Self-Calibration and Simultaneous Motion Estimation for C-Arm CT Using Fiducial Markers

Christopher Syben[1], Bastian Bier[1], Martin Berger[1], André Aichert[1]
Rebecca Fahrig[2], Garry Gold[2], Marc Levenston[2] and Andreas Maier[1]

[1]Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg
[2]Radiological Sciences Lab, Stanford University, Stanford, USA
christopher.syben@fau.de

Abstract. C-arm cone-beam CT systems have an increasing popularity in the clinical environment due to their highly flexible scan trajectories. Recent work used these systems to acquire images of the knee joint under weight-bearing conditions. During the scan, the patient is in a standing or in a squatting position and is likely to show involuntary motion, which corrupts image reconstruction. The state-of-the-art fully automatic motion compensation relies on fiducial markers for motion estimation. Due to the not reproducible horizontal trajectory, the system has to be calibrated with a calibration phantom before or after each scan. In this work we present a method to incorporate a self-calibration into the existing motion compensation framework without the need of prior geometric calibration. Quantitative and qualitative evaluations on a numerical phantom as well as clinical data, show superior results compared to the current state-of-the-art method. Moreover, the clinical workflow is improved, as a dedicated system calibration for weight-bearing acquisitions is no longer required.

1 Introduction

The high flexibility of C-arm cone-beam CT (CBCT) systems allow their usage in a wide range of new applications. Recently, these systems have been used to acquire data from knee joints under weight-bearing conditions [1, 2, 3]. For this purpose, the C-arm has to move on a horizontal trajectory around the standing patient. During the scan, involuntary patient motion can occur, which causes blurring, double edges and streaks in the 3D image reconstruction. Estimation and compensation of patient motion improves the quality of the reconstructed images. The state-of-the-art method estimates motion based on fiducial markers, which are attached to the patients knee [4].

However, the method requires a time consuming calibration with a calibration phantom for each scan, since the horizontal trajectory is not supported and thus, not reliably reproducible with the used C-arm system [5]. A self-calibration

K.H. Maier-Hein et al. (Hrsg.), *Bildverarbeitung für die Medizin 2017*,
Informatik aktuell, DOI 10.1007/978-3-662-54345-0_19

approach would be beneficial in such a setting. Current approaches can be divided into methods, which use external tools, like calibration markers or tracking systems [6], or rely only on the acquired data [7, 8].

The proposed approach in this work uses fiducial markers to calibrate the system, while simultaneously compensating for the patient's motion. Hence, dedicated time consuming calibration scans are dispensable.

2 Materials and methods

2.1 State-of-the-art motion compensation framework

The state-of-the-art motion estimation framework is introduced in the following [4, 3]. First, a reference 3D marker position is estimated for each marker by backprojecting the detected 2D marker positions into the volume. Then, the 3D reference positions are registered with the detected 2D positions [9]. Afterwards, the rigid motion consisting of three translation and three rotation parameter is estimated, such that the reprojection error (RPE) of the projected 3D reference marker positions on the 2D detected marker is minimized

$$\begin{aligned} \arg\min_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}) &= \arg\min_{\boldsymbol{\alpha}} \frac{1}{2} \sum_{j=1}^{J} \sum_{i=1}^{M} ||h(\boldsymbol{n}) - \boldsymbol{u}_{ij}||_2^2 \\ \boldsymbol{n} &= (n_1 \quad n_2 \quad n_3)^\top = \mathbf{P}_j \cdot \mathbf{M}_j(\boldsymbol{\alpha}) \cdot (\boldsymbol{v}_i \quad 1)^\top \end{aligned} \tag{1}$$

where vector $\boldsymbol{\alpha} \in \mathbb{R}^{6J}$ contains three rotation and translation parameters for each projection. The matrix $\mathbf{M}_j(\alpha)$ applies the rigid motion to the calibrated projection matrix $\mathbf{P}_j \in \mathbb{R}^{3\times4}$ for projection j. The estimated i-th 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, i.e., a division by the third component $h(\boldsymbol{n}) = \left(\frac{n_1}{n_3} \quad \frac{n_2}{n_3}\right)^\top$. In a last step, motion is compensated by incorporating the estimated motion into the projection matrices and using these updated projection matrices for the reconstruction [3, 4].

2.2 Joint motion estimation and system calibration

We face two problems if no calibration scan is performed and thus no initial estimation of the projection matrices is available. Valid system matrices are needed for the backprojection in the marker detection and for the forward projection to evaluate the objective function. To overcome the missing calibration, we propose to initialize with an ideal circular trajectory and to decompose the projection matrices into an extrinsic and intrinsic matrix, such that the intrinsic parameter estimation can be incorporated in the estimation process.

Initialization. The projection matrices $\mathbf{P}_j$ are initialized with the ideal horizontal circular trajectory based on the systems properties. The 3D marker detection and the evaluation of the objective function can be performed sufficiently well.

Intrinsic camera model. For the self-calibration of the system, the projection matrices need to be estimated. The decomposed projection matrix is $\mathbf{P} = \mathbf{K}\big[\mathbf{R}|\mathbf{t}\big]$, where $\mathbf{K} \in \mathbb{R}^{3\times3}$ is the intrinsic camera matrix and $\mathbf{R} \in \mathbb{R}^{4\times3}$ and $\mathbf{t} \in \mathbb{R}^{4\times1}$ contain the extrinsic parameters. For the extrinsic matrix, three rotation and three translation parameters are needed. For the projection matrices, it is not distinguishable whether the patient or the CT system itself is moving. These parameters are able to cover both rigid scanner and patient motion. The intrinsic camera matrix has five degree of freedom. Assuming isotropic detector pixels it can be further reduced to three parameters, i.e., the focal length f and the location of the principal point c_x and c_y. These parameters are estimated using an extended model by Wein et al. [7], which is well suited to the source detector geometry. The model is given by

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

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 to which the detector is tilted with respect to the plane orthogonal to the principal axis around the vertical detector axis $\boldsymbol{v}$. Similarly angle θ describes the rotation around the horizontal detector axis $\boldsymbol{u}$.

Joint motion and system estimation. To achieve a simultaneous motion and system calibration, the calibrated projection matrices in Eq. 1 are replaced with the decomposition of $\mathbf{P}_j$. The cost function for the joint estimation is given by

$$\begin{aligned} \underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min}\ f(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\arg\min}\ \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{M} ||h(\boldsymbol{n}) - \boldsymbol{u}_{ij}||_2^2 \\ \boldsymbol{n} = \boldsymbol{P}_j(\boldsymbol{\alpha},\boldsymbol{\beta}) \cdot (\boldsymbol{v}_i \quad 1)^\top &= \mathbf{K}_j(\boldsymbol{\beta}) \cdot \Big[\mathbf{R}_j\,|\boldsymbol{t}_j\Big] \cdot \mathbf{M}_j(\boldsymbol{\alpha}) \cdot (\boldsymbol{v}_i \quad 1)^\top \end{aligned} \tag{3}$$

where vector $\boldsymbol{\beta} \in \mathbb{R}^{3J}$ contains the three intrinsic parameters of presented model, for each projection. Hence, $\mathbf{K}_\beta$ describes the intrinsic matrix for the j-th projection. The patient motion and the mechanical deviations from the system are applied by multiplying with $\mathbf{M}_\alpha$, which describes the deviation from the ideal circular trajectory. Further $\big[\mathbf{R}_j|\boldsymbol{t}_j\big]$ describes the initial circular trajectory and is fixed during optimization.

The function is solved using a gradient-based optimizer. An analytical derivative of the objective function is computed, which achieves a remarkable speedup. The optimization is carried out in two steps. First, the cost function is minimized with respect to $\boldsymbol{\beta}$ and afterwards $\mathbf{K}_\beta$ is fixed and the function is optimized w.r.t. $\boldsymbol{\alpha}$. Additionally, an iterative scheme is applied, where the new projection

matrices are used to obtain new, more precise 3D marker positions. In the next iteration the updated 3D marker positions are used for the optimization. After two iterations the resulting projection matrices contain the patient motion as well as the system matrices for the scan.

2.3 Data and experiments

For the evaluation a numerical phantom and clinical data from three healthy patients are used. The phantom models a simplified leg, that consists of three encapsulated cylinders with a radii of 36, 40, and 100 mm and attenuations of bone marrow, femur and water, respectively. Beads, with an attenuation of stainless steel, are included into the phantom on a helical trajectory on the surface of the outer cylinder. Additionally, a wire with a diameter of 0.2 mm and with material set to bone is placed in the isocenter, pointing in direction $(1\ 1\ 1)^\top$ [8]. The projections are generated with the CONRAD framework [10] using calibrated projection matrices with additional estimated patient motion from clinical scans. Additionally, clinical datasets were acquired with a Siemens Artis zeego system (Siemens Healthcare GmbH, Forchheim, Germany). The scan range of the acquisition is 200° acquiring 248 projections in 10 s. Each projection image has 1240×960 pixels with isotropic pixel spacing of 0.308 mm.

For evaluation we use quantitative metrics and visual inspection. The first metric is the RPE, which is the result of the cost function. The second is the Full-width at Half-Maximum (FWHM) obtained averaging 720 line profiles through the cross-section of the wire in case of the numerical phantom and the metallic beads in case of the clinical datasets. Note for visual inspection that row one is rotated due to different initialized trajectories.

3 Results

We compared the proposed method with the results of the state-of-the-art method and the results if no correction is applied. The first column in Fig. 3 shows reconstructions of the phantom dataset with the wire in the ROI. Without correction, motion artifacts are clearly visible: streaks are present and the shape of the wire cannot be identified. The state-of-the-art method and the proposed approach were able to reconstruct the elliptical cross-section of the wire. However, our method could reconstruct the shape more clear.

The RPE's for the phantom dataset are shown in Tab. 1. Both correction methods could improve registration accuracy, where we obtained the best results for the proposed method. The FWHM results are shown in Tab. 2. Without application of motion correction, the FWHM value could not be measured, due to strong artifacts. The state-of-the-art and the proposed method achieve comparable FWHM results.

Columns two to four in Fig. 3 show reconstructions of clinical datasets. Without any correction the reconstructions of all three datasets are affected by streaking artifacts and the bone outline can not be restored properly. Both

Table 1. RPE in pixel for the described methods and datasets.

Dataset	No Correction	Berger et al.	Proposed Method
Phantom	84.85	1.36	0.07
Clinical 1	96.70	6.07	0.25
Clinical 2	71.29	0.72	0.19
Clinical 3	38.20	0.59	0.43

the state-of-the-art method and the proposed method are able to improve image quality at the bone structures for the clinical datasets. However, the proposed method could further eliminate residual streaking artifacts that could not be fully corrected by the state-of-the-art approach.

Without correction, the highest RPE values for the clinical datasets can be obtained (Tab. 1). The state-of-the-art method achieves for all three clinical datasets good RPE values. However, the RPE values of the proposed method are the best for all three clinical datasets. The FWHM results in Tab. 2 for the clinical dataset 1 are similar for the state-of-the-art method and the proposed method. For clinical datasets 2 and 3 the median FWHM and the standard deviation (std) of the proposed method are lower compared to the state-of-the-art-method.

4 Discussion

Incorporating a self-calibration component into the state-of-the-art method shows promising results. The proposed method is able to achieve reconstruction results, which are superior to the results of the state-of-the-art method, yet, it does not

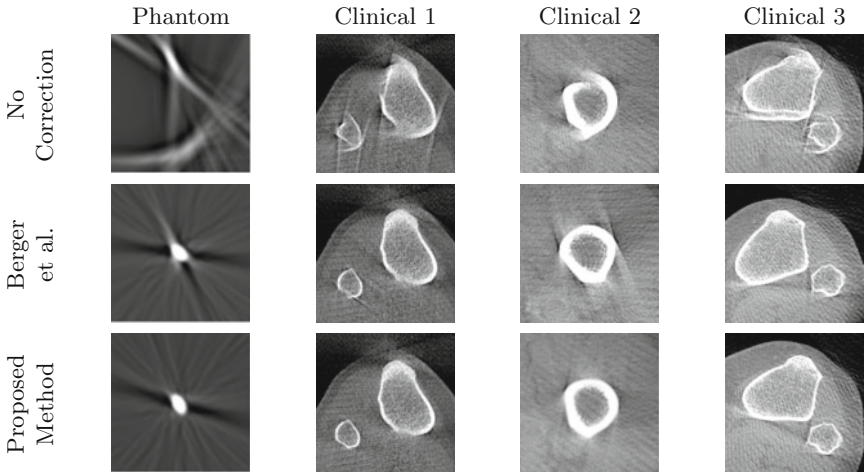

Fig. 1. ROI reconstruction for the described methods and datasets.

Dataset	Berger et al. (median ± std)	Proposed Method (median ± std)
Phantom	0.36 ± 1.80	0.35 ± 1.21
Clinical 1	0.62 ± 3.32	0.63 ± 3.11
Clinical 2	0.84 ± 2.68	0.81 ± 1.31
Clinical 3	0.88 ± 3.22	0.81 ± 1.57

Table 2. FWHM (median ± std) for the described methods and datasets.

require a separate calibration scan. The proposed method consists of two improvements: an self-calibration component and the iterative scheme for updated 3D marker positions. These improvements leads to the superior performance of the proposed method and to a less time consuming procedure. Future work will extend the state-of-the-art method with the iterative scheme using updated 3D reference marker positions to improve the estimation results.

References

1. Choi JH, Fahrig R, Keil A, et al. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees - Part I - Numerical model-based optimization. Med Phys. 2013;40(9):091905–n/a.
2. Choi JH, Maier A, Keil A, et al. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees - II - Experiment. Med Phys. 2014;41(6):061902–n/a.
3. Müller K, Berger M, Choi JH, et al. Automatic motion estimation and compensation framework for weight-bearing C-arm CT scans using fiducial markers. IFMBE Proc. 2015; p. 58–61.
4. Berger M, Müller K, Aichert A, et al. Marker-free motion correction in weight-bearing cone-beam CT of the knee joint. Med Phys. 2016;43(3):1235–48.
5. Maier A, Choi JH, Keil A, et al. Analysis of vertical and horizontal circular C-arm trajectories. Proc SPIE. 2011;7961:7961231–8.
6. Mitschke M, Navab N. Recovering the X-ray projection geometry for three-dimensional tomographic reconstruction with additional sensors: attached camera versus external navigation system. Med Image Anal. 2003;7(1):65–78.
7. Wein W, Ladikos A, Baumgartner A. Self-Calibration of geometric and radiometric parameters for cone-beam computed tomography. Fully3D Proc. 2011; p. 327–30.
8. Ouadah S, Stayman JW, Gang GJ, et al. Self-Calibration of cone-beam CT geometry using 3D–2D image registration. Phys Med Biol. 2016;61(7):2613.
9. Berger M, Forman C, Schwemmer C, et al. Automatic removal of externally attached fiducial markers in cone beam C-arm CT. Proc BVM. 2014; p. 168–73.
10. Maier A, Hofmann HG, Berger M, et al. CONRAD–A software framework for cone-beam imaging in radiology. Med Phys. 2013;40(11):111914.