which may originate from the misregistration-based streaking artifacts seen in Figure 5j and Figure 5f. Even though the MB method yields better reconstructions with less streaking, the NGI method shows higher UQI values for the 0° and 35° dataset. This discrepancy is analyzed in more detail in the following section.

C. Deformation of bone outline for MB method

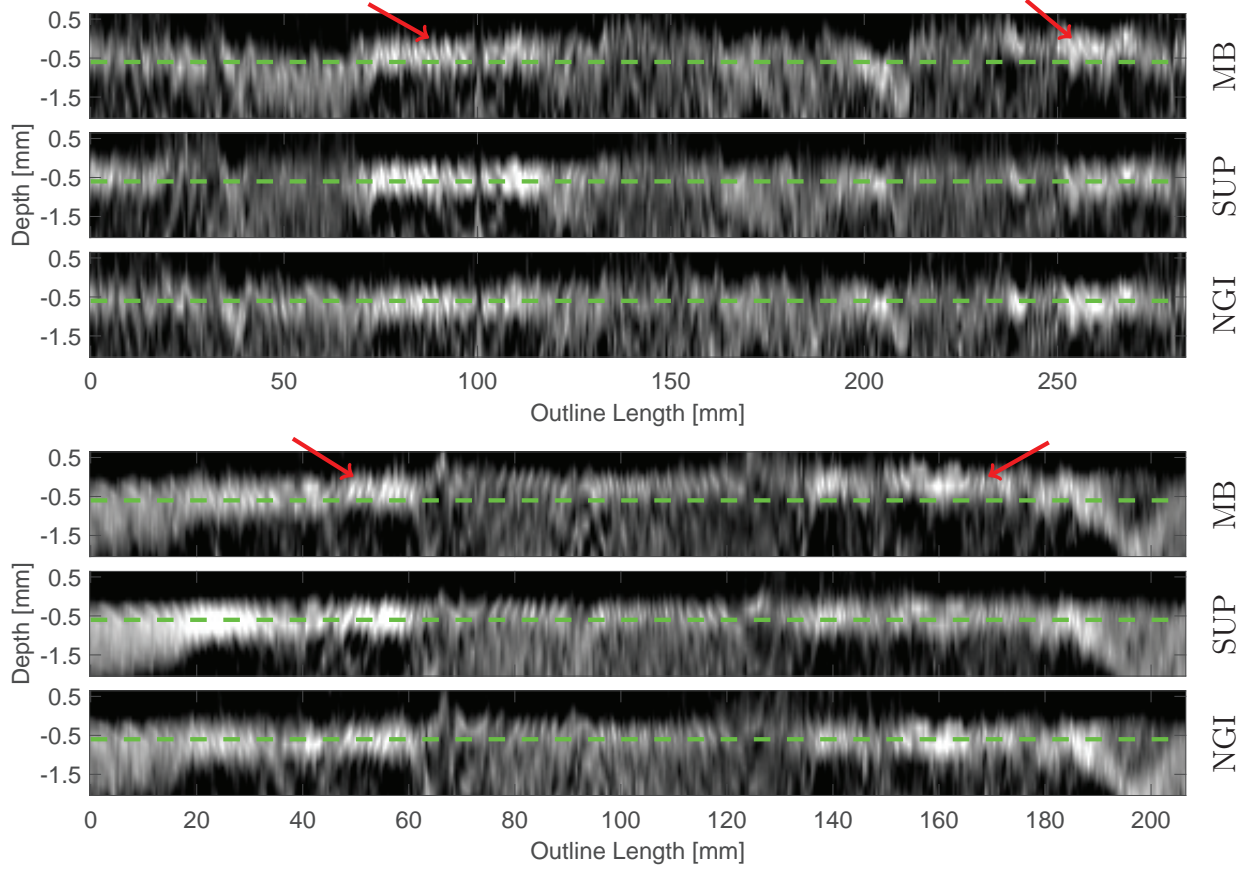
In contrast to the visual results, the highest UQI values were achieved by the NGI method with a mean distance to the MB correction of 6.3 in the 0° dataset, 3.1 in the 35° dataset and comparable results for the 60° dataset. This could be explained by a small deformation of the reconstructed bones in case of the MB approach. To analyze this deformation we extracted line profiles equidistantly and orthogonal to the femur’s outline. The semi-automatic segmentation generally yielded an outline that slightly extended outside the visually determined bone outline. Therefore, we manually refined the segmentation results for a selected axial and coronal slice of the supine acquisition, such that the bone’s outline is covered exactly. The refinement was performed using the manual segmentation functionality of ITK-Snap [36]. Then, 2D spline models of the refined segmentation mesh were obtained for the selected axial and coronal slice. We also extracted the corresponding image slices for NGI, MB and the supine data. Finally, we sampled line profiles perpendicular to the fitted spline, equidistantly along the whole spline curve.

Figure 7a shows the axial and coronal slice of the motion free reconstruction with an overlay of the extracted spline curves and a subset of the line profiles. In Figure 7b the extracted profiles are shown. We incorporated a dashed reference line at 0.5mm distance to the spline, which corresponds roughly to the center of the femur’s cortical bone edge in the supine scan.

The resulting profiles show that for the MB approach the edge intensities shift upwards



(a) Line profile measurement in an axial and coronal slice.



(b) Edge profiles along the right femur's outline in an axial (top) and coronal (bottom) slice.

Figure 7. Edge profiles along the outline of the right femur for NGI-0°, MB-0° and the supine, motion-free reference (SUP). The depth axis points from the bone outwards. The starting point and direction of the x-axis is indicated by arrows in 7a. Compared to the NGI method the edge shifts upwards for the MB method indicating a scaling effect.

at multiple locations with respect to the reference line. This is not the case for the NGI approach. Shifting upwards corresponds to a deformation perpendicular to the bone's surface. It can be interpreted as scaling or distortion if it occurs uniformly along the spline. In

TRE [mm]			
dataset	GC	NGI	MB
0°	0.45±0.29	0.37±0.24	0.15±0.08
35°	1.45±1.32	0.92±0.86	0.35±0.17
60°	1.55±2.46	0.75±1.07	0.29±0.16

Table II. Mean and standard deviations for the 2D TRE of reprojected marker locations. 3D marker detection was done in the motion corrected reconstructions for GC and GI. For MB the existing 3D estimates have been used.

addition to the distortion, we can observe that the edge of the NGI approach is more similar to the motion-free scan, than the edge of the MB method.

D. Target Registration Error

As expected, for all datasets the MB method yielded the smallest TRE values and also the smallest deviations. The maximum TRE of the GC method amounts to 1.55mm, whereas all TRE values were no larger than 0.92mm for the NGI method. Compared to the MB approach the standard deviations for NGI and GC increased substantially. Yet, for the NGI method the highest standard deviation (1.07mm) is still considerably smaller than for the GC case (2.46mm).

E. Relative Bone Motion

In Figure 8 we show boxplots based on all relative motion parameters over all bones and all projections. We limited the analysis to the NGI method using the 60° case, as this corresponded to the highest amount of motion estimated with the best performing registration approach. The boxes depict the 25th and the 75th percentiles and the whiskers cover $\approx 99\%$ of the samples in case of a normal distribution. The red line shows the median of the samples. Almost all rotation angles lie within a range of 1°, which corresponds to the maximum rotational deviation between the bones. The translations in vertical direction, i.e. Δt_z show only little difference with a maximum range below 1mm. In contrast, translations in x- and

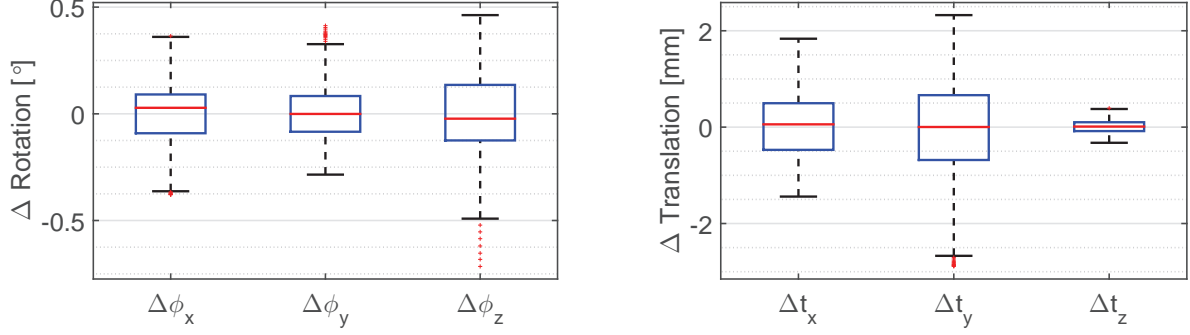


Figure 8. Relative bone motion during the 60° scan estimated by the NGI approach. The motion parameters show the deviation to the average rigid transform over all four bones.

y-axis are in the range of $[-1.44\text{mm}, 1.84\text{mm}]$ and $[-2.89\text{mm}, 2.32\text{mm}]$, respectively.

IV. DISCUSSION

CBCT scanning of knees under weight-bearing conditions poses a difficult motion estimation and compensation problem. In previous work on image-based motion correction, we tried to correct for involuntary patient motion without markers using 2D/2D registration of projection images and maximum intensity DRRs of a motion-corrupted reconstruction [19]. However, the improvement in image quality was limited due to the amount of motion artifacts present in the initial reconstruction. All of our weight-bearing studies include a non-weight bearing scan in supine position, which serves as a reference for the investigation of functional parameters, such as joint space analysis.

A non-weight-bearing scan in a supine patient position is part in all our weight-bearing studies to obtain a reference for further investigation, e.g., a joint space analysis. Therefore, we propose a marker-free motion correction method based on piecewise-rigid 2D/3D registration of a motion-free reference volume to all projection images acquired in a weight-bearing scan. The proposed method builds on previous work by Berger et al. [20], where a proof-of-concept for the GC approach is presented and evaluated on a noise-free numerical phantom. We developed the method further to allow its application to real patient acquisitions. Our main contributions are an improved similarity measure (NGI), a noise reduction approach for the forward projected gradient images and a more sophisticated estimation of a global motion field using non-rigid TPS extrapolation. Further, we extended

the MB motion estimation as shown in Müller et al. [4] by an analytic gradient computation and an improved outlier detection scheme. Moreover, this work gives a complete overview of existing motion-compensation methods for weight-bearing CBCT of the knee-joint and shows a thorough evaluation and comparison between the state-of-the-art MB and the novel, image-based approach using 2D/3D registration.

Our results show a substantial improvement of image quality for GC and NGI as well as for the MB approach compared to reconstructions without motion correction. The NGI method yielded higher UQI values and showed less streaking at the bone boundaries than the GC method. We believe that the NGI is indeed more robust to overlapping structures as indicated by Otake et al. [15]. This is of special importance in lateral views where we see increased overlap of left and right leg accompanied by an increased number of misregistrations in case of the GC method (see Figure 6).

The goal of the proposed method was to achieve an image quality similar to that of the MB approach. As expected the MB method was able to accurately restore the skin outline whereas the registration-based methods showed inaccuracies at the skin boundary that also led to slight streaking artifacts. In contrast, we see the smallest amount of streaking for the MB approach. The NGI method performed comparable to the MB method in reconstructing the bones' outline. This is of special interest as a potential application is a 3D analysis of the joint space between the femur and tibia as outlined in section II C.

Although the MB approach produces better visual results, we have noticed a lower UQI compared to the NGI method, especially for smaller flexion angles. Line profiles along a femur's boundary revealed a deformation of the reconstructed bone outline with respect to the reference scan which can be seen as an intensity shift in Figure 7. Note that the edge of the MB approach is generally higher than that of the reference which shows that the intensity shift is not due to a misalignment but due to a real distortion effect. A possible cause can be the assumption of a global, rigid transform per projection image which does not allow for non-rigid motion. If we incorrectly assume a rigid movement per projection, this can lead to a distortion in the reconstruction. In a future study we plan to perform a statistical analysis of the directions and lengths of the residual distances in 2D. This may provide insight on the amount of non-rigidity present in the marker motion.

We deem MB methods unsuitable for the problem of reconstructing the knee under weight-bearing conditions for a number of reasons. First and foremost, markers are cumber-

486 some and too time-intensive to attach by a physician in clinical routine. If not done carefully,
 487 markers may overlap in projections from some angles, leading to potential mismatching and
 488 lower estimation accuracy. Second, MB methods are restricted to a single rigid motion
 489 model. Rigid transformations of individual bones or even deformable models would require
 490 a considerably larger number of markers, exacerbating the problem of overlapping markers.
 491 On top of that, markers can only be attached to the skin, while the accurate relationship
 492 of bone and skin motion is unclear. Finally, markers degrade image quality due to metal
 493 artifacts. This work remedies the above mentioned problems by entirely image-based meth-
 494 ods, thus possibly allowing a fully automatic system in the future. We currently rely on
 495 a semi-automatic segmentation and initialization of the supine bones. Methods that allow
 496 for a fully automatic segmentation of the bones and an automatic initial alignment could
 497 eliminate manual interaction and will be part of future work.

498 We have used the UQI for a qualitative evaluation of the reconstruction results. Note that
 499 the pipeline described in section II C 2 requires a truly motion-free supine scan preferably of
 500 high quality. Based on the patient’s supine position and the good reconstruction quality of
 501 the supine scans (cf. Figure 3), we could not identify patient motion in the supine scan. The
 502 image quality was superior to the standing scans, as the supine acquisition protocol used
 503 twice as many projection images. Another limitation of the evaluation method is that it also
 504 includes potential registration errors of the 3D/3D registration. To eliminate this subsequent
 505 registration we would require a ground truth motion of the bones, which is hardly possible
 506 during an in vivo study. Finally, the reported UQI values describe the mean values over the
 507 bone ROIs in Figure 3. The ROIs include the bones and adjacent soft tissue, yet we do not
 508 claim that the UQI is an accurate measure for soft tissue deformation.

509 To evaluate the TPS extrapolation we computed the reprojection error of the markers
 510 (TRE) for the GC, NGI and MB methods. As expected the MB method performed best
 511 with a maximum TRE of 0.35mm for the 35° case. The NGI method yielded a maximum
 512 TRE of 0.92mm for the same dataset, which is still acceptable. Note, that the markers
 513 represent the boundary of the anatomically meaningful area and therefore have the largest
 514 distances to the bone surfaces on which the extrapolation is based. As is generally the case
 515 in extrapolation methods, the confidence of the TPS extrapolation decreases with increasing
 516 distance to the control points. Additionally, the MB cost-function aims to minimize exactly
 517 this error, whereas GC and NGI do not consider the markers in their optimization.

518 Even though the proposed NGI method creates an accurate reconstruction of the bone
 519 outline, we still need to consider the effect of 2D/3D registration errors. As all our bone
 520 positions are estimated independently, errors in 2D/3D registrations may lead to deviations
 521 in relative bone position, e.g., between tibia and femur. Our analysis of relative bone motions
 522 in section III E shows rather small rotational differences between the bones and only little
 523 variation in z-axis translations. Note that the estimated relative motion will always consist
 524 of a combination of the real ground truth motion and the residual error of the registration.
 525 Hence, for an exact measurement of relative bone motions we would need access to the
 526 ground truth motion directly, which was not possible during our in vivo, weight-bearing
 527 acquisitions. Thus, it remains unclear if the increased deviations of translations in x- and
 528 y-axis correspond to registration errors or real variability of motion. A study using cadaver
 529 legs, where tibia and femur are fixed to a device that applies a predefined motion pattern
 530 would allow an exact measurement of registration errors and is planned as future work.

531 As explained in section II B 4, we reduced the bone segmentations to the bone outlines
 532 to limit the noise level in the DRRs. This makes the method similar to a mesh-to-image
 533 2D/3D registration where the meshes are directly registered to the acquired projection im-
 534 ages without the need of an expensive DRR generation step. An extensive overview of these
 535 methods is given in [39], where it is referred to as “feature-based” 2D/3D registration. A
 536 comparison of the two methods would be interesting for future work, yet, we expect that the
 537 current method is more robust to variations in the segmentation quality as it does not only
 538 consider segmented 3D positions but also the measured intensities. Moreover, most mesh-
 539 based registration methods require feature point detection in the projection images which
 540 will be difficult considering the high degree of overlapping structures in our acquisitions.

541 We will also investigate if we can perform the motion correction using a cost-function
 542 based on data consistency conditions (DCC) instead of 2D/3D registration. Methods based
 543 on image moments [40], filtering in the Fourier domain [41] or by using the Epipolar geometry
 544 [42, 43] will be investigated. The motion is then estimated directly in the projection domain
 545 without the need for a motion-free scan and bone segmentations.

V. CONCLUSION

We presented a novel motion correction scheme to allow for weight-bearing CBCT imaging of the knee joint. The involuntary patient motion is estimated with respect to a motion-free reference scan in supine patient position. Left and right femur and tibia are segmented and registered to the acquired weight-bearing projections. Thus, six rigid motion parameters were estimated for each bone and each projection resulting in a total of 5952 parameters. To improve registration results we also incorporated a regularizer that ensured smoothness of the motion-parameters over time. The motion was then used to estimate a TPS-based non-rigid deformation field for each projection which was directly incorporated into the backprojection step, yielding a motion-compensated reconstruction.

Our study included a thorough comparison between two versions of our proposed method and a state-of-the-art MB motion estimation method [4]. All correction methods substantially improved image quality compared to reconstructions without motion correction. The GC similarity measure proved to be less robust to overlapping bone structures than the NGI similarity metric. Our quantitative evaluation over ROIs around the bones showed a mean UQI of 18.4 for no correction, 53.3 and 56.1 for the proposed method using GC and NGI, respectively, and 53.7 for the MB reference approach. Increased streaking was observed for GC, whereas the visual image quality for NGI was close to that of the MB approach. In contrast to the MB method, the proposed method does not require the attachment of markers which will improve the clinical workflow and patient comfort. Further, we found that the MB method causes small, non-rigid deformations at the bone outline which indicates that markers may not accurately reflect the internal motion at tibia and femur. Therefore, we believe that the proposed method is a promising alternative to MB motion management.

For future work we plan further improvements of the 2D/3D registration algorithms, e.g. by incorporating an analytic gradient computation. Furthermore, we plan a thorough evaluation of the impact of residual registration errors on the relative positioning of bones.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding of the Research Training Group 1773 Heterogeneous Image Systems and of the Erlangen Graduate School in Advanced Optical Technolo-

gies (SAOT) by the German Research Foundation (DFG) in the framework of the German
excellence initiative. Further, we acknowledge funding support from NIH Shared Instrument
Grant S10 RR026714 supporting the zeego@StanfordLab, and Siemens AX.

-
- [1] J.-H. Choi, A. Maier, A. Keil, S. Pal, E. J. McWalter, G. S. Beaupré, G. E. Gold, and
R. Fahrig, “Fiducial marker-based correction for involuntary motion in weight-bearing C-arm
CT scanning of knees. II. Experiment,” *Medical Physics* **41**, 061902 (2014).
- [2] J.-H. Choi, R. Fahrig, A. Keil, T. F. Besier, S. Pal, E. J. McWalter, G. S. Beaupre, 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.” *Medical Physics* **40**,
091905 (2013).
- [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,” in *SPIE Medical
Imaging*, edited by N. J. Pelc, E. Samei, and R. M. Nishikawa (International Society for
Optics and Photonics, 2011) pp. 796123–796123–8.
- [4] K. Müller, M. Berger, J.-H. Choi, A. Maier, and R. Fahrig, “Automatic Motion Estimation
and Compensation Framework for Weight-bearing C-arm CT scans using Fiducial Markers,”
in *IFMBE Proceedings*, edited by D. A. Jaffray (2015) pp. 58–61.
- [5] M. Berger, C. Forman, C. Schwemmer, J. H. Choi, K. Müller, A. Maier, J. Hornegger, and
R. Fahrig, “Automatic Removal of Externally Attached Fiducial Markers in Cone Beam C-
arm CT,” in *Bildverarbeitung für die Medizin 2014*, edited by H. H. T. Deserno (2014) pp.
168–173.
- [6] C. Schwemmer, C. Rohkohl, G. Lauritsch, K. Müller, and J. Hornegger, “Residual motion
compensation in ECG-gated interventional cardiac vasculature reconstruction,” *Physics in
Medicine and Biology* **58**, 3717 (2013).
- [7] E. Hansis, D. Schäfer, O. Dössel, and M. Grass, “Projection-based motion compensation for
gated coronary artery reconstruction from rotational x-ray angiograms.” *Physics in Medicine
and Biology* **53**, 3807–20 (2008).
- [8] T.-Y. Tsai, T.-W. Lu, C.-M. Chen, M.-Y. Kuo, and H.-C. Hsu, “A volumetric model-based
2D to 3D registration method for measuring kinematics of natural knees with single-plane

- 604 fluoroscopy,” *Medical Physics* **37**, 1273 (2010).
- 605 [9] C.-C. Lin, S. Zhang, J. Frahm, T.-W. Lu, C.-Y. Hsu, and T.-F. Shih, “A slice-to-volume
606 registration method based on real-time magnetic resonance imaging for measuring three-
607 dimensional kinematics of the knee.” *Medical Physics* **40**, 102302 (2013).
- 608 [10] Z. Zhu and G. Li, “An automatic 2D-3D image matching method for reproducing spatial knee
609 joint positions using single or dual fluoroscopic images.” *Computer methods in biomechanics
610 and biomedical engineering* **15**, 1245–56 (2012).
- 611 [11] Z. Zhu, D. F. Massimini, G. Wang, J. J. P. Warner, and G. Li, “The accuracy and repeata-
612 bility of an automatic 2D-3D fluoroscopic image-model registration technique for determining
613 shoulder joint kinematics.” *Medical Engineering & Physics* **34**, 1303–9 (2012).
- 614 [12] J. Wang, A. Borsdorf, B. Heigl, T. Köhler, and J. Hornegger, “Gradient-Based Differential
615 Approach for 3-D Motion Compensation in Interventional 2-D/3-D Image Fusion,” in *Inter-
616 national Conference on 3D Vision*, edited by I. C. P. Services (Japan, 2014) pp. 293–300.
- 617 [13] P. Bifulco, M. Sansone, M. Cesarelli, R. Allen, and M. Bracale, “Estimation of out-of-plane
618 vertebra rotations on radiographic projections using CT data: a simulation study,” *Medical
619 Engineering & Physics* **24**, 295–300 (2002).
- 620 [14] P. Bifulco, M. Cesarelli, R. Allen, M. Romano, A. Fratini, and G. Pasquariello, “2D-3D Regis-
621 tration of CT Vertebra Volume to Fluoroscopy Projection: A Calibration Model Assessment,”
622 *EURASIP Journal on Advances in Signal Processing* **2010**, 1–8 (2010).
- 623 [15] Y. Otake, A. S. Wang, J. Webster Stayman, A. Uneri, G. Kleinszig, S. Vogt, A. J. Khanna,
624 Z. L. Gokaslan, and J. H. Siewerdsen, “Robust 3D-2D image registration: application to spine
625 interventions and vertebral labeling in the presence of anatomical deformation.” *Physics in
626 Medicine and Biology* **58**, 8535–53 (2013).
- 627 [16] Y. Otake, A. S. Wang, A. Uneri, G. Kleinszig, S. Vogt, N. Aygun, S.-f. L. Lo, J.-P. Wolinsky,
628 Z. L. Gokaslan, and J. H. Siewerdsen, “3D2D registration in mobile radiographs: algorithm
629 development and preliminary clinical evaluation.” *Physics in Medicine and Biology* **60**, 2075–
630 90 (2015).
- 631 [17] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, and D. J. Hawkes, “A comparison
632 of similarity measures for use in 2-D-3-D medical image registration.” *IEEE Transactions on
633 Medical Imaging* **17**, 586–95 (1998).
- 634 [18] S. Ouadah, J. W. Stayman, G. Gang, A. Uneri, T. Ehtiati, and J. H. Siewerdsen, “Self-

calibration of cone-beam CT geometry using 3D-2D image registration: development and application to task-based imaging with a robotic C-arm,” in *SPIE Medical Imaging*, edited by Z. R. Yaniv and R. J. Webster (International Society for Optics and Photonics, 2015) p. 94151D.

[19] M. Unberath, J.-H. Choi, M. Berger, A. Maier, and R. Fahrig, “Image-based Compensation for Involuntary Motion in Weight-bearing C-arm Cone-beam CT Scanning of Knees,” in *Proc. SPIE Medical Imaging 2015*, Vol. 9413, edited by S. C. H. D. Manning (2015).

[20] M. Berger, K. Müller, J.-H. Choi, A. Aichert, A. Maier, and R. Fahrig, “2D/3D Registration for Motion Compensated Reconstruction in Cone-Beam CT of Knees Under Weight-Bearing Condition,” in *IFMBE Proceedings*, edited by D. A. Jaffray (2015) pp. 54–57.

[21] G. Loy and A. Zelinsky, “Fast radial symmetry for detecting points of interest,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**, 959–973 (2003).

[22] K. Müller, C. Schwemmer, J. Hornegger, Y. Zheng, Y. Wang, G. Lauritsch, C. Rohkohl, A. Maier, C. Schultz, and R. Fahrig, “Evaluation of interpolation methods for surface-based motion compensated tomographic reconstruction for cardiac angiographic C-arm data,” *Medical Physics* **40** (2013), 10.1118/1.4789593.

[23] H. Livyatan, Z. Yaniv, and L. Joskowicz, “Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT,” *IEEE Transactions on Medical Imaging* **22**, 1395–406 (2003).

[24] W. Wein, B. Roeper, and N. Navab, “2D/3D Registration Based on Volume Gradients,” in *Medical Imaging*, edited by J. M. Fitzpatrick and J. M. Reinhardt (International Society for Optics and Photonics, 2005) pp. 144–150.

[25] <https://www5.cs.fau.de/research/software/java-parallel-optimization-package/>.

[26] B. Lorch, M. Berger, J. Hornegger, and A. Maier, “Projection and Reconstruction-Based Noise Filtering Methods in Cone Beam CT,” in *Bildverarbeitung für die Medizin 2015*, edited by H. Handels (2015) pp. 59–64.

[27] M. Davis, A. Khotanzad, D. Flamig, and S. Harms, “A physics-based coordinate transformation for 3-d image matching,” *IEEE Transactions on Medical Imaging* **16**, 317–328 (1997).

[28] E.-P. Rührnschopf and K. Klingenberg, “A general framework and review of scatter correction methods in cone beam CT. part 2: Scatter estimation approaches,” *Medical Physics* **38** (2011).

[29] D. L. Parker, “Optimal short scan convolution reconstruction for fan beam CT,” *Medical Physics* **9**, 254–257 (1982).

- [30] B. Ohnesorge, T. Flohr, K. Schwarz, J. P. Heiken, and K. T. Bae, “Efficient correction for CT image artifacts caused by objects extending outside the scan field of view,” *Medical Physics* **27**, 39–46 (2000).
- [31] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging* (Society for Industrial and Applied Mathematics, 2001).
- [32] H. Scherl, B. Keck, M. Kowarschik, and J. Hornegger, “Fast GPU-Based CT Reconstruction using the Common Unified Device Architecture (CUDA),” in *2007 IEEE Nuclear Science Symposium Conference Record*, Vol. 6 (IEEE, 2007) pp. 4464–4466.
- [33] D. Schaefer, J. Borgert, V. Rasche, and M. Grass, “Motion-compensated and gated cone beam filtered back-projection for 3-D rotational X-ray angiography,” *IEEE Transactions on Medical Imaging* **25**, 898–906 (2006).
- [34] Q. Cao, G. Thawait, G. J. Gang, W. Zbijewski, T. Reigel, T. Brown, B. Corner, S. Demehri, and J. H. Siewerdsen, “Characterization of 3D joint space morphology using an electrostatic model (with application to osteoarthritis).” *Physics in Medicine and Biology* **60**, 947–60 (2015).
- [35] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig, “CONRAD—a software framework for cone-beam imaging in radiology.” *Medical Physics* **40**, 111914 (2013).
- [36] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, “User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability.” *NeuroImage* **31**, 1116–28 (2006).
- [37] S. Pieper, M. Halle, and R. Kikinis, “3D Slicer,” in *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No. 04EX821)*, Vol. 2 (IEEE, 2004) pp. 632–635.
- [38] Z. Wang and A. Bovik, “A universal image quality index,” *Signal Processing Letters, IEEE* **9**, 81–84 (2002).
- [39] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, “A review of 3D/2D registration methods for image-guided interventions.” *Medical image analysis* **16**, 642–61 (2012).
- [40] R. Clackdoyle and L. Desbat, “Data consistency conditions for truncated fanbeam and parallel projections,” *Medical Physics* **42**, 831–845 (2015).
- [41] M. Berger, A. Maier, Y. Xia, J. Hornegger, and R. Fahrig, “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*, edited by F. Noo (2014) pp. 329–332.

[42] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. Maier, “Epipolar Consistency in Transmission Imaging.” *IEEE Transactions on Medical Imaging* (2015).

[43] A. Aichert, N. Maass, Y. Deuerling-Zheng, M. Berger, M. Manhart, J. Hornegger, A. K. Maier, and A. Doerfler, “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*, edited by F. Noo (2014) pp. 333–337.