Motion Compensation using Range Imaging in C-arm Cone-Beam CT

Bastian Bier[1], Mathias Unberath[1,2], Tobias Geimer[1,2], Jennifer Maier[1], Garry Gold[3], Marc Levenston[3], Rebecca Fahrig[3,*] and Andreas Maier[1,2]

[1]Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany
[2]Erlangen Graduate School in Advanced Optical Technologies, Friedrich-Alexander-University Erlangen-Nuremberg, Germany
[3]Department of Radiology, School of Medicine, Stanford University, USA
[*]Now with Siemens Healthcare GmbH, Erlangen, Germany

bastian.bier@fau.de

Abstract. Cone-beam C-arm CT systems allow to scan patients in weight-bearing positions to assess knee cartilage health under more realistic conditions. Involuntary patient motion during the acquisition results in motion artifacts in the reconstructions. The current motion estimation method is based on fiducial markers. They can be tracked with a high spatial accuracy in the projection images, but only deliver sparse information. Further, placement of the markers on the patient's leg is time consuming and tedious. Instead of relying on a few well defined points, we seek to establish correspondences on dense surface data to estimate 3D displacements.

In this feasibility study, motion corrupted X-ray projections and surface data are simulated. We investigate motion estimation by registration of the surface information. The proposed approach is compared to a motion free, an uncompensated, and a state-of-the-art marker-based reconstruction using the SSIM.

The proposed approach yields motion estimation accuracy and image quality close to the current state-of-the-art, reducing the motion artifacts in the reconstructions remarkably. Using the proposed method, Structural Similarity improved from 0.887 to 0.975 compared to uncorrected images. The results are promising and encourage future work aiming at facilitating its practical applicability.

1 Introduction

C-arm cone-beam CT (CBCT) offers a great variety of applications in interventional and diagnostic medicine. One of the reasons for that is their high degree of flexibility, which allows the C-arm to scan objects on arbitrary trajectories, even enabling scanning of patients in upright, weight-bearing positions [6, 7]. Weight-bearing imaging was found to be particularly beneficial for the assessment of knee cartilage health, since the knee joint has different mechanical properties under load [19]. To this end, the C-arm rotates around the standing

patient on a horizontal trajectory and acquires images from different views [13]. A schematic scheme of the imaging geometry is shown in Figure 1. One of the major problems is that patients tend to show motion during the scan time of approximately ten seconds. This results in motion artifacts in the reconstructed volumes, which manifest as streaks, double edges, and blurring, as can be seen in Figure 3(b). In order to increase the image quality and the diagnostic value of the reconstructions, motion has to be estimated and compensated for.

Recently, motion correction in extremity imaging under weight-bearing conditions gained increasing attention. Choi et al. and Berger et al. use fiducial markers placed on the patient's knees. These markers can be detected in the 2D projection images and their position can be used for alignment with their respective 3D reference position. Although this method has been used in various clinical studies [7, 16, 17, 4], several limitations exist. Marker placement is time consuming, associated with patient discomfort, and tedious since overlapping markers in the projections would result in inaccurate motion estimates. Another category of approaches do not rely on markers and are image-based only. Unberath et al. estimate motion by aligning the projection images with the maximum intensity projections of a motion corrupted reconstruction [22]. However, only 2D detector shifts are evaluated. Berger et al. uses 2D/3D registration of a segmented bone from a motion free supine acquisition with the acquired projection images [4, 18]. The reconstructions show a much sharper bone outline, however, a previously acquired supine scan might not always be available. Sisniega et al. proposed a method to estimate 6D rigid motion by optimizing the image sharpness and entropy in a region of interest in the reconstruction [20, 21]. Further, epipolar consistency conditions were investigated to estimate motion and showed promising results [1, 5]. However, motion estimation proved to be robust only parallel to the detector.

Range imaging has been used in different applications in medicine [2] such as augmented reality [10], motion estimation in PET/SPECT imaging [15], and patient positioning and motion estimation in radiotherapy [12]. Recently, Fotouhi et al. investigated the feasibility on improving iterative reconstruction quality by incorporating information of a RGBD camera [11]. In this work, a feasibility study is conducted to investigate the motion estimation capability of range imaging in the scenario of acquisitions under weight-bearing conditions. Therefore, surface point cloud data and X-ray projections are simulated under the influence of motion. Using registration of the point clouds, a motion estimate is obtained. We compare the reconstructed images of the proposed method with the results of the uncompensated reconstruction and the result of a marker-based method [7]. Further, the estimated motion signals are compared to the marker-based method using the correlation coefficient. To investigate the effects of motion compensation on the image quality, we compare the reconstructed images of the uncompensated case, the marker-based method and the proposed approach to the ground truth using Structural Similarity (SSIM).

2 Materials and Methods

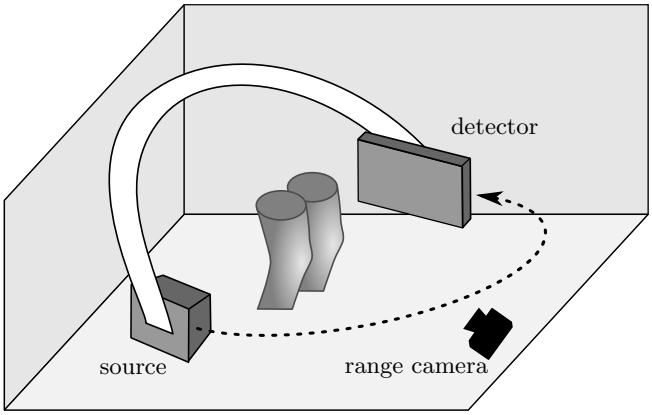
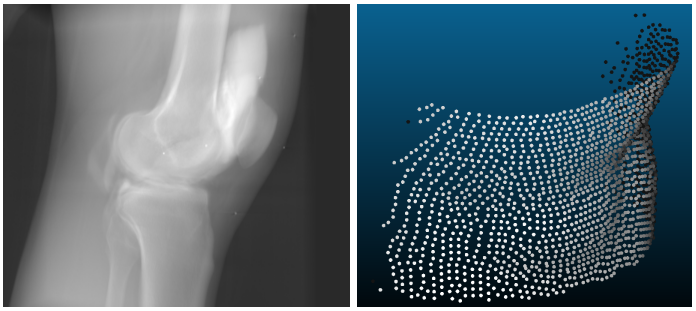


Fig. 1. Schematic scheme of the imaging setup. The X-ray source and the detector rotate around the standing patient. A range imaging camera is placed statically on the ground in front of the object.

The method relies on 3D point clouds and X-ray projection images generated for each time point in the same motion state. Exemplary input data are shown in Figure 2. The section is structured as follows: First, we describe the range camera simulation; second, we present the point cloud registration, and finally, the X-ray simulation and reconstruction is described.

2.1 Depth Camera Simulation



(a) Projection image. (b) Point cloud.

Fig. 2. Simulated projection image (a) with its corresponding point cloud (b).

Camera position. Different camera positions are feasible in combination with a C-arm CT system. Either the camera is mounted on the detector or statically on the ground. If mounted on the C-arm, only one calibration between the two modalities is necessary. However, in the weight-bearing scenario occlusion from the second leg occurs. Further, the registration task is more difficult due to partial overlap. Especially for imaging cylindrical, smooth objects such as the knees, ambiguities in the surface will make the registration task difficult. For a static camera position, this is not the case. Therefore, in this study, the camera is placed in front of the standing patient, such that the object of interest is always completely in the field of view, as shown in Figure 1. Further, the registration problem simplifies since partial overlap at the surface border is smaller and some anatomical structures, e.g. the patella, are visible.

Creation of the point clouds. Depth images are simulated by casting rays through a mask volume, which is obtained from a motion free supine reconstruction, where one leg is segmented. The geometry of the camera is described by projection matrices similar to the ones of the C-arm. For each ray, a depth value is obtained, which can be represented in a depth map or a point cloud. The camera and its noise properties are selected to be similar to the Microsoft Kinect One v2 [24]: The sampled points lie on a grid like pattern, as shown in Figure 2(b). In the area of the object, the resolution is 160×120 pixels with an approximate distance of the sampling points of 1.4 mm in image space and a depth resolution of 1 mm. Noise free and noise corrupted point clouds are created for our experiments. For this purpose, Gaussian noise with a zero mean and a standard deviation of $\sigma = 1$ and $\sigma = 2$ is added on the depth coordinates. The camera to object distance is 75 cm.

2.2 Motion Simulation

In order to simulate rigid 3D translation on the depth camera, the volume is shifted for each time point by a reference motion vector in the 3D volume, resulting in a different point cloud for each time point. The applied reference motion is 3D translation derived from real patient motion captured with a motion capture system [6].

The same reference motion is used to create motion corrupted X-ray projections. To this end, the projection matrices $\mathbf{P} \in \mathbb{R}^{4\times3}$, which describe the imaging geometry of the C-arm system, are modified. Rigid motion can be directly incorporated in the projection matrices using the following formula:

$$\hat{\mathbf{P}}_j = \mathbf{P}_j \cdot \begin{pmatrix} \mathbf{R}_j & \mathbf{t}_j \\ \mathbf{0} & 1 \end{pmatrix} \tag{1}$$

$\mathbf{P}_j$ is the calibrated projection matrix from the system and $\mathbf{R}_j \in \mathbb{R}^{3\times3}$ and $\mathbf{t}_j \in \mathbb{R}^{1\times3}$ are the rotation and translation for each time step j, respectively. These motion corrupted projection matrices are used to create motion corrupted projection images, shown in Figure 2(a).

2.3 Point Cloud Registration

The registration is initialized by aligning the centroids of all point clouds. In order to speed up the registration, all point clouds are registered to the point cloud corresponding to the first frame that is assigned the translation vector $\mathbf{t}_0 = (0, 0, 0)$. Afterwards, a point-to-surface Iterative Closest Point (ICP) algorithm [3] is used to refine the estimation result. In contrast to the common ICP algorithm that optimizes for point-to-point distance, point-to-surface ICP seeks to minimize the distance of points to the other point cloud's surface. This strategy proved more accurate in this application. Further, the common ICP is heavily influenced by the grid pattern of the point clouds, which results in an estimate that is a multiple of the spacing between the points. The optimal motion estimate is found by solving the following objective function for a pair of point clouds:

$$\hat{\boldsymbol{t}} = \arg\min_{\boldsymbol{t}} \left(\sum_{i=1}^{N} ||((\boldsymbol{p}_i + \boldsymbol{t}) - \boldsymbol{q}_i)\boldsymbol{n}_i|| \right), \tag{2}$$

where $\hat{\boldsymbol{t}}$ is the translation to be estimated, N is the number of points in the source point cloud with $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ being a point on the source point cloud and its corresponding point on the target point cloud, respectively. The corresponding point is defined by the projection of $\boldsymbol{p}_i$ along the normal vector $\boldsymbol{n}_i$ of the target point cloud. This is implemented by triangulating the point cloud and computing the distance of the point to the respective triangles with a defined normal vector. The normal vector is obtained from the cross product of two vectors of the triangle. The function is solved with a gradient descent algorithm[1].

2.4 Reconstruction

Images are reconstructed using the Feldkamp-David-Kress backprojection algorithm [9]: the projection data is preprocessed using cosine weighting, Parker redundancy weighting, and row-wise ramp filtering. In a last step the projections are backprojected. The projection matrices used are real calibrated projection matrices from a clinical C-arm system, which has been operated on a horizontal trajectory. In order to incorporate the motion estimate into the reconstruction, Equation 1 can be used to modify the projection matrices before the backprojection step.

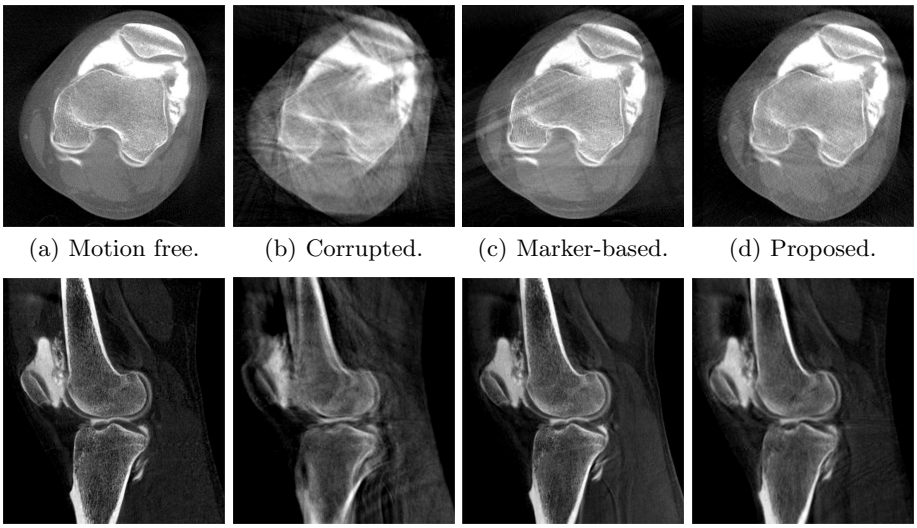
2.5 Experiment

The simulation is performed on a segmented knee, extracted from a clinical high quality supine reconstruction. The data was acquired on a clinical C-arm CT system (Artis Zeego, Siemens Healthcare GmbH, Erlangen, Germany). We use the real scanner geometry with 248 projection images on a horizontal trajectory over

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

a rotation of 200°. Detector resolution is 1240×960 pixels with an isotropic pixel spacing of 0.308 mm. Volumes with a size of 512^3 and an isotropic voxelsize of 0.5 mm are reconstructed. The entire processing, simulation, and reconstruction pipeline is implemented in the open-source framework CONRAD, dedicated to the simulation and reconstruction of CBCT data [14]. In order to compare the image quality of the results, the SSIM [23] is computed. To this end, all reconstructions are registered to the motion free reference reconstruction using the open-source software 3D slicer [8].

3 Results



(a) Motion free. (b) Corrupted. (c) Marker-based. (d) Proposed.

(e) Motion free. (f) Corrupted. (g) Marker-based. (h) Proposed.

Fig. 3. Axial and sagittal slices of the reconstruction results of the motion free reference volume (a), the motion corrupted (b), the marker-based corrected (c), and the proposed method (d).

In Figure 3, axial and sagittal slices of the reconstruction results of the motion free scan, the motion corrupted acquisition, the marker-based method, and the proposed method (on noise free data) are shown. Severe streaking and blurring artifacts are present if no correction is applied. Both, the marker-based method as well as the new proposed method substantially reduce the corruption by artifacts. In the marker-based result (Figure 3(c)) streaks from one direction are visible. They appear due to marker overlap in these views resulting in an inaccurate estimation. In contrast, images obtained using the proposed approach have more but smaller streak artifacts, which appear especially at the bone outline.

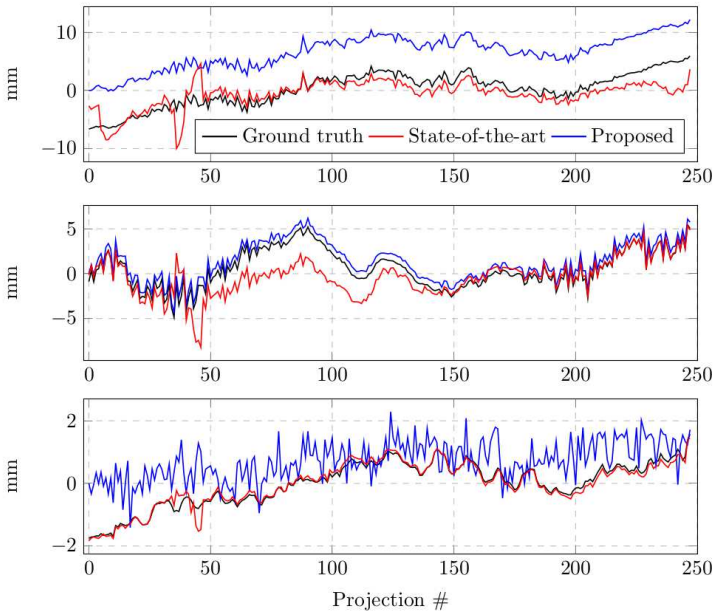


Fig. 4. Ground truth motion parameters compared to the estimation of the marker-based and the proposed method for the x (top), y (middle), and z (bottom) direction.

In Figure 4, the estimated motion parameters of the marker-based and proposed method are compared with the ground truth motion: x is the object motion parallel to the detector, y is the depth and z corresponds to the direction parallel to the rotation axis. Considering the estimation results for x and y direction, both methods were able to recover the motion well, especially the high frequencies. Note, however, that due to the current optimization scheme of the proposed method, all projections are aligned to the first frame resulting in a constant offset in the estimated signal. This offset will cause a shift of the reconstruction in the volume. The views in which the markers overlap are visible in the motion estimation around projection number 40, where sudden jumps are visible in the marker-based estimation. They correspond to the heavy streaks in the reconstruction images. In contrast, the z direction is estimated poorly with the proposed method. One reason for that might be that the cylindrical shape does not contain enough information to nicely align the point clouds. Further, the object is also truncated in the z direction, which might lead to these errors.

In Table 1 the SSIM between the shown reconstructions compared to the motion free volume are shown. The quantitative values are in agreement with the visual impression of the reconstructions: the marker-based and the proposed method show comparably good results that are far superior to the uncompensated reconstruction. The SSIM decreases with the level of noise applied to the

Table 1. Structural Similarity (SSIM) results of the reconstructed images.

Method	SSIM
Uncorrected	0.887
Marker-based [7]	0.971
Proposed	0.975
Proposed with noise $\sigma = 1$	0.961
Proposed with noise $\sigma = 2$	0.952

Table 2. Correlation coefficients (CC) between the estimated and the ground truth motion signals.

Method	x-direction	y-direction	z-direction
Marker-based [7]	0.77	0.72	0.98
Proposed	0.98	0.98	0.59

point clouds. However, achieved results are superior towards the uncorrected reconstruction.

In order to quantify the accuracy of the motion estimation, we calculated the correlation coefficient between the estimated and the ground truth signal. The results are shown in Table 2, and show that the proposed method outperforms the marker-based method in the motion estimate of the x and y direction, but achieved only poor results in the z direction with a correlation coefficient of 0.58 only.

4 Discussion and Conclusion

In this feasibility study, a novel method to estimate patient motion in acquisitions under weight-bearing conditions using range imaging is investigated. Our method registers point clouds to a reference frame to estimate motion, without the need to place markers on the patient's skin. We validated our approach on simulated projection and point cloud data of a high quality supine reconstruction. The method reduces the motion artifacts notably and suggests high potential of using depth imaging for motion correction in CBCT reconstruction.

Surface and projection data are simulated with the same motion pattern for a full acquisition. All point clouds are registered using a point-to-surface ICP algorithm resulting in a motion estimate used to obtain the motion corrected reconstruction. In the performed experiments, the proposed method showed an improvement of the image quality comparable to the marker-based approach.

Future work will address the current limitations of the method. Compared to the 6D rigid motion estimation of the marker-based method, we tested translational motion only in our experiments that does not accurately reflect real knee motion. However, the method might have the potential to estimate more complex non-rigid motion of the knee as the dense surface information could be used to estimate more sophisticated displacements. Currently, motion along z direction is estimated insufficiently. One possibility to address this issue would

be to combine the presented approach with image-based consistency conditions, which usually estimate this direction robustly. Further, a more comprehensive study on the method's robustness against sensor noise has to be conducted. This is especially important when dealing with real depth data containing a high amount of noise and physical artifacts. Finally, for practical applicability in the dynamic clinical environment, the camera has to be placed such that the field of view is not temporarily obstructed and in a distance, which allows to acquire data in good quality. Synchronization and cross calibration of the depth camera with the C-arm system have to be considered.

Despite the mentioned limitations, this work is yet another step towards motion estimation in imaging under weight-bearing conditions without the need of prior knowledge or the placement of fiducial markers.

References

1. Aichert, A., Wang, J., Schaffert, R., Dörfler, A., Hornegger, J., Maier, A.: Epipolar Consistency in Fluoroscopy for Image-Based Tracking. In: Proc BMVC. pp. 82.1–82.10 (2015)
2. Bauer, S., Seitel, A., Hofmann, H.G., Blum, T., Wasza, J., Balda, M., Meinzer, H.P., Navab, N., Hornegger, J., Maier-Hehn, L.: Real-Time Range Imaging in Health Care: A Survey. Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications Lecture Notes in Computer Science pp. 228–254 (2013)
3. Bellekens, B., Spruyt, V., Weyn, M.: A Survey of Rigid 3D Pointcloud Registration Algorithms. AMBIENT 2014, The Fourth International Conference on Ambient Computing, Applications, Services and Technologies. 2014. pp. 8–13 (2014)
4. Berger, M., Müller, K., Aichert, A., Unberath, M., Thies, J., Choi, J.H., Fahrig, R., Maier, A.: Marker-free motion correction in weight-bearing cone-beam CT of the knee joint. Medical Physics 43(3), 1235–1248 (2016)
5. Bier, B., Aichert, A., Felsner, L., Unberath, M., Levenston, M., Gold, G., Fahrig, R., Maier, A.: Epipolar Consistency Conditions for Motion Correction in Weight-Bearing Imaging. In: Bildverarbeitung für die Medizin 2017 (2017)
6. Choi, J.H., Fahrig, R., Keil, A., Besier, T.F., Pal, S., McWalter, E.J., Beaupré, G.S., Maier, A.: Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization. Medical Physics 41(6), 061902 (2014)
7. Choi, J.H., Maier, A., Keil, A., Pal, S., McWalter, E.J., Beaupré, G.S., Gold, G.E., Fahrig, R.: Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. II. Experiment. Medical Physics 41(6), 061902 (2014)
8. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R.: 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic Resonance Imaging 30(9), 1323–1341 (2012)
9. Feldkamp, L.a., Davis, L.C., Kress, J.W.: Practical cone-beam algorithm. Journal of the Optical Society of America A 1(6), 612 (1984)
10. Fischer, M., Fuerst, B., Lee, S.C., Fotouhi, J., Habert, S., Weidert, S., Euler, E., Osgood, G., Navab, N.: Preclinical usability study of multiple augmented reality

concepts for K-wire placement. International Journal of Computer Assisted Radiology and Surgery 11(6), 1007–1014 (2016)
11. Fotouhi, J., Fuerst, B., Wein, W., Navab, N.: Can real-time RGBD enhance intraoperative Cone-Beam CT? International Journal of Computer Assisted Radiology and Surgery (2017)
12. Geimer, T., Unberath, M., Taubmann, O., Bert, C., Maier, A.: Combination of Markerless Surrogates for Motion Estimation in Radiation Therapy. In: Computer Assisted Radiology and Surgery (CARS) 2016. pp. 59–60 (2016)
13. Maier, A., Choi, J.H., Keil, A., Niebler, C., Sarmiento, M., Fieselmann, A., Gold, G., Delp, S., Fahrig, R.: Analysis of vertical and horizontal circular C-arm trajectories. SPIE Medical Imaging 7961, 796123–1 – –796123–8 (2011)
14. Maier, A., Hofmann, H.G., Berger, M., Fischer, P., Schwemmer, C., Wu, H., Müller, K., Hornegger, J., Choi, J.H., Riess, C., Keil, A., Fahrig, R.: CONRAD - A software framework for cone-beam imaging in radiology. Medical Physics 40(11), 111914 (2013)
15. McNamara, J.E., Pretorius, P.H., Johnson, K., Mukherjee, J.M., Dey, J., Gennert, M.A., King, M.A.: A flexible multicamera visual-tracking system for detecting and correcting motion-induced artifacts in cardiac SPECT slices. Medical Physics 36(5), 1913–1923 (2009)
16. Müller, K., Berger, M., Choi, J., Maier, A., Fahrig, R.: Automatic Motion Estimation and Compensation Framework for Weight-bearing C-arm CT scans using Fiducial Markers. In: IFMBE Proceedings. pp. 58–61 (2015)
17. Müller, K., Berger, M., Choi, J.h., Datta, S., Gehrisch, S., Moore, T., Marks, M.P., Maier, A.K., Fahrig, R.: Fully Automatic Head Motion Correction for Interventional C-arm Systems using Fiducial Markers. In: Proceedings of the 13th Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine. pp. 534–537 (2015)
18. Ouadah, S., Stayman, J.W., Gang, G.J., Ehtiati, T., Siewerdsen, J.H.: Self-calibration of cone-beam CT geometry using 3D2D image registration. Physics in Medicine and Biology 61(7), 2613–2632 (2016)
19. Powers, C.M., Ward, S.R., Fredericson, M.: Knee Extension in Persons With Lateral Subluxation of the Patella : A Preliminary Study. Journal of Orthopaedic and Sports Physical Therapy 33(11), 677–685 (2013)
20. Sisniega, A., Stayman, J.W., Cao, Q., Yorkston, J., Siewerdsen, J.H., Zbijewski, W.: Image-based motion compensation for high-resolution extremities cone-beam CT. SPIE Medical Imaging 9783, 97830K (2016)
21. Sisniega, A., Stayman, J., Yorkston, J., Siewerdsen, J., Zbijewski, W.: Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion. Physics in Medicine and Biology 62(9) (2017)
22. Unberath, M., Choi, J.H., Berger, M., Maier, A., Fahrig, R.: Image-based compensation for involuntary motion in weight-bearing C-arm cone-beam CT scanning of knees. In: SPIE Medical Imaging. vol. 9413, p. 94130D (2015)
23. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
24. Wasenmüller, O., Stricker, D.: Comparison of Kinect v1 and v2 Depth Images in Terms of Accuracy and Precision. In: Asian Conference on Computer. pp. 1–12 (2016)