JOINT CALIBRATION AND MOTION ESTIMATION IN WEIGHT-BEARING CONE-BEAM CT OF THE KNEE JOINT USING FIDUCIAL MARKERS

C. Syben[1], B. Bier[1], M. Berger[1], A. Aichert[1], R. Fahrig[2], G. Gold[2], M. Levenston[2] and A. Maier[1]

[1]Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg
[2]Radiological Sciences Lab, Stanford University

ABSTRACT

Recently, C-arm cone-beam CT systems have been used to acquire knee joints under weight-bearing conditions. For this purpose, the C-arm acquires images on a horizontal trajectory around the standing patient, who shows involuntary motion. The current state-of-the-art reconstruction approach estimates motion based on fiducial markers attached to the knee. A drawback is that this method requires calibration prior to each scan, since the horizontal trajectory is not reproducible. In this work, we propose a novel method, which does not need a calibration scan. For comparison, we extended the state-of-the-art method with an iterative scheme and we further introduce a closed-form solution of the compensated projection matrices. For evaluation, a numerical phantom and clinical data are used. The novel approach and the extended state-of-the-art method achieve a reduction of the reprojection error of 94% for the phantom data. The improvement for the clinical data ranged between 10% and 80%, which is followed by the visual impression. Therefore, the novel approach and the extended state-of-the-art method achieve superior results compared to the state-of-the-art method.

Index Terms— C-arm CBCT, Image Reconstruction, Motion Estimation, Self-Calibration, Knee-Joint Imaging

1. INTRODUCTION

Recent work on 3D acquisition of the knee joint under weight-bearing conditions use C-arms on a horizontal trajectory around the standing or squatting patient [1, 2, 3]. C-arms are not originally designed to scan on horizontal trajectories, which makes the scans less repeatable. In addition, the standing patient position favors involuntary motion. To improve the 3D image quality, both scanner and patient motion have to be estimated and compensated.

The motion can be estimated using purely image-based or marker-based approaches. Image-based methods estimate the motion using e.g. data consistency conditions [4, 5, 6]. Marker-based approaches have been used often in medical imaging [2, 7, 8, 9], but are not implemented on real clinical C-arm CT systems [3]. In order to estimate the scanner motion a self-calibration approach would be beneficial. For a self-calibration of the C-arm CT system either image-based approaches or external tracking techniques can be used. Image-based approaches use 2D-3D registration or optimize image-based metrics to calibrate the system [10, 11]. Also, external-tracking techniques were published [12, 13], but rely on external cameras.

The state-of-the-art approach is a fully automatic marker-based motion estimation and compensation framework [3, 7, 14]. The current approach shows a clear improvement of image quality, yet, we identified two main limitations. The method estimates 3D marker positions from the detected, motion-corrupted, 2D positions, which are then used to estimate 2D-3D point correspondences. These correspondences are assumed to remain constant during motion estimation, even though their assignment is directly dependent on the currently estimated motion. Additionally, the system has to be calibrated with a calibration phantom at the beginning of a scan series [15]. In order to improve the accuracies of the 3D marker positions and the assignment of point correspondences and to overcome the calibration effort, we propose a novel optimization model with an iterative scheme to update the 3D positions and the point correspondences during optimization. Furthermore, we propose a self-calibration which makes a prior calibration scan obsolete. For comparability we extend the state-of-the-art method with the iterative scheme and also compare the approach to a closed form solution.

2. MATERIALS AND METHODS

2.1. State-of-the-Art Method

2.1.1. Motion Estimation and Compensation

The state-of-the-art method published by Müller et al. [3] combines the motion estimation presented by Choi et al. [2] with the automatic marker detection introduced by Berger et al. [14] and is summarized in the following. First, a reference 3D marker position is estimated for each marker by backprojecting the detected 2D marker positions. The detected 2D marker positions are then assigned to the 3D locations, such that the reprojection error (RPE) between projected 3D points and detected 2D points is minimal. The rigid motion is then estimated by further minimization of the RPE given by

$$\arg\min_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}) = \arg\min_{\boldsymbol{\alpha}} \frac{1}{2} \sum_{j=1}^{J} \sum_{i=1}^{M} ||h(\boldsymbol{n}) - \boldsymbol{u}_{ij}||_2^2 \ , \quad (1)$$

where

$$\boldsymbol{n} = (n_1 \quad n_2 \quad n_3)^\top = \mathbf{P}_j \cdot \mathbf{M}_j(\boldsymbol{\alpha}) \cdot (\boldsymbol{v}_i \quad 1)^\top \ , \quad (2)$$

with vector $\boldsymbol{\alpha} \in \mathbb{R}^{6J}$ containing three rotation and translation parameters for each projection. The matrix $\mathbf{M}_j(\alpha) \in \mathbb{R}^{4\times4}$ applies the rigid motion to the calibrated projection matrix $\mathbf{P}_j \in \mathbb{R}^{3\times4}$ for the $j-$th of J projections. The estimated i-th of M 3D marker position is given by $\boldsymbol{v}_i$ and the corresponding detected 2D position on the j-th projection is given by $\boldsymbol{u}_{ij}$. The function $h : \mathbb{R}^3 \mapsto \mathbb{R}^2$ describes the mapping from 3D homogeneous coordinates to 2D coordinates $h(\boldsymbol{n}) = \left(\frac{n_1}{n_3} \quad \frac{n_2}{n_3}\right)^\top$. The new projection matrices $\mathbf{P}_j^{\text{new}} = \mathbf{P}_j \cdot \mathbf{M}_{\boldsymbol{\alpha}}$ can be directly used to obtain a motion compensated reconstruction (illustrated in Fig. 1a). An analytical derivative of the cost function is available [7]. In the following this method is referred to as the reference method.

2.1.2. Proposed Extension: Update the 3D marker positions

The initial estimated 3D marker positions and the point correspondences are negatively affected by patient motion. We propose an extension of the reference method (illustrated in Fig. 1b), where the estimated motion compensated projection matrices $\mathbf{P}_j^{\text{new}}$ are used for a second estimation step. With these new projection matrices more precise 3D marker positions and thus, more accurate point correspondences can be obtained, which improves the overall estimation.

2.2. Joint Motion Estimation and System Calibration

We present a method, which contains a self-calibration component and thus is independent of a calibration scan. Two problems occur if no calibrated projection matrices are available: first, the marker detection needs initial, viable projection matrices to estimate the 3D marker positions. Second, an evaluation of the objective function given by Equation 1 also require an initial set of projection matrices.

2.2.1. Initialization and Intrinsic Camera Model

The projection matrices are initialized with an ideal horizontal circular trajectory, which is created based on known parameters of the system's geometry. We further introduce a camera model, to be able to estimate deviations of the camera calibration. The projection matrix can be written as: $\mathbf{P} = \mathbf{K}\left[\mathbf{R}|\mathbf{t}\right]$, where $\mathbf{K} \in \mathbb{R}^{3\times3}$ is the intrinsic matrix and $\mathbf{R} \in \mathbb{R}^{4\times3}$ and $\mathbf{t} \in \mathbb{R}^{4\times1}$ contain six extrinsic parameters. From the perspective of the projection matrices it is not distinguishable whether the patient or the CT system is moving. Therefore, relative scanner and patient motion are both represented by the extrinsic parameters. Assuming isotropic detector pixels reduces the number of degrees of freedom (DoF) of $\mathbf{K}$ from 5 down to 3, i.e., the focal length f and the position of the central point c_x and c_y. For the estimation of these parameters, we use an extended model by Wein et al. [11], which is well suited for a cone-beam CT. The intrinsic calibration matrix $\mathbf{K}$ is then given by

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \ , \tag{3}$$

which can be written as

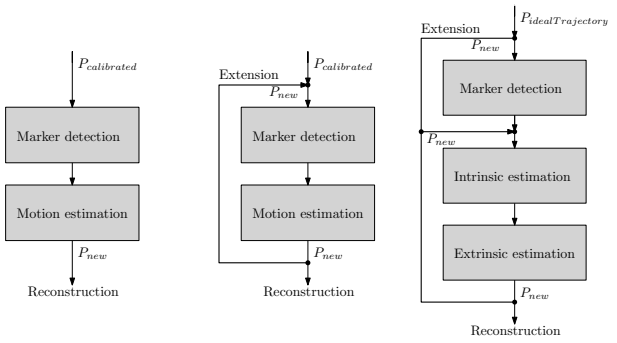
$$\begin{aligned} f_x &= \frac{p_x}{s_x} \cdot d, \qquad f_y = \frac{p_y}{s_y} \cdot d \\ c_x &= \frac{p_x}{2} + \frac{p_x}{s_x} \cdot \tan\eta \cdot d, \qquad c_y = \frac{p_y}{2} + \frac{p_y}{s_y} \cdot \tan\theta \cdot d \ , \end{aligned} \tag{4}$$

where d is the source detector distance (SDD) in mm. The width and height of the detector is given by s_x and s_y in mm and p_x and p_y are the width and height of the projection image in pixels. Further, η is the angle, at which the detector is tilted around the vertical detector axis $\boldsymbol{v}$. Similarly, angle θ describes the rotation around the horizontal detector axis $\boldsymbol{u}$.

2.2.2. Cost-Function for Motion Estimation and System Calibration

To incorporate the self-calibration into the current objective function, the projection matrix of Equation 1 is replaced with the decomposition of $\mathbf{P}_j$. The cost function for the joint estimation is given

$$\underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min}\ f(\boldsymbol{\alpha},\boldsymbol{\beta}) = \underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min}\ \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{M} ||h(\boldsymbol{n}) - \boldsymbol{u}_{ij}||_2^2 \ , \tag{5}$$



(a) Reference method. (b) Extended reference method. (c) Proposed method with extension.

Fig. 1: Flowchart for the three introduced methods.

where

$$\begin{aligned} \boldsymbol{n} &= \mathbf{P}_j(\boldsymbol{\alpha},\boldsymbol{\beta}) \cdot (\boldsymbol{v}_i \quad 1) \\ &= \mathbf{K}_j(\boldsymbol{\beta}) \cdot \left[\mathbf{R}_j \,|\boldsymbol{t}_j\right] \cdot \mathbf{M}_j(\boldsymbol{\alpha}) \cdot (\boldsymbol{v}_i \quad 1)^\top \ , \end{aligned} \tag{6}$$

where $\boldsymbol{\beta} \in \mathbb{R}^{3J}$ encodes three intrinsic parameters for each projection. Hence, $\boldsymbol{K}_{\boldsymbol{\beta}}$ describes the intrinsic matrix for the j-th projection depending on $\boldsymbol{\beta}$. The patient motion and the mechanical deviations from the system are applied by multiplying with $\boldsymbol{M}_{\boldsymbol{\alpha}}$, which describes the deviation from the circular trajectory. $\left[\mathbf{R}_j \,|\boldsymbol{t}_j\right]$ describes the circular trajectory and remains fixed during optimization.

The function is then solved using a gradient-based optimizer. An analytical derivative of the objective function is available [7]. The optimization is carried out in two steps. First, the cost function is minimized w.r.t. $\boldsymbol{\beta}$ and afterwards $\mathbf{K}_\beta$ is fixed and the function is optimized w.r.t. $\boldsymbol{\alpha}$. The iterative extension mentioned for the reference method is also applied for the proposed approach (illustrated in Fig. 1c), since inaccuracies of the 3D marker position have to be assumed due to the initialization and the patient motion. A complete overview of the introduced methods is illustrated in Fig. 1.

2.3. Closed-Form Solution

We compare our method also with a direct determination of the projection matrices by solving a system of linear equations. For the estimation of the 3D marker positions the projection matrices are initialized with the ideal horizontal circular trajectory. Given the 3D reference marker positions $\boldsymbol{v}_i$ and the corresponding detected 2D positions in the j-th projection $\boldsymbol{u}_{ij}$ we can set up a system of linear equations, which can be solved by

$$\begin{aligned} &\underset{\boldsymbol{p}}{\arg\min}\, ||\mathbf{A}\boldsymbol{p}||^2, \quad \text{subject to} \quad ||\boldsymbol{p}_j||^2 = 1 \\ &\underset{\boldsymbol{p}}{\arg\min}\, \boldsymbol{p}^\top\mathbf{A}^\top\mathbf{A}\boldsymbol{p} - \lambda(\boldsymbol{p}^\top\boldsymbol{p} - 1) \end{aligned} \tag{7}$$

where $\mathbf{A}_j \in \mathbb{R}^{2M\times12}$ is the measurement matrix, which contains the information of the 3D reference marker positions $\boldsymbol{v}_i$ and the corresponding detected 2D positions $\boldsymbol{u}_{ij}$. Further, $\boldsymbol{p}_j$ containing all 12 entries of the j-th projection matrix $\mathbf{P}_j$. Computing the zero crossings of the gradient and simplifying the result yields an eigenvalue problem

$$\mathbf{A}^\top\mathbf{A}\boldsymbol{p} = \lambda\boldsymbol{p} \ . \tag{8}$$

The components of the projection matrix $\mathbf{P}$ are then gained from the eigenvector belonging to the smallest eigenvalue. The iterative extension for the reference method is also applied.

2.4. Data Generation

For the experiment, 248 projections are acquired on $200°$. Each projection image has 1240×960 pixels with 0.308 mm isotropic pixel spacing.

We used a numerical phantom that consists of three cylinders with the materials bone marrow, femur and water and 36, 40, and 100 mm radius, respectively. The markers are steel beads placed along a helix on the surface of the outer cylinder. A thin wire with material bone and 0.2 mm diameter is placed in the isocenter, pointing in direction $(1\ 1\ 1)^\top$ [10]. The projections are generated using the CONRAD framework [16]. Calibrated projection matrices from a clinical scanner and previously estimated patient motion from clinical scans are used for the simulation.

Further, three clinical scans are acquired with a Siemens Artis zeego system (Siemens Healthcare GmbH, Forchheim, Germany) using a 10 s protocol.

3. RESULTS

3.1. Qualitative Results

In Fig. 2, reconstruction results of the motion corrupted, closed-form solution, reference, extended reference and of the proposed approach are shown. The first row in Fig. 2 shows part of the reconstructions of the phantom dataset with the wire in the ROI. The wire's shape is not accurately reconstructed using the closed form approach and the non-corrected reconstruction. The reference method is able to reconstruct the elliptical cross-section of the wire but the reconstruction is affected by streaking artifacts. The extended reference and the proposed method are able to reconstruct the shape of the wire more clear than the reference method.

In the three bottom rows in Fig. 2, the results of the clinical datasets are shown. Without correction, severe motion artifacts are present. The results of the closed form method show sustained reconstructed bone outlines but the streaking artifacts remain. The extended reference and the proposed method are able to reconstruct the bone outlines for all clinical datasets, although slight streaks remain. For the clinical dataset 2, the streaking artifacts for the reference method are stronger compared to the other two methods. Note that column 3 and 4 are rotated due to the different initialized trajectories.

3.2. Quantitative Results

The RPE's are listed in Tab. 1. Without correction the RPE is in general the highest compared to the other methods. The RPE for the closed form method is lower than the RPE of the reference method, except for the first clinical dataset, but is in each case higher compared to the other two methods. A clear improvement could be achieved using the extended reference method. The proposed method shows superior results than the reference method. However, the introduced iterative extension of the reference method yields comparable results than the proposed method.

The point spread function (PSF) is evaluated with the full-width at half-maximum (FWHM) [10]. In case of the phantom, the cross-section of the wire serve as the PSF and in the clinical cases the attached markers serve as the PSF. The FWHM values are listed in Tab. 2. Without correction, the FWHM is not measurable. The FWHM of the closed form approach is worse than the other two proposed methods, except for Clinical 3, where all methods yield comparable results. The FWHM of the reference method is slightly higher compared to the extended reference and the proposed method.

Table 1: RPE's in pixel for the different methods and datasets.

	Phantom	Clinical 1	Clinical 2	Clinical 3
No Correction	84.85	96.70	71.29	38.20
Closed Form	0.135	9.174	0.396	0.591
Reference	1.367	4.597	0.726	0.617
Ext. reference	0.088	2.099	0.143	0.561
Proposed	0.088	3.283	0.324	0.535

Table 2: FWHM (median $\pm$ std) for the different methods and datasets.

	Phantom	Clinical 1	Clinical 2	Clinical 3
Closed Form	0.40 ± 444	0.64 ± 3.5	0.82 ± 1.6	0.80 ± 1.7
Reference	0.36 ± 1.7	0.63 ± 2.7	0.84 ± 3.7	0.81 ± 3.3
Ext. reference	0.35 ± 2.6	0.63 ± 2.7	0.82 ± 1.3	0.82 ± 2.4
Proposed	0.35 ± 2.6	0.62 ± 2.9	0.81 ± 2.6	0.81 ± 1.7

4. DISCUSSION

We presented a novel motion estimation and system calibration method, which refines an ideal circular trajectory and does not require a calibration scan. For comparability we extend the reference method with an iterative scheme to update the 3D positions and further present a closed form solution for the projection matrices. The proposed and the extended reference method show comparable results, which are superior to the reference method and the closed form solution. A similar performance of the two methods indicates that our proposed self-calibration is able to accurately estimate the intrinsic system parameters. Thus, a clear benefit of the proposed method is its independence of a calibration scan, which is needed in case of the extended reference method. This will greatly improve the clinical work flow, required for weight-bearing acquisitions. The improved image quality of the extended reference method and of the novel approach will facilitate further processing steps, e.g. measuring cartilage deformation. Both methods estimate a coherent rigid motion using all attached markers, although the knees move independently. Future work could extend the estimation model to non-rigid motion to allow for independent knee motion.

5. REFERENCES

[1] J.-H. Choi, R. Fahrig, A. Keil, T. F. Besier, S. Pal, E. J. McWalter, G. S. Beaupré, and A. Maier, "Fiducial marker-based correction for involuntary motion in weight-bearing c-arm ct scanning of knees. part i. numerical model-based optimization," *Med Phys*, vol. 40, no. 9, pp. 091905, 2013.

[2] 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. part ii. experiment," *Med Phys*, vol. 41, no. 6, pp. 061902, 2014.

[3] 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 Proc*, 2015, pp. 58–61.

[4] M. Berger, A. Maier, Y. Xia, J. Hornegger, and R. Fahrig, "Motion Compensated Fan-Beam CT by Enforcing Fourier Properties of the Sinogram," in *Proc. CT Meeting 2014*, Frederic Noo, Ed., 2014, pp. 329–332.

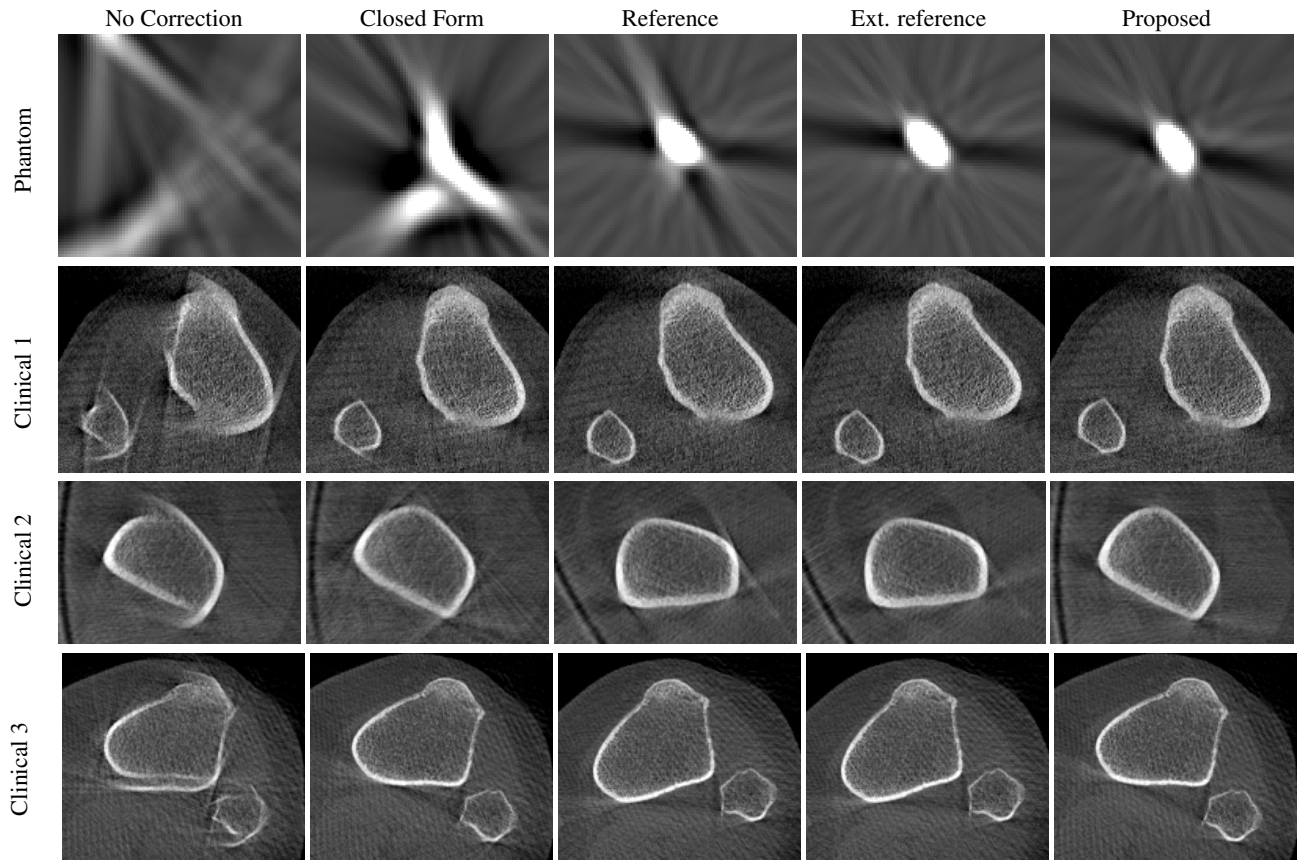


Fig. 2: ROI reconstruction for the different methods and datasets.

[5] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. Maier, "Epipolar Consistency in Transmission Imaging," *IEEE Trans Med Imaging*, vol. 34, no. 10, pp. 1–15, 2015.

[6] R. Clackdoyle, S. Rit, J. Hoskovec, and L. Desbat, "Fanbeam data consistency conditions for applications to motion detection," *Proc. CT Metting*, pp. 324–328, 2014.

[7] M. Berger, K. Müller, A. Aichert, M. Unberath, J. Thies, J.-H. Choi, R. Fahrig, and A. Maier, "Marker-free motion correction in weight-bearing cone-beam ct of the knee joint," *Med Phys*, vol. 43, no. 3, pp. 1235–1248, 2016.

[8] A. Sisniega, J.W. Stayman, Q. Cao, J. Yorkston, J.H. Siewerdsen, and W. Zbijewski, "Image-based motion compensation for high-resolution extremities cone-beam ct," in *Proc. SPIE Medical Imaging*, 2016, pp. 97830K–97830K.

[9] S. Sawall, M. Knaup, and M. Kachelrieß, "A robust geometry estimation method for spiral, sequential and circular cone-beam micro-ct," *Med Phys*, vol. 39, no. 9, pp. 5384–5392, 2012.

[10] S. Ouadah, J. W. Stayman, G. J. Gang, T. Ehtiati, and J. H. Siewerdsen, "Self-calibration of cone-beam ct geometry using 3d–2d image registration," *Phys Med Biol*, vol. 61, no. 7, pp. 2613, 2016.

[11] W. Wein, A. Ladikos, and A. Baumgartner, "Self-calibration of geometric and radiometric parameters for cone-beam computed tomography," in *Fully3D 2011 Proceedings*, 2011, vol. 2.

[12] F. Albiol, A. Corbi, and A. Albiol, "Geometrical calibration of x-ray imaging with rgb cameras for 3d reconstruction," *IEEE Trans Med Imaging*, vol. 35, no. 8, pp. 1952–1961, Aug 2016.

[13] M. Mitschke and N. Navab, "Recovering the x-ray projection geometry for three-dimensional tomographic reconstruction with additional sensors: Attached camera versus external navigation system," *Med Image Anal*, vol. 7, no. 1, pp. 65–78, 2003.

[14] 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*, 2014, pp. 168–173.

[15] 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 *Proc. SPIE Vol. 7961*, SPIE, Ed., 2011, pp. 7961231–7961238.

[16] A. Maier, H. 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," *Med Phys*, vol. 40, no. 11, pp. 111914, 2013.