Motion Compensation for Cone-Beam CT Using Fourier Consistency Conditions

M. Berger[1,2], Y. Xia[1,3], W. Aichinger[1], K. Mentl[1], M. Unberath[1,3], A. Aichert[1], C. Riess[1,2], J. Hornegger[1,2,3], R. Fahrig[4], A. Maier[1,2,3]

[1]Pattern Recognition Lab, Friedrich-Alexander-Universtät Erlangen-Nürnberg, 91058 Erlangen, Germany
[2]Graduate School 1773, Heterogeneous Image Systems, 91058 Erlangen, Germany
[3]Erlangen Graduate School in Advanced Optical Technologies (SAOT), 91052 Erlangen, Germany
[4]Radiological Sciences Laboratory, Stanford University, Stanford, California 94305

E-mail: martin.berger@cs.fau.de

May 26, 2017

Abstract. In cone-beam CT, involuntary patient motion and inaccurate or irreproducible scanner motion substantially degrades image quality. To avoid artifacts this motion needs to be estimated and compensated during image reconstruction. In previous work we showed that Fourier Consistency Conditions (FCC) can be used in fan-beam CT to estimate motion in the sinogram domain. This work extends the FCC to 3-D cone-beam CT. We derive an efficient cost function to compensate for 3-D motion using 2-D detector translations.

The extended FCC method have been tested with five translational motion patterns, using a challenging numerical phantom. We evaluated the root-mean-square-error (RMSE) and the structural-similarity-index (SSIM) between motion corrected and motion-free reconstructions. Additionally, we computed the mean-absolute-difference (MAD) between the estimated and the ground-truth motion. The practical applicability of the method is demonstrated by application to respiratory motion estimation in rotational angiography, but also to motion correction for weight-bearing imaging of knees. Where the latter makes use of specifically modified FCC version which is robust to axial truncation.

The results show a great reduction of motion artifacts. Accurate estimation results were achieved with a maximum MAD value of 708 µm and 1184 µm for motion along the vertical and horizontal detector direction, respectively. The image quality of reconstructions obtained with the proposed method is close to that of motion corrected reconstructions based on the ground-truth motion. Simulations using noise-free and noisy data demonstrate that FCC are robust to noise. Even high-frequency motion was accurately estimated leading to a considerable reduction of streaking artifacts. The method is purely image-based and therefore independent of any auxiliary data.

PACS numbers: XXXXX

Submitted to: *Phys. Med. Biol.*

1. Introduction

Classical CT reconstruction algorithms assume a static object and an ideal scanning trajectory. In many clinical applications this assumption is violated due to movements of the patient, such as respiratory motion (Rit et al.; 2009), leading to inconsistencies in the projection data. Depending on the amplitude and frequency of the underlying motion, this can cause streaking and blurring artifacts in the reconstruction.

Different approaches exist to reduce motion-induced artifacts. Typically, the first step is to detect and estimate the motion, followed by a motion compensated reconstruction that takes the estimated motion into account. The majority of methods require additional auxiliary signals for motion estimation. In the field of cardiovascular imaging, artifacts can be reduced by clustering consistent projection images using an ECG signal (Schaefer et al.; 2006). Subsequently, the image may be further improved via registration in projection (Schwemmer et al.; 2013) or reconstruction domain (Müller et al.; 2014). Another type of motion estimation uses fiducial markers, that are identified in projection images and tracked over time (Li et al.; 2006; Choi et al.; 2014). This approach is known to be accurate but requires the attachment or even surgical placement of metallic markers to the anatomy of interest (Shirato et al.; 2003). A key problem of using auxiliary signals is that they usually reduce patient comfort and increase acquisition time thus hindering clinical workflow. Further, these signals are dedicated to very specific use cases or anatomies and may not be used in other acquisition types. For example, an ECG signal will not correlate with skeletal motion while implanting markers for conventional heart imaging does not seem reasonable. Another approach is to perform 2-D/3-D registration of a known volume or an initial reconstruction to the individual projection images (Zeng et al.; 2005; Gendrin et al.; 2012; Ouadah et al.; 2016; Berger et al.; 2016). Registration approaches have been investigated extensively for tumor tracking in the field of radiotherapy, where we refer to Rit et al. (2013) for a comprehensive review.

In principle, motion estimation methods relying on the acquired projection images only can alleviate most of these problems. Only little work has been done on purely image-based motion estimation. Some methods impose assumptions on the imaged object and optimize entropy or positivity measures for the 3-D reconstruction (Rohkohl et al.; 2013; Ens et al.; 2010). Others use mathematically formulated data consistency conditions (CC) that describe redundancies in projection or sinogram domain.

The best known CC are the Helgason-Ludwig consistency conditions (HLCC), derived by Helgason (1980) and Ludwig (1966). The HLCC state that after building image moments along the detector direction of a parallel-beam sinogram, the Fourier series expansion along the projection angles needs to be zero at well defined positions. An intuitive representation of different image moments is given by Welch et al. (1998) who use the HLCC for attenuation correction in PET acquisitions. Several extensions of the HLCC exist for the fan-beam geometry and have been used for motion estimation or detection by (Yu and Wang; 2007; Leng et al.; 2007; Clackdoyle and Desbat; 2015).

Recently, an extension to three-dimensional cone-beam geometries has been proposed by Clackdoyle and Desbat (2013). It requires that all X-ray source positions are on a plane which does not intersect the object. Therefore it is not yet applicable for circular cone-beam CT.

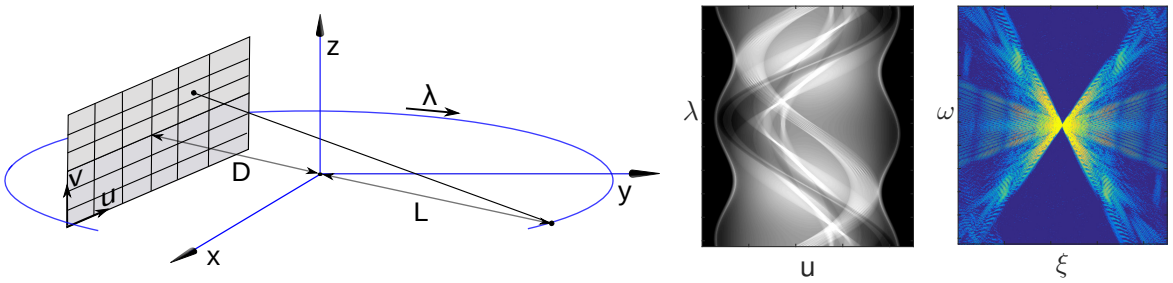
In addition to the HLCC there also exist FCC that are defined in the Fourier domain of the sinogram. It has been shown by Edholm et al. (1986) and Natterer (1986) that specific triangular regions in the spectrum of a sinogram have to have an absolute value close to zero. The size of these regions depends on the scanner geometry and the maximum extent of the object with respect to the system's center of rotation. Thus, knowing the object's extent the triangular zero regions can be precomputed. Extensions of FCC to the fan-beam geometry have been proposed by Hawkins et al. (1988); Natterer (1993); Mazin and Pelc (2010). Recently, we used the FCC for motion estimation in fan-beam CT (Berger et al.; 2014). Detector shifts are estimated for each projection angle, such that the energy in the triangular regions of the sinogram's spectrum is minimized. Initial work for an extension of the FCC theorem to cone-beam projections was presented by Brokish and Bresler (2006). However, to our knowledge, the FCC have never been used for artifact correction in cone-beam geometries.

Both HLCC and FCC have initially been proposed for parallel-beam geometries and were extended to fan- and cone-beam geometries. In contrast, John (1938) presented full CC directly for cone-beam geometries that are related to second-order partial differential equations. The theorem provides a system of ultrahyperbolic partial differential equations for the projection images. The applicability to an actual reconstruction was however limited. Yet, extensions to a more practical formulation have more recently been made by Patch (2002b) with the application of estimating missing projection data (Patch; 2002b,a).

In this work, we present an extension of the FCC to the cone-beam geometry and apply it for motion estimation of simulated cone-beam CT data. Besides the thorough evaluation of translational motion using a challenging numerical phantom, we apply our method for respiratory motion estimation in rotational angiography, which was recently published by Unberath et al. (2017). Additionally, we show its applicability to motion correction for cone-beam CT of knees under weight-bearing conditions. Parts of this work were recently shown in Berger (2016). The extension to cone-beam CT is based on the findings of Brokish and Bresler (2006). We define a cost function that estimates horizontal and vertical detector shifts to correct for translational object motion in 3-D. The optimization is gradient-based and uses an analytic gradient computation.

2. Methods and Materials

In the following we present the FCC and its extension to the 3-D cone-beam geometry in Section 2.1. Further, we describe a cost function for the estimation of arbitrary motion models in Section 2.2. An efficient variant of the cost function and its gradient with respect to the motion parameters is presented in Section 2.3. Implications of



(a) Cone-beam CT (b) Sinogram (c) Spectrum

Figure 1: Cone-beam CT geometry and example of triangular regions.

sampling and discretization are outlined in Section 2.4. Additionally, Section 2.5 provides information about the numeric optimization followed by an approach to increase the cost function's robustness to data truncation in Section 2.6. Finally, we present implementation details in Section 2.7.

2.1. Extended Fourier Properties of the Sinogram

It has been shown that the Fourier transform of a parallel- and fan-beam sinogram contains triangular regions with an absolute value close to zero (Edholm et al.; 1986; Natterer; 1986; Mazin and Pelc; 2010). An example of such regions is shown in Figure 1c, showing the logarithmically scaled spectrum of the sinogram in Figure 1b. The size and orientation of the empty triangular spectral regions depend on the maximum distance of the object to the center of rotation r_p, as well as the source-isocenter-distance L and the detector-isocenter-distance D. According to (Mazin and Pelc; 2010) the triangular regions for a flat detector fan-beam geometry can be described by

$$\left| \frac{\gamma}{\gamma - \xi(L+D)} \right| > \frac{r_p}{L} \quad , \tag{1}$$

where γ and ξ are the frequency variables corresponding to the projection angles and the detector rows, respectively. These triangular regions exist for perfect sinograms, that is, a perfectly circular scan without motion, truncation or beam-hardening. As shown by Berger et al. (2014) potential patient or scanner motion violates this requirement. Hence, the overall energy in the triangular spectral regions increases.

Figure 1a gives an overview of the cone-beam CT acquisition geometry and the corresponding mapping from the world to the detector coordinate system. Let ψ be the frequency variable that corresponds to the vertical detector directions. We now want to use the 2-D fan-beam properties of Eqn. 1 for the 3-D Fourier space of a cone-beam projection volume acquired with a circular source trajectory. Brokish and Bresler (2006) investigated the essential support of this 3-D Fourier domain and derived minimum sampling conditions that ensure an aliasing-free sampling of the object. The support is the opposite to the triangular zero regions and described by reversing the inequality of Eqn. 1. Further, Brokish and Bresler showed that for the majority of

cone-beam imaging devices it is sufficient to determine the support of γ for $\psi = 0$, i.e., in the fan-beam case. From their numerical simulation results it becomes clear that the variation of the support along the ψ-direction is negligible. Therefore, we use the fan-beam properties from Eqn. 1 and expand them in the direction of ψ to construct the triangular zero regions for the 3-D Fourier transform.

2.2. Cost Function for Arbitrary Motion Models

To minimize the energy in the regions described by Eqn. 1, we define a cost function for a set of K projection images of size $I \times J$, i.e., for a total of $M = I \cdot J \cdot K$ measurements. The cost function is given by

$$e(\boldsymbol{\alpha}) = \left\| \boldsymbol{W} \left(\boldsymbol{F}\, \boldsymbol{p}(\boldsymbol{\alpha}) \right) \right\|_2^2 \ , \tag{2}$$

where $\boldsymbol{F} \in \mathbb{C}^{M \times M}$ is a symmetric matrix that performs the 3-D discrete Fourier transform (DFT), $\boldsymbol{p}(\boldsymbol{\alpha}) \in \mathbb{R}^M$ is the projection data in vector format with respect to the motion parameters given by $\boldsymbol{\alpha}$, and $\boldsymbol{W} = \operatorname{diag}(\boldsymbol{w}) \in \mathbb{R}^{M \times M}$ is the diagonal matrix of vector $\boldsymbol{w} \in \mathbb{R}^M$ that represents the 3-D mask for the zero regions. The function $\operatorname{diag}(\cdot)$ takes a vector and returns a diagonal matrix with the vector elements on its diagonal. Further, let ξ_i, ψ_j and γ_k be the frequency values of the Fourier transforms' discrete axes, over detector u- and v-directions and the projection angles λ, respectively. Using the detector spacings Δu and Δv and the angular spacing $\Delta\lambda$ they can be computed by

$$\begin{aligned} \xi_i &= f(i, \Delta u, I) \\ \psi_j &= f(j, \Delta v, J) \qquad \text{with } f(l, \Delta, N) = \frac{l - 1 - \lfloor \frac{N}{2} \rfloor}{N\Delta} \ , \\ \gamma_k &= f(k, \Delta\lambda, K) \end{aligned}$$

where N is the signal's length, $l \in [1, \ldots, N]$ its index, and Δ the sampling period. The mask $\boldsymbol{w}$ is defined as

$$\boldsymbol{w} = \Big(w_{1,1,1}, \quad \ldots, \quad w_{i,j,k}, \quad \ldots, \quad w_{I,J,K} \Big)^{\mathrm{T}} \ , \tag{3}$$

where the entries of the mask are computed by an extension of the 2-D mask given in Eqn. 1

$$w_{i,j,k} = \begin{cases} 1, & \text{if } \left| \frac{\gamma_k}{\gamma_k - \xi_i (L+D)} \right| > \frac{r_p}{L} \\ 0, & \text{otherwise} \end{cases} \ . \tag{4}$$

2.3. Efficient Cost Function to Estimate Detector Shifts

Given a reasonably large L, shifts in detector u- and v-direction can be used to approximate 3-D translations of moderate amplitude. Our image coordinate system is aligned with the plane of rotation for CT acquisition. In that case, 3-D motion in direction of the detector u-axis can be associated with detector u-shifts. In contrast, a 3-D translation towards and away from the detector makes the object or patient

appear bigger and smaller, which could be roughly approximated by an image scaling. However, the distance an object has to be moved orthogonal to the detector to introduce an error of about one pixel depends on the cone angle and is typically larger than the corresponding distance of motion parallel to the detector. Additionally, 2-D translations are applied directly in the 2-D Fourier domain of the projection images using the shift theorem. This reduces the computational cost as the 2-D-FFT of the projection images can be precomputed.

We obtain a parameter vector $\boldsymbol{\alpha} \in \mathbb{R}^{2K}$,

$$\boldsymbol{\alpha} = \begin{pmatrix} s_1, & t_1, & s_2, & t_2, & \dots, & s_K, & t_K \end{pmatrix}^{\mathrm{T}} , \tag{5}$$

where s_k and t_k are the detector shifts for projection index k in u- and v-direction, respectively. A single cost function evaluation requires only the 1-D Fourier transforms over the rotation angles, i.e., over index k. The complexity of the FFT thus reduces from $\mathcal{O}(M \log M)$ to $\mathcal{O}(M \log K)$. We can rewrite Eqn. 2 to

$$\hat{e}(\boldsymbol{\alpha}) = \left\| \boldsymbol{x}(\boldsymbol{\alpha}) \right\|_2^2 = \left\| \boldsymbol{W} \left(\boldsymbol{F}_\lambda \left(\boldsymbol{T}(\boldsymbol{\alpha}) \, \boldsymbol{F}_{uv} \boldsymbol{p} \right) \right) \right\|_2^2 , \tag{6}$$

where $\boldsymbol{F}_\lambda \in \mathbb{C}^{M \times M}$ and $\boldsymbol{F}_{uv} \in \mathbb{C}^{M \times M}$ denote the 1-D Fourier transforms over the angles and the 2-D Fourier transforms over the projection images, respectively. Further, $\boldsymbol{p} \in \mathbb{R}^M$ represents the original projection data and $\boldsymbol{T}(\boldsymbol{\alpha}) \in \mathbb{C}^{M \times M}$ is a diagonal matrix that holds the phase factors which encode the shifts in $\boldsymbol{\alpha}$. Assuming that $\boldsymbol{p}$ is ordered first over the rows u, then over the columns v and finally over the angles λ, $\boldsymbol{T}(\boldsymbol{\alpha})$ is given by

$$\begin{aligned} \boldsymbol{T}(\boldsymbol{\alpha}) = \operatorname{diag}\Big(\Big(& e^{-\imath 2\pi(\xi_1 s_1 + \psi_1 t_1)}, \quad \dots, \quad e^{-\imath 2\pi(\xi_I s_1 + \psi_J t_1)}, \\ & e^{-\imath 2\pi(\xi_1 s_2 + \psi_1 t_2)}, \quad \dots, \quad e^{-\imath 2\pi(\xi_I s_K + \psi_J t_K)} \Big)^{\mathrm{T}} \Big) . \end{aligned} \tag{7}$$

To allow for a gradient-based minimization of Eqn. 6 we need to compute the gradient with respect to the shifting parameter vector $\boldsymbol{\alpha}$. It is defined by

$$\nabla \hat{e}(\boldsymbol{\alpha}) = \left(\frac{\partial \hat{e}(\boldsymbol{\alpha})}{\partial s_1}, \quad \frac{\partial \hat{e}(\boldsymbol{\alpha})}{\partial t_1}, \quad \dots , \quad \frac{\partial \hat{e}(\boldsymbol{\alpha})}{\partial s_K}, \quad \frac{\partial \hat{e}(\boldsymbol{\alpha})}{\partial t_K} \right)^{\mathrm{T}} . \tag{8}$$

Hence, the computation of the gradient requires the partial derivatives of Eqn. 6 with respect to the individual shifting parameters. The partial derivatives are computed by

$$\begin{aligned} \frac{\partial \hat{e}(\boldsymbol{\alpha})}{\partial \alpha_l} &= \frac{\partial}{\partial \alpha_l} \left\| \boldsymbol{x}(\boldsymbol{\alpha}) \right\|_2^2 \\ &= \frac{\partial}{\partial \alpha_l} \left(\boldsymbol{x}(\boldsymbol{\alpha})^{\mathrm{H}} \boldsymbol{x}(\boldsymbol{\alpha}) \right) \\ &= 2 \, \boldsymbol{x}(\boldsymbol{\alpha})^{\mathrm{H}} \left(\frac{\partial}{\partial \alpha_l} \boldsymbol{x}(\boldsymbol{\alpha}) \right) \\ &= 2 \left(\boldsymbol{p}^{\mathrm{H}} \boldsymbol{F}_{uv}{}^{\mathrm{H}} \boldsymbol{T}(\boldsymbol{\alpha})^{\mathrm{H}} \boldsymbol{F}_\lambda{}^{\mathrm{H}} \boldsymbol{W}^{\mathrm{H}} \right) \boldsymbol{W} \boldsymbol{F}_\lambda \left(\frac{\partial}{\partial \alpha_l} \boldsymbol{T}(\boldsymbol{\alpha}) \right) \boldsymbol{F}_{uv} \boldsymbol{p} \\ &= 2 \, \boldsymbol{p}^{\mathrm{H}} \boldsymbol{F}_{uv}{}^{\mathrm{H}} \boldsymbol{T}(\boldsymbol{\alpha})^{\mathrm{H}} \boldsymbol{F}_\lambda{}^{\mathrm{H}} \boldsymbol{W} \boldsymbol{F}_\lambda \left(\frac{\partial}{\partial \alpha_l} \boldsymbol{T}(\boldsymbol{\alpha}) \right) \boldsymbol{F}_{uv} \boldsymbol{p} , \end{aligned} \tag{9}$$

where α_l is an element of vector $\boldsymbol{\alpha}$ and corresponds to either s_l or t_l, i.e. a shift of the l-th projection image in u- or in v-direction, respectively. Further, $\square^{\mathrm{H}}$ denotes the transposed and complex conjugated of a vector or matrix. Due to the diagonal and binary structure of $\boldsymbol{W}$ it holds that $\boldsymbol{W}^{\mathrm{H}} = \boldsymbol{W}$ and that $\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W} = \boldsymbol{W}$ which was used in the last step of Eqn. 9. Note, that $\boldsymbol{x}(\boldsymbol{\alpha})$ has already been computed in the cost function evaluation and does not need to be recomputed for the evaluation of the gradient. The partial derivatives of $\boldsymbol{T}(\boldsymbol{\alpha})$ are given by

$$\frac{\partial}{\partial \alpha_l}\boldsymbol{T}(\boldsymbol{\alpha}) = \operatorname{diag}(\boldsymbol{s})\ \boldsymbol{T}(\boldsymbol{\alpha}) \tag{10}$$

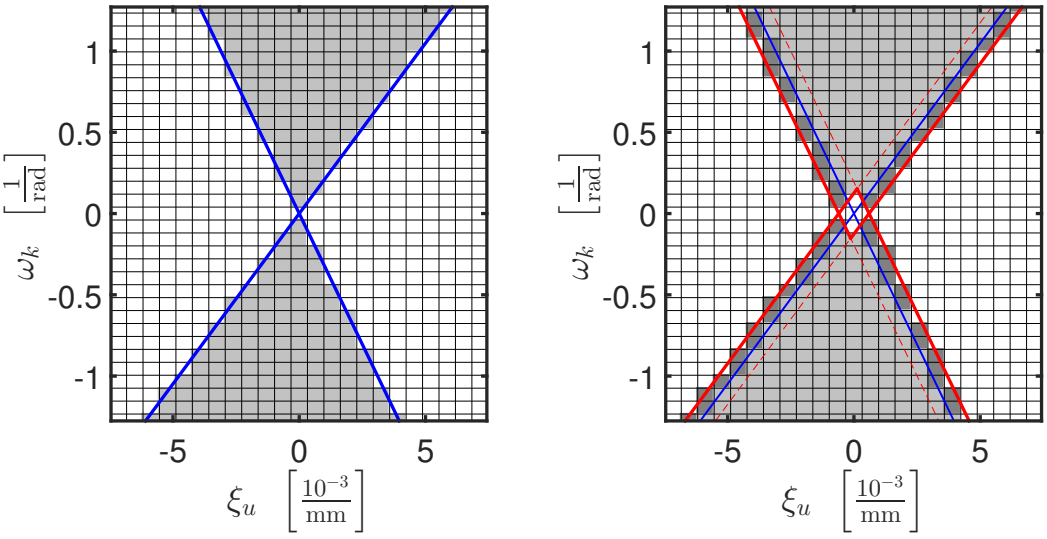
$$\boldsymbol{s} = \begin{cases} (0, \ldots, 0, -\imath 2\pi\xi_1, \ldots, -\imath 2\pi\xi_I, 0, \ldots, 0)^{\mathrm{T}} & \text{if } \alpha_l = s_l \\ (0, \ldots, 0, -\imath 2\pi\psi_1, \ldots, -\imath 2\pi\psi_J, 0, \ldots, 0)^{\mathrm{T}} & \text{if } \alpha_l = t_l \end{cases} \tag{11}$$

where $\boldsymbol{s} \in \mathbb{C}^M$ holds the partial derivatives of the arguments of the exponential functions in $\boldsymbol{T}(\boldsymbol{\alpha})$. Note, that a change in the parameters s_l or t_l only effects the 2-D Fourier transform of the l-th projection image. Therefore, $\boldsymbol{s}$ is only non-zero at elements in the range of $[(l-1)IJ+1, (l)IJ]$. Note also, that the multiplication in Fourier domain by $\operatorname{diag}(\boldsymbol{s})$, corresponds to a spatial derivative of the l-th projection image in u- or v-direction.

2.4. Discretization and Adjustment of the Mask

The implementation of the mask $\boldsymbol{w}$ according to Eqn. 4 represents a discretization of the continuous function given in Eqn. 1. Discretization includes the intersection of the continuous straight lines at non-integer positions with the discretized DFT grid, making the determination of the mask's boundaries challenging. This process is illustrated in Figure 2a, where we can see the sampling grid as well as the discretized mask along with the continuous straight lines defined by Eqn. 1. We can see that some pixels of the mask extend over the boundary defined by the straight lines, whereas others barely touch the boundary. The implications of this inaccuracy are two-fold. First, a mask that is too large prevents the motion-induced spectral energy to be redistributed close to the support boundary. Thus, we implicitly assume that these frequencies are zero even for ideal motion-free data which may not be entirely true. Second, if the mask is too narrow, energy may be shifted to spectral bins that are close to the boundary but not accurately covered by the mask. Hence, these bins have none or only little influence on the cost function and are therefore not important for the actual consistency. During our experiments we have noticed that the latter, i.e., a too narrow mask is less robust for motion estimation leading to a low-frequency sinusoidal signal with high amplitude which interferes with the actual motion signal. Please see the discussion in Section 5 for more details.

As a result we present an adjustment of Eqn. 4 such that the mask region slightly extends over the lines. The outcome can be seen in Figure 2b, where the adjusted



(a) Direct discretization (b) Adjusted mask

Figure 2: (a) Direct discretization of triangular regions, (b) Heuristic adjustment of mask to ensure that all motion related energies are covered.

mask is again visualized as gray area. The mask now extends over the solid blue lines that indicate the original boundary. As new discretization boundaries we define four shifted lines (solid, red lines) parallel to the original triangle boundary (dashed, red lines). These lines form two overlapping triangles which define our new mask region as depicted in Figure 2b. To ensure that the zero-frequencies are not contained in the adjusted mask, we exclude frequencies that are covered by both triangles. The adjustment can be easily done by modifying the distance ϵ of ideal triangle boundaries to the shifted, parallel lines.

2.5. Regularization, Optimization and Motion Compensation

By definition the FCC cost function in Eqn. 6 is invariant to shifts in u- or v-direction which are constant over the projection images, e.g. a global detector offset. Let $\varphi_{i,j,n}$ be the complex Fourier coefficient of the 2-D Fourier transform of the n-th projection image along u- and v-direction and c_1 and c_2 are constant shift values, respectively. Further, $x_{i,j,k}$ is the 3-D Fourier coefficient of frequency bin (i, j, k) computed by

$$x_{i,j,k}(\boldsymbol{\alpha}) = w_{i,j,k} \sum_{n=1}^{K} \varphi_{i,j,n}\, \mathrm{e}^{-\imath 2\pi(\xi_i s_n + \psi_j t_n)}\, \mathrm{e}^{-\imath 2\pi k n/K} \quad . \tag{12}$$

If we set $s_k = c_1$ and $t_k = c_2$, we can factorize the phase factor out of the Fourier transform's summation

$$\begin{aligned} x_{i,j,k} &= w_{i,j,k}\, \mathrm{e}^{-\imath 2\pi(\xi_i c_1 + \psi_j c_2)} \sum_{n=1}^{K} \varphi_{i,j,n}\, \mathrm{e}^{-\imath 2\pi k n/K} \\ x_{i,j,k} &= w_{i,j,k}\, \mathrm{e}^{-\imath 2\pi(\xi_i c_1 + \psi_j c_2)}\ \Phi_{i,j,k} \end{aligned} \quad , \tag{13}$$

where $\Phi_{i,j,k}$ is the 3-D Fourier coefficient of the shift-free projection image stack at frequency bin (i, j, k). The FCC theorem considers only the absolute value of the spectrum which is why the cost function in Eqn. 6 uses the complex, squared L2-norm. This is equal to the sum of element-wise complex conjugated multiplications

$$\begin{aligned} \hat{e}(\boldsymbol{\alpha}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} w_{i,j,k}\, x_{i,j,k}\, \overline{x_{i,j,k}} \\ \hat{e}(\boldsymbol{\alpha}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} w_{i,j,k}\, \Phi_{i,j,k}\, \overline{\Phi_{i,j,k}}\, \underbrace{\mathrm{e}^{\imath 2\pi(\xi_i c_1 + \psi_j c_2)}\, \mathrm{e}^{-\imath 2\pi(\xi_i c_1 + \psi_j c_2)}}_{=1} \end{aligned} \quad , \tag{14}$$

showing that the cost function is independent of any constant shifts c_1 or c_2.

While a constant offset in v-direction corresponds only to an object shift in z-direction (not considering cone-beam artifacts), a constant shift in u-direction leads to distortions in the reconstruction. To account for this limitation and to ensure stability of the algorithm we decided to provide our method with constant motion parameters for the first projection image, i.e., $s_1 = \widetilde{s}$ and $t_1 = \widetilde{t}$.

We incorporate $\widetilde{s}$ and $\widetilde{t}$ as regularizer that extends the cost function in Eqn. 6. Our final cost function including the regularization is given by

$$\widetilde{e}(\boldsymbol{\alpha}) = \hat{e}(\boldsymbol{\alpha}) + \kappa\, \frac{1}{2} \left\| \begin{pmatrix} s_1 - \widetilde{s} \\ t_1 - \widetilde{t} \end{pmatrix} \right\|_2^2 \quad . \tag{15}$$

As the regularizer is incorporated with the Lagrangian multiplier κ, the gradient can be computed in a straightforward manner yielding

$$\nabla\widetilde{e}(\boldsymbol{\alpha}) = \nabla\hat{e}(\boldsymbol{\alpha}) + \kappa \begin{pmatrix} (s_1 - \widetilde{s}) & (t_1 - \widetilde{t}) & 0 & \cdots & 0 \end{pmatrix}^{\mathrm{T}} \quad . \tag{16}$$

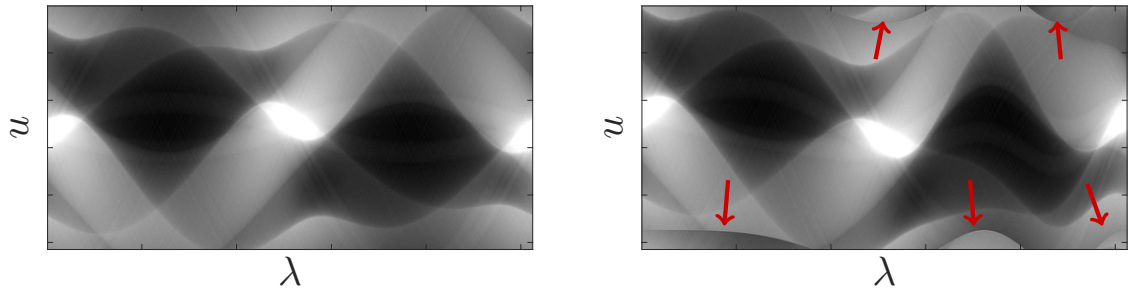
We use an unconstrained gradient-based optimizer for the minimization of Eqn. 15. The gradient is computed analytically following Eqn. 16 whilst the inverse Hessian is approximated from the gradient using a BFGS scheme. To determine the step direction we attempt a Newton step based on the inverted Hessian. Finally, the step size is given by a line-search method. Further information about the optimizer may be obtained from Maier and Berger (2010); Verrill (2005).

The final motion estimates are directly incorporated into the projection matrices to compensate motion during the backprojection step. The function $p(\boldsymbol{p}_k, \boldsymbol{\alpha})$ is used to update the original, motion corrupted projection matrices $\boldsymbol{p}_k \in \mathbb{R}^{3\times 4}$ according to

$$p(\boldsymbol{p}_k, \boldsymbol{\alpha}) = \begin{pmatrix} 1 & 0 & -\frac{s_k}{\Delta u} \\ 0 & 1 & -\frac{t_k}{\Delta v} \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{p}_k \quad . \tag{17}$$

2.6. Axial Truncation: Problems and Solutions

FCC are not theoretically defined for truncated projection data. For the proposed method the truncation problem becomes visible after application of detector



(a) Truncated fan-beam sinogram (b) Sinogram after Fourier-based shifting

Figure 3: (a) Truncated fan-beam sinogram. (b) Reappearance of sinogram at its boundaries after applying translations in Fourier domain. The erroneously introduced intensity variations over λ may violate FCC.

translations. A detector translation using phase factors, causes image content at the detector boundary to reappear at the opposite side. This is shown in in Figures 3a and 3b, using a truncated fan-beam sinogram and its shifted version using a simulated detector translation. Application of the translation in spatial domain would also not solve the problem as the periphery is still unknown. We propose a practical extension to FCC, that allow their application to truncated data. In principle the concept would also be suitable for lateral truncation, however, within the scope of this work we focus on axial truncation only. Thus, the method becomes applicable to more realistic data, where axial truncation is usually unavoidable.

To ensure that the projection image is of limited extent, we multiply each image with an apodization function that smoothly fades out to zero in horizontal detector direction. This weighting emulates a maximum object size along the scanner's rotation axis. However, apodization has to be performed after application of estimated motion parameters in order to avoid intensity variations due to the windowing function itself. For motion estimation we use a Blackman-Harris apodization function as depicted in Fig. 4. The function is then multiplied with each column of all projection images, after applying the estimated translations. However, instead of a regular Blackman-Harris window, that covers the full detector height, we limit the range by t_{max} at its borders. As a result, any translation $t_k \in [-t_{max}, t_{max}]$ will not cause a periodic reappearance at bottom or top of the detector.

Apodization is easily integrated into the cost functions in Equations 2 and 6. Let $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{a}) \in \mathbb{R}^{IJ \times IJ}$ be the diagonal matrix which holds all weights $\boldsymbol{a} \in \mathbb{R}^{IJ}$ of the apodization function. Following Harris (1978) it is defined by,

$$a_{i,j,k} = \begin{cases} 0 & \text{if } j \notin \left[\lceil \frac{t_{max}}{\Delta v} \rceil + 1, J - \lceil \frac{t_{max}}{\Delta v} \rceil\right] \\ \sum_{l=0}^{3} c_l \cdot \cos\left(\frac{l\,(2\pi j')}{J'}\right) & \text{otherwise} \end{cases} \tag{18}$$

with adjusted index $j' = j - \lceil t_{max}/\Delta v \rceil - 1$ and reduced window width $J' = J - 2\lceil t_{max}/\Delta v \rceil$, where $\lceil \cdot \rceil$ represents rounding to the next highest integer. The coefficients

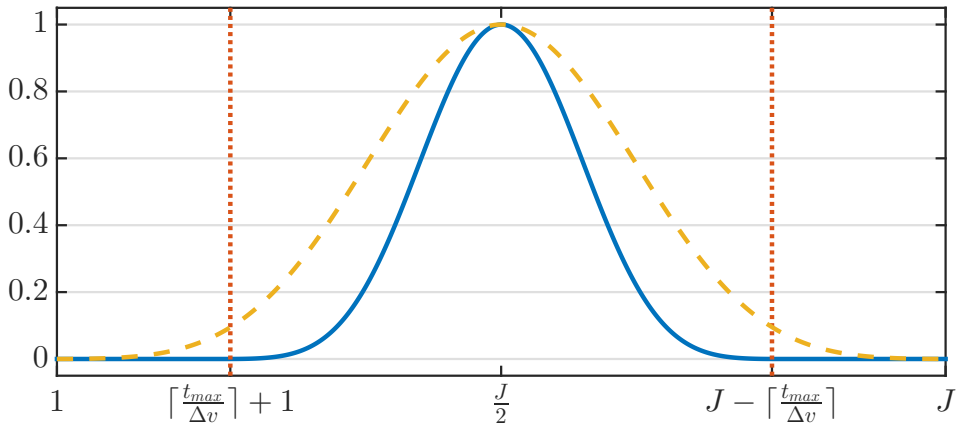


Figure 4: Adjusted window function to allow for translations $t_k \in [-t_{max}, t_{max}]$ on truncated data . Yellow, dashed line: Original Blackman-Harris function with a width of $j \in [1, J]$. Blue, solid line: Adjusted function ranging from $j \in \left[\lceil \frac{t_{max}}{\Delta v} \rceil + 1, J - \lceil \frac{t_{max}}{\Delta v} \rceil\right]$.

are given by $c_0 = 0.35875, c_1 = -0.48829, c_2 = 0.14128, c_3 = -0.01168$.

The cost functions in Eqn. 2 is now adjusted to

$$e(\boldsymbol{\alpha}) = \left\| \boldsymbol{W} \left(\boldsymbol{F}\, \boldsymbol{A}\, \boldsymbol{p}(\boldsymbol{\alpha}) \right) \right\|_2^2 . \tag{19}$$

However, incorporating $\boldsymbol{A}$ into the efficient cost function described by Eqn. 6, requires an additional inverse (i.e., $\boldsymbol{F}_{uv}^{\mathrm{H}}$) and forward (i.e., $\boldsymbol{F}_{uv}$) 2-D FFT, yielding

$$\hat{e}(\boldsymbol{\alpha}) = \frac{1}{2} \left\| \boldsymbol{W} \left(\boldsymbol{F}_{\lambda} \underbrace{\left(\boldsymbol{F}_{uv} \boldsymbol{A} \boldsymbol{F}_{uv}^{\mathrm{H}} \right)}_{\text{apodization}} \left(\boldsymbol{T}(\boldsymbol{\alpha})\, \boldsymbol{F}_{uv} \boldsymbol{p} \right) \right) \right\|_2^2 . \tag{20}$$

Consequently, the computational benefit of this description would vanish. Yet we can exploit the convolution theorem and apply the window by convolution with its Fourier transform. It is particularly beneficial that the Blackman-Harris window, has only a small number of nonzero Fourier coefficients (Harris; 1978). For $t_{max} = 0$ it holds that only four unique spectral coefficients are present, requiring seven multiplications and six summations at each Fourier coefficient to perform apodization. This is no longer true for $t_{max} > 0$, as the window's spectrum is no longer sparse. However, in practice using only a small number of most significant coefficients, reduces complexity and still leads to a good approximation of the actual window. In practice, we chose to retain eight unique spectral coefficients, which led already to a very close approximation to the actual window function. Note, that the cost function's gradient in Eqn. 9 is still valid. Yet, it now also includes an apodization step, thus, obtaining partial derivatives also requires described convolution in spectral domain.

An unwanted side effect of apodization is that the overall energy of the spectrum may not remain constant. Hence, energy in the triangular regions can also be reduced by translations that move high intensities such that they observe a lower weight

during apodization. To reduce this effect, we propose regularization with the motion parameters' squared L2 norm. The updated cost function is given by

$$e'(\boldsymbol{p}) = \widetilde{e}(\boldsymbol{p}) + \kappa' \left(\frac{1}{2IJK} \boldsymbol{p}^{\mathrm{T}} \boldsymbol{p} \right) \quad . \tag{21}$$

Consequently, the overall gradient of the cost function may be computed by

$$\nabla e'(\boldsymbol{p}) = \nabla \widetilde{e}(\boldsymbol{p}) + \kappa' \left(\frac{1}{IJK} \boldsymbol{p} \right) \quad . \tag{22}$$

Similar to Eqn. 15 an additional Lagrangian multiplier κ' is used to create a single combined cost function for optimization.

2.7. Implementation Details

The runtime of a single cost function and gradient evaluation is crucial for the practical applicability of our algorithm. First tests based on a single-threaded CPU implementation revealed a high runtime of several minutes for only a single evaluation of the cost function due to multiple iterations over the projection volume and the 3-D Fourier transform. We save time by exploiting the shift theorem and use Eqn. 6 so that $\boldsymbol{F}_{uv}\boldsymbol{p}$ can be precomputed prior to the optimization. To further speed up the recurring 1-D-Fourier transform over the angular direction, we also swapped dimensions such that the projection angles are stored sequentially in memory. Further, we decided to exploit the parallelizable nature of the problem and implemented all necessary computations on the GPU using OpenCL.

2.7.1. Cost Function Evaluation For large data (in our case $640 \times 480 \times 512 \times 2 \times 4\,\mathrm{Byte}/1024^2 \approx 1200\,\mathrm{MB}$), it may be infeasible to store multiple copies of the projection stack in GPU memory. Therefore, we only keep a single copy of the data on the GPU, representing the shifted projection images. The parameter vector for the n-th iteration, i.e. $\boldsymbol{\alpha}^n$ is given by the optimizer and builds upon the original motion-corrupted data. As a first step we build a relative parameter vector $\Delta\boldsymbol{\alpha}^n = \boldsymbol{\alpha}^n - \boldsymbol{\alpha}^{n-1}$, where $\boldsymbol{\alpha}^{n-1}$ denotes the parameter vector of the previous iteration. Instead of applying $\boldsymbol{\alpha}^n$, we use the relative shifts in $\Delta\boldsymbol{\alpha}^n$ and apply them directly to the already shifted projection volume.

The algorithm for a single cost function evaluation consists of the following steps:

(i) Compute relative parameter vector: $\Delta\boldsymbol{\alpha}^n = \boldsymbol{\alpha}^n - \boldsymbol{\alpha}^{n-1}$

(ii) Apply shifts: Performs the complex multiplication with the phase factors derived from $\Delta\boldsymbol{\alpha}^n$. We precompute two vectors holding all ψ and ξ values. In every iteration, only the relative parameter vector $\Delta\boldsymbol{\alpha}^n$ needs to be transferred to the GPU.

(iii) Angular 1-D FFTs: We use a GPU-based OpenCL library for the 1-D FFT along the projection angles. This implementation does not support memory efficient in-place transforms, which is why we store the 3-D FFT result separately.

(iv) L2-norm and mask: The sum of all magnitudes in the mask region is built using a GPU-based parallel reduction. Note, that only the summation result needs to be transferred back to the host. The mask indicating the frequency-region we want to minimize is precomputed according to Eqn. 3, hence, only the final result of the summation needs to be transferred back to the host.

2.7.2. Gradient Computation Eqns. 8 and 9 state that the gradient vector requires computation of $\boldsymbol{x}(\boldsymbol{\alpha})\frac{\partial}{\partial\alpha_l}\boldsymbol{x}(\boldsymbol{\alpha})$ for each of the $2K$ partial derivatives. $\boldsymbol{x}(\boldsymbol{\alpha})$ is already computed, as the gradient computation is executed after the cost function evaluation. $\frac{\partial}{\partial\alpha_l}\boldsymbol{x}(\boldsymbol{\alpha})$ consists of a multiplication of the precomputed 2-D Fourier transforms $\boldsymbol{F}_{uv}\,\boldsymbol{p}$, with the partial derivatives of the shifting matrix $\frac{\partial}{\partial\alpha_l}\boldsymbol{T}(\boldsymbol{\alpha})$, followed by the angular Fourier transform $\boldsymbol{F}_\lambda$. Thus, naive implementation requires one Fourier transform per partial derivative. However, a mathematical trick allows to avoid computation of the angular FFTs entirely.

We exploit the fact that $\frac{\partial}{\partial\alpha_l}\boldsymbol{T}(\boldsymbol{\alpha})$ will only be non-zero for values corresponding to the l-th projection image. The same holds for $\frac{\partial}{\partial\alpha_l}\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{F}_{uv}\,\boldsymbol{p}$. Using the notation in Section 2.5, we define the angular FFT of the partial derivative as

$$\begin{aligned}\tfrac{\partial}{\partial\alpha_l}x_{i,j,k}(\boldsymbol{\alpha}) &= w_{i,j,k}\sum_{n=1}^{K}\varphi_{i,j,n}\underbrace{\left(\frac{\partial}{\partial\alpha_l}\mathrm{e}^{-\imath 2\pi(\xi_i s_n+\psi_j t_n)}\right)}_{=0\ \forall\, n\neq l}\mathrm{e}^{-\imath 2\pi kn/K}\ ,\\ \tfrac{\partial}{\partial\alpha_l}x_{i,j,k}(\boldsymbol{\alpha}) &= w_{i,j,k}\,\varphi_{i,j,l}\left(\tfrac{\partial}{\partial\alpha_l}\mathrm{e}^{-\imath 2\pi(\xi_i s_l+\psi_j t_l)}\right)e^{-\imath 2\pi kl/K}\ .\end{aligned} \tag{23}$$

As a result, the angular Fourier transform breaks down to a simple multiplication by $e^{-\imath 2\pi kl/K}$ which drastically reduces the computational complexity.

Further, we make sure that the partial derivatives $\frac{\partial}{\partial s_l}\hat{e}(\boldsymbol{\alpha})$ and $\frac{\partial}{\partial t_l}\hat{e}(\boldsymbol{\alpha})$ are computed within the same GPU kernel call. As both values are based on the same data, and data transfer is the dominant time factor, this step almost halves the overall runtime.

3. Simulation and Evaluation Protocol

To evaluate our method, we perform a numerical phantom study using the 3-D FORBILD head phantom (Lauritsch and Bruder; 2001). Overall, we created five projection image stacks, where three of them suffer from high-frequency object motion and two contain only low-frequency motion. We also rendered motion-free projections to obtain ground-truth images and a reference reconstruction.

All projection images have been created using a monochromatic absorption model with a photon energy of 80 keV. The rendering was performed as described by Maier et al. (2012). It uses analytic descriptions of the phantoms which allows for a direct computation of intersection lengths. We performed noise-free and noisy simulations using a Poisson distributed noise model. The number of photons emitted from the X-ray source was limited to 5000 per detector pixel which resulted in a high noise level.

A small number of rays had large intersection lengths with bone material, so that we encountered photon starvation. Those rays produced infinite line integral values which were subsequently limited to the maximum value of the pixels that did not suffer from photon starvation.

One application field we aim for with the development of FCC-based motion estimation, is knee imaging under weight-bearing conditions. Therefore, we also evaluate the proposed method on a simulated but highly realistic dataset of two knees under load. The data is generated with the 4-D XCAT model, which is based on segmentations of real anatomies (Segars et al.; 2010). Moreover, the simulated motion is taken from a real subject by incorporating motion recordings of subjects holding a squat for 20 s using an optical camera tracking system. For details on the acquisition process and registration of the recorded to the XCAT motion we refer to (Choi et al.; 2013).

In contrast to the FORBILD phantom, axial truncation is inevitable when imaging knees. Yet, the datasets do not suffer from lateral truncation as the detector spacing was specifically increased to allow full coverage of both legs. Note, that the FCC method is not yet applicable to real knee acquisitions, as that would require full scans and possibly also truncation-free acquisitions in lateral direction. Nevertheless, we consider this experiment as a first step towards an application in real scenarios. We are specifically interested, if the proposed extension based on dynamic apodization of projection images in detector v-direction can account for axial truncation.

In addition to weight-bearing imaging of the knee, we recently proposed the usage of FCC for respiratory motion estimation in rotational angiography of coronary arteries (Unberath et al.; 2017). In this field, the respiratory motion of the heart is often approximated by only vertical detector translations. Thus, we restricted the optimization to only vertical translations by removing s_k from the parameter vector. A synthetic rotational angiography dataset, also based on the XCAT phantom, was used for projection generation. The background structures, e.g., soft tissue and rips, were removed by a novel approach for virtual, digital subtraction angiography (DSA) to achieve truncation-free projection images, depicting contrast filled coronary arteries only (Unberath et al.; 2016). For further information on data acquisition and the setup of the experiment, we refer to (Unberath et al.; 2017). Note, that this study did not include subsequent image reconstruction, which is still challenging due to the heart's contraction.

Table 1 shows the geometry and reconstruction parameters for the simulations. We estimate the object extent r_p directly from the motion-corrupted data. We estimate r_p by using the motion that shows the largest displacement in horizontal detector direction. This way, r_p will be rather overestimated, causing smaller triangular regions, which prevents the mask from extending into spectral energies originating from the scanned object. First, we build the sum over dimension J and dimension K, resulting in a vector of size I. Then, we manually identify the distances of the left and right zero-edge to the center, i.e. $\tilde{u}_l$ and $\tilde{u}_r$, respectively. The maximum object extent is estimated by fitting

Parameter	Symbol	Unit	Ideal / Noisy	XCAT Knee
SID	L	mm	600	780
DID	D	mm	600	418
Object size	r_p	mm	125.0	112.3
Mask adj.	ϵ	-	3×10^{-3}	1×10^{-3}
Detector Size	$I \times J$	-	640×480	620×480
Pixel size	$\Delta u \times \Delta v$	mm^2	1.2×1.2	1.2×1.2
#Projections	K	-	512	256
Angular spacing	$\Delta\lambda$	deg	0.703	1.406
Reconstr. size	$R_x \times R_y \times R_z$	-	$512 \times 512 \times 512$	$512 \times 512 \times 256$
Isotr. voxel size	-	mm	0.5	0.5
#Photons per pixel	-	-	- / 5000	-
Absorption model	-	-	monochromatic	
Photon Energy	-	keV	80	

Table 1: Parameters used for FORBILD and XCAT knee cases. For the rotational angiography dataset we refer to (Unberath et al.; 2017).

a circle to the backprojection of $\max(\tilde{u}_l, \tilde{u}_r)$. This can be formalized by

$$\hat{r}_p = \max(\tilde{u}_l, \tilde{u}_r) \frac{L}{L+D} \tag{24}$$

$$r_p = \hat{r}_p \frac{L}{\sqrt{L^2 + (\hat{r}_p)^2}} \ , \tag{25}$$

where Eqn. 24 corrects for magnification and Eqn. 25 for the fan-angle.

As described in Section 2.5, constants $\tilde{s}$ and $\tilde{t}$ have to be selected for the shift values of the first projection. We decided to set these values to the 2-D ground-truth motion to allow for an accurate quantitative comparison. Further implications on determining $\tilde{s}$ and $\tilde{t}$ are discussed in Section 5. We normalized the data term $\hat{e}(\boldsymbol{\alpha})$ such that it evaluates to 100 in the first iteration and the Lagrangian multiplier of the regularizer was set to $\kappa = 1$. Finally, the gradient-based optimizer was restricted to a maximum of 10000 iterations, yet, this limit was never reached in practice (see Section 4.6).

For the XCAT knee dataset the Lagrangian for L2 regularization κ' as well as ϵ were heuristically determined, whereas other parameters remained constant and identical to the FORBILD cases. We did not have access to ground-truth detector translations. Thus, we decided to remove the regularizer presented in Eqn. 15 by setting its Lagrangian multiplier $\kappa = 0$. However, a constant bias in motion parameters is also prevented by usage of the L2 regularizer.

Respiratory motion estimation on coronary artery data generally followed the same pipeline as was used for the FORBILD cases. Apodization was not necessary as truncation could already be removed using a virtual DSA approach.

3.1. Simulation of Translational 3-D Motion for FORBILD Phantom

In addition to a motion-free case, we simulated three high-frequency translational motion patterns and two low-frequency cases. When selecting the motion patterns we made sure that their visual effect to the triangular regions is widespread, as can be obtained from the differences of the uncorrected spectra in Figure 8 and Figure 9. We use concepts of projective geometry to efficiently describe the imaging geometry of a cone-beam CT, i.e., the mapping of 3-D world to 2-D detector coordinates, using projection matrices and homogeneous coordinates (Hartley and Zisserman; 2004). Thus, the motion can be easily incorporated by a right-multiplication with the projection matrices

$$\boldsymbol{P}_k = \hat{\boldsymbol{P}}_k \begin{pmatrix} 1 & 0 & 0 & \Delta x_k \\ 0 & 1 & 0 & \Delta y_k \\ 0 & 0 & 1 & \Delta z_k \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \forall k \in [1, \cdots, K] \ , \tag{26}$$

where $\hat{\boldsymbol{P}}_k \in \mathbb{R}^{3\times4}$ are the original projection matrices of the circular trajectory. The matrices $\boldsymbol{P}_k$ are then used for the creation of the motion-corrupted projection images. Note, that Eqns. 27 to 31 are normalized such that they take $\tau_k = \frac{k-1}{K-1} \in [0, 1]$ as a temporal parameter, which would effectively equal a total scan time of 1 s. Thus, motion frequencies reported below have to be interpreted in perspective and would have to be divided by a typically longer scan time. Fig. 5 provides an overview of all simulated 3-D translational motion patterns.

Oscillating Motion (Oscil) Our first motion type represents an accelerated, oscillating translation along the x-, y- and z-direction. It is a straightforward 3-D extension of the 2-D motion used in our fan-beam study (Berger et al.; 2014). This motion has a constant cosine oscillation with zero-mean. Low-frequency contributions are small due to a moderate setting for the acceleration factor. The translations are computed as

$$\Delta x_k = \Delta y_k = \Delta z_k = m_{\text{oscil}}(a, b, f) = a\Big(\frac{2}{1 + \exp(b\cos(2\pi f \tau_k))} - 1\Big) \ , \tag{27}$$

where a controls the amplitude, b the acceleration and f is the frequency. We set $a = 3\,\text{mm}$, $b = 4$ and $f = 16\,\text{Hz}$. Note, that the amplitude is given by $a(\exp(b) - 1)/(\exp(b) + 1) \approx 2.89\,\text{mm}$.

Linear Increase in Frequency (Chirp) This motion describes an oscillating cosine motion with the same amplitude along the x-, y- and z- direction. It has a zero-mean and only a moderate amplitude but the frequency increases linearly between $[0, p]$ over the acquisition time. We use this type of translation characterize the behavior of our method for a combination of low- and high-frequency motion. It is define by

$$\Delta x_k = \Delta y_k = \Delta z_k = a\cos(2\pi f_k \tau_k) \ , \tag{28}$$

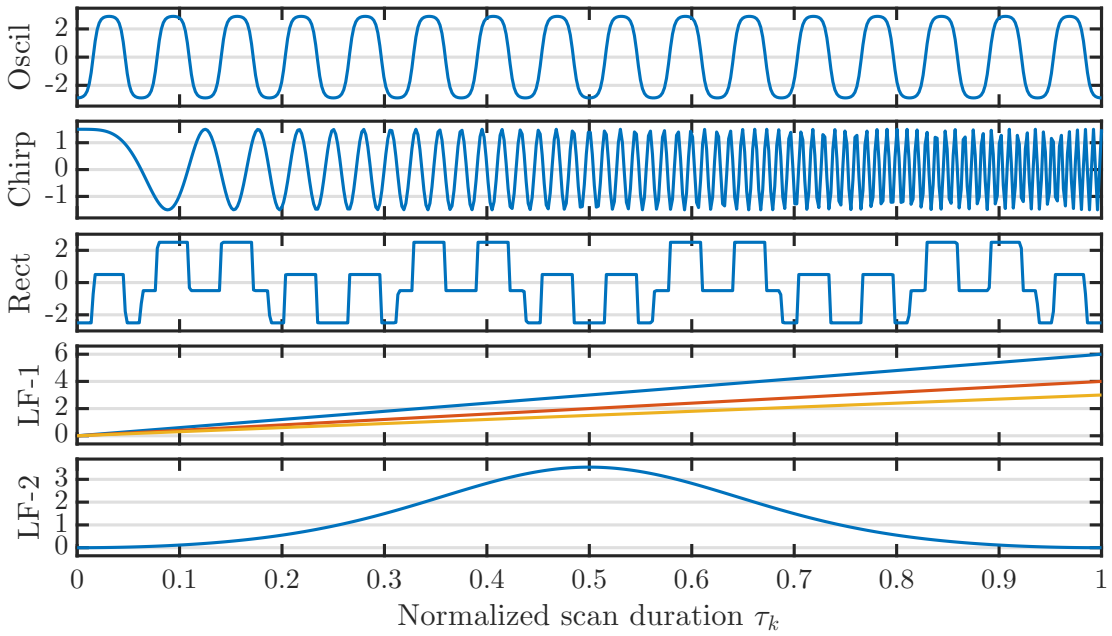


Figure 5: Plot of all simulated 3-D translations. Except for *LF-1*, all translations have been simultaneously applied to x-, y- and z-direction, to ensure that the object motion covers all possible directions.

where we set the amplitude to $a = 1.5\,\text{mm}$, the frequency to $f_k = (p\,\tau_k)$ and the frequency pitch to $p = 64\,\text{s}^{-2}$.

Rectangular Steps (Rect) With this motion, we model sharp transitions of motion states by two overlapping rectangular functions of different amplitude and frequency. The rectangular function is emulated by the *Oscil* function in Eqn. 27 with an extreme setting for the acceleration factor b, i.e.,

$$\Delta x_k = \Delta y_k = \Delta z_k = m_{\text{oscil}}(a_1, b, f_1) + m_{\text{oscil}}(a_2, b, f_2) \ , \tag{29}$$

where a_1 and a_2 as well as f_1 and f_2 correspond to the amplitude and frequency of the first and second rectangular signal, respectively. We set $a_1 = 1.5\,\text{mm}$, $f_1 = 16\,\text{Hz}$, $a_2 = 1.0\,\text{mm}$, $f_2 = 4\,\text{Hz}$ and $b = 128$.

Low-Frequency Motion 1 (LF-1) We adopted this motion from Yu et al. (2006) who used the x- and y-axis translation for their evaluation using fan-beam CC. The object moves linearly over time in along x-, y- and z-axis with three different slopes. The motion is therefore low-frequent and has a nonzero-mean. The starting and end position are not identical, leading to a sharp transition in the periodic extension over the projection angles. This leads to strong streaking artifacts and influences the angular Fourier transform as it assumes a periodic extension. The motion is defined by

$$\Delta x_k = q_1\tau_k, \ \ \Delta y_k = q_2\tau_k, \ \ \Delta z_k = q_3\tau_k \ , \tag{30}$$

where q_1, q_2 and q_3 are the slopes for x-, y- and z-axis which we set to $q_1 = 6\,\mathrm{mm\,s^{-1}}$, $q_2 = 4\,\mathrm{mm\,s^{-1}}$ and $q_3 = 3\,\mathrm{mm\,s^{-1}}$.

Low-Frequency Motion 2 (LF-2) In contrast to *LF-1*, this low-frequency motion has the same start and end position. It performs a single forth-and-back transition along x-, y- and z-axis. The reconstruction artifacts are limited to motion blur and low-frequency intensity gradients as no sharp transitions exist in the angular direction of the projection stack. The motion is given by

$$\Delta x_k = \Delta y_k = \Delta z_k = a \left(\frac{\exp(-\cos(2\pi f \tau_k) + 1) - 1}{\exp(2) - 1} \right) \quad , \tag{31}$$

where we set the amplitude to $a = \sqrt{\frac{25}{2}}\mathrm{mm}$ and the frequency to $f = 1\,\mathrm{Hz}$.

3.2. Evaluation Protocol of Estimated 2-D Motion Parameters

We perform a qualitative comparison of the estimated motion parameters using plot overlays. Additionally, we compute the mean-absolute-distance (MAD) and its standard deviation (SD) between the estimated and the ground-truth shifting parameters. The measures are calculated separately for each detector direction for all cases, including motion-free, noise-free and noisy data. They are defined by

$$\begin{aligned} \mathrm{MAD}(\boldsymbol{\alpha}) &= \tfrac{1}{K} \sum_{k=1}^{K} \begin{pmatrix} |s_k - g_{k_1}| \\ |t_k - g_{k_2}| \end{pmatrix} \\ \mathrm{SD}(\boldsymbol{\alpha}) &= \tfrac{1}{K-1} \sum_{k=1}^{K} \begin{pmatrix} (|s_k - g_{k_1}| - \mathrm{MAD}(\boldsymbol{\alpha})_1)^2 \\ (|t_k - g_{k_2}| - \mathrm{MAD}(\boldsymbol{\alpha})_2)^2 \end{pmatrix} \end{aligned} \quad . \tag{32}$$

where $\boldsymbol{g}_k = \begin{pmatrix} g_{1_k} & g_{2_k} \end{pmatrix}^{\mathrm{T}}$ denotes the 2-D ground-truth motion. We compute $\boldsymbol{g}_k$ by projecting the coordinate center as well as the 3-D translations $\begin{pmatrix} \Delta x_k & \Delta y_k & \Delta z_k \end{pmatrix}^{\mathrm{T}}$ to 2-D detector coordinates. Due to the structure of the projection matrices, the results are directly given by the last columns of $\hat{\boldsymbol{P}}_k$ and $\boldsymbol{P}_k$, which we call $\hat{\boldsymbol{c}}_k = \begin{pmatrix} \hat{c}_{k,1} & \hat{c}_{k,2} & \hat{c}_{k,3} \end{pmatrix}^{\mathrm{T}}$ and $\boldsymbol{c}_k = \begin{pmatrix} c_{k,1} & c_{k,2} & c_{k,3} \end{pmatrix}^{\mathrm{T}}$, respectively. The projection to 2-D is then given as division by the third component. As shown in Eqn. 33, $\boldsymbol{g}_k$ can be computed by the difference of the projected 2-D points, scaled by the detector pixel sizes.

$$\boldsymbol{g}_k = \begin{pmatrix} \Delta u & 0 \\ 0 & \Delta v \end{pmatrix} \left(\begin{pmatrix} c_{k,1}/c_{k,3} \\ c_{k,2}/c_{k,3} \end{pmatrix} - \begin{pmatrix} \hat{c}_{k,1}/\hat{c}_{k,3} \\ \hat{c}_{k,2}/\hat{c}_{k,3} \end{pmatrix} \right) \tag{33}$$

3.2.1. Image-Based Evaluation Protocol We performed three different reconstructions for all cases using the reconstruction settings mentioned in Table 1. The standard Feldkamp-Davis-Kress (FDK) algorithm (Feldkamp et al.; 1984) was used for image reconstruction, followed by clamping negative reconstruction intensities to zero.

Note that the upper bound for the image quality after a correction is determined by the limitations of the motion model. Detector shifts do not take the scaling effects into account that occur from motion orthogonal to the detector. To allow for a separate evaluation of the motion model and the motion estimation we perform three types of reconstructions. Initially, we created uncorrected reconstructions using the original projection matrices $\hat{\boldsymbol{P}}_k$. Thereafter, the parameter vector as obtained from the minimization of Eqn. 6 was used to create the FCC-based reconstruction. Additionally, we built reconstructions using the extracted ground-truth motion $\boldsymbol{g}_k$ directly as input to Eqn. 17. These reconstructions serve as an upper bound for the image quality that can be achieved given this motion model.

On top of a qualitative comparison we added a quantitative evaluation of the reconstructed volumes with respect to the motion-free reconstruction. We computed the root-mean-squared-error (RMSE) as well as the structural-similarity index (SSIM) proposed by Wang et al. (2004). Let $f(x, y, z)$ be the reconstruction that we want to evaluate and $r(x, y, z)$ the motion-free reference reconstruction. Then, the RMSE can be computed as

$$\mathrm{RMSE}(f, r) = \left(\frac{1}{R_x R_y R_z} \sum_{x=1}^{R_x} \sum_{y=1}^{R_y} \sum_{z=1}^{R_z} \left(f(x, y, z) - r(x, y, z) \right)^2 \right)^{\frac{1}{2}} . \tag{34}$$

It has been shown that the RMSE may not reflect the visual perception of image differences (Wang et al.; 2004). In contrast, the SSIM was designed to incorporate this information by merging information about intensity, contrast and structural differences of two images. It is defined by

$$\mathrm{SSIM}(f, r) = \frac{(2\mu_f \mu_r + C_1)(2\sigma_{fr} + C_2)}{(\mu_f^2 + \mu_r^2 + C_1)(\sigma_f^2 + \sigma_r^2 + C_2)} , \tag{35}$$

where f and r correspond to volume regions, μ_f and μ_r are the mean intensity values of the regions, σ_f and σ_r are their standard deviations and σ_{fr} their correlation. Further, C_1 and C_2 are constants that assure stability for regions which either have a zero-mean ($\mu_f^2 + \mu_r^2 \approx 0$) or are free of variations ($\sigma_f^2 + \sigma_r^2 \approx 0$). We adjusted them to $C_1 = (0.01B)^2$ and $C_2 = (0.03B)^2$ as proposed by Wang et al. (2004), where B is the intensity range of the full reference volume $r(x, y, z)$. We evaluate the SSIM (see Eqn. 35) locally for each voxel, using a cubical neighborhood of size $9 \times 9 \times 9$. The final value is obtained by the mean over all voxel-wise SSIM values.

4. Results

4.1. Qualitative Results for the FORBILD Cases

To provide a qualitative comparison we extracted the central axial slices of the reconstructed volumes. The first row of Figure 6 shows the reconstructions of the motion-free reference and the following rows show the reconstruction of high-frequency

motion cases *Oscil*, *Chirp* and *Rect*, respectively. Similarly, Figure 7 depicts two cases that have initially been corrupted with a low-frequency motion pattern, namely *LF-1* and *LF-2*.

On the columns of Figure 6 and Figure 7 we show reconstructed slices without correction (NoCorr), corrected using the proposed method (FCC), and the reference reconstruction using the ground-truth 2-D motion (CorrGT). The fourth column shows zoomed details of the head's anterior region, containing a high amount of bone as well as a low-contrast region representing the left and right eyeball. Similarly, the fifth column depicts details of the ear region which contains high intensity variations due to bone material and inserted air bubbles. The zoom regions are indicated by yellow (dotted) and green (dashed) bounding boxes on the corresponding NoCorr reconstructions. From top to bottom the zoomed images correspond to NoCorr, FCC, and CorrGT.

For the *Oscil* (cf. Figure 6(f-j)) as well as for the *Rect* motion (cf. Figure 6(p-t)), we can see clear streaking and blurring artifacts in case no correction was applied. The *Chirp* motion has a lower amplitude, yet, a broad range of motion frequencies is observed, mainly resulting in streaking artifacts (cf. Figure 6(k-o)). The proposed method was able to improve image quality for all motion types. The detailed views of the anterior and the ear show that the structural loss of the *Rect* and *Oscil* motion could be well restored and the low-contrast eyeballs and the cerebral fluid can be identified after the correction. The structural image quality of the corrected reconstructions is close to that of the reconstructions using the ground-truth motion. Streaking artifacts are also reduced to a large extent, which can be clearly seen when comparing Figure 6k with Figure 6l. From the reconstructions based on the ground-truth 2-D motion we can see that a more accurate motion estimation could further improve streak reduction.

When applied to motion-free projection images our method slightly misestimates motion in detector u-direction as can be seen in Figure 11a. Note that Figure 6a is identical to the reconstruction in Figure 6c as the 2-D ground-truth motion was zero. Both correspond to the motion-free reference volume $r(x, y, z)$.

The reconstructions of the low-frequency motion in Figure 7 are dominated by blurring artifacts. Streaking is only present for the *LF-1* motion due to an angular discontinuity based on the different starting and end positions of the object. The streaking could be well reduced by our correction as can be seen in Figure 7b and Figure 7d. The low-frequency intensity variations resulting from the *LF-2* motion appear well restored when comparing Figure 7f and Figure 7g. The zoomed images of the ear region in Figure 7e and Figure 7j reveal that the spatial alignment could be enhanced for both motion types. However, there is still a noticeable difference in edge sharpness when comparing reconstructions based on the ground-truth motion and the proposed method.

In addition to the reconstruction results we show the sinograms and the logarithmically scaled magnitude of their 2-D spectra. The sinograms and spectra correspond to the central detector row. They are depicted for the high- and low-frequency motion in Figure 8 and Figure 9, respectively. The first three columns show

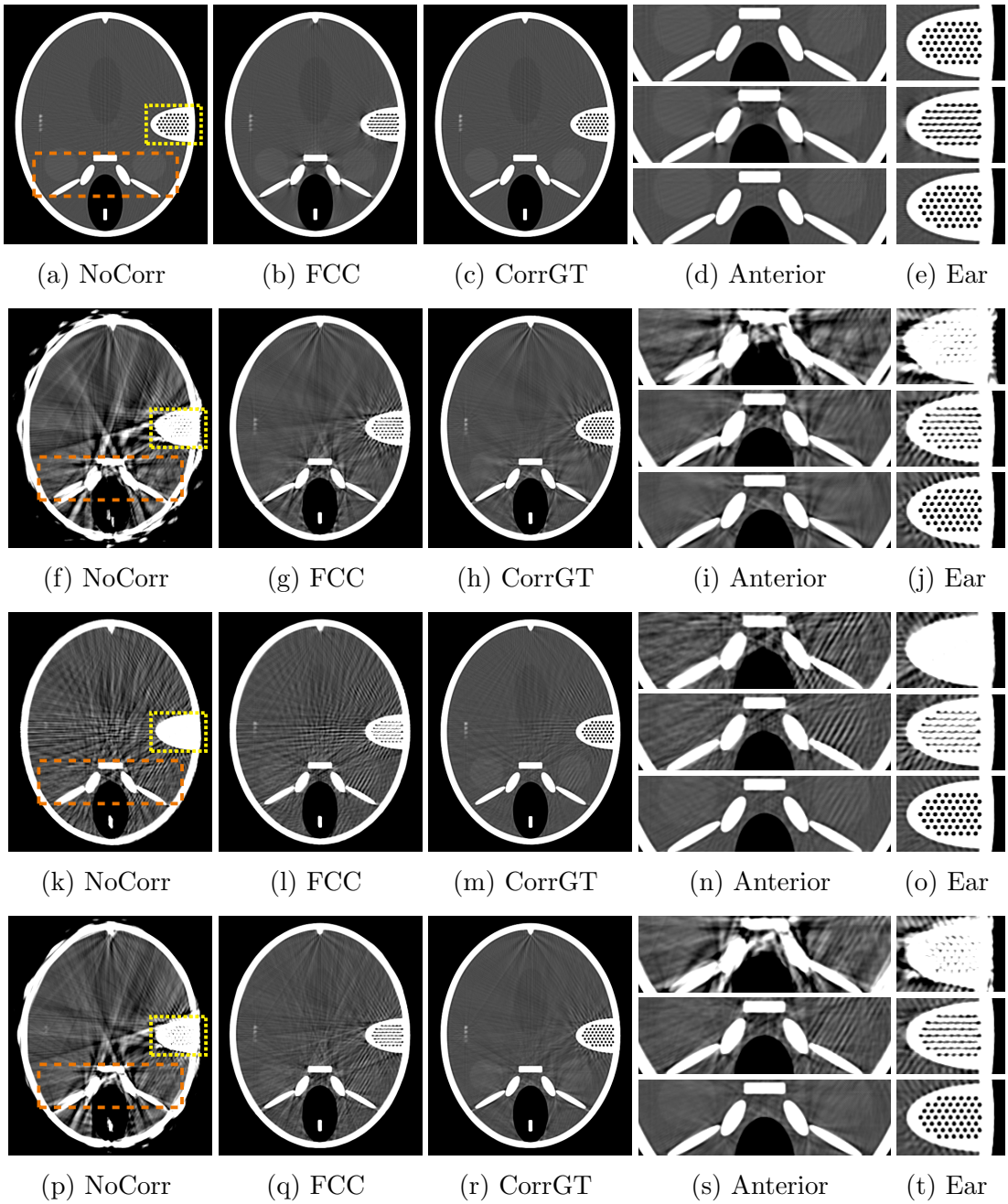
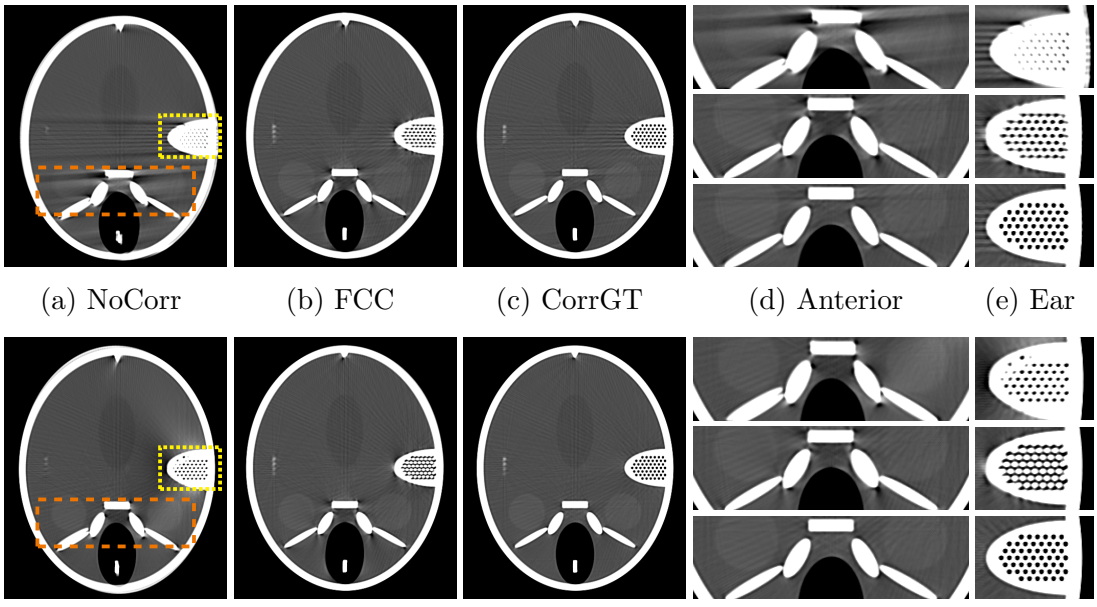


Figure 6: Top to bottom: Axial reconstructions for motion-free (*None*) and high-frequency motion data (*Oscil*, *Chirp* and *Rect*). Left to right: Reconstructions without correction (NoCorr), corrected using FCC and corrected with the ground-truth motion (CorrGT). Zoomed regions of Anterior and Ear regions with results from NoCorr, FCC and CorrGT as rows. (W: 697 HU, C: 105 HU).

the motion-corrupted sinogram, the corrected sinogram and a zoom to the sinograms' region highlighted by a red (dotted) box in the motion-corrupted sinograms. The next two columns depict the corresponding spectra with a zoom to a low-frequency region in the last column.



(a) NoCorr (b) FCC (c) CorrGT (d) Anterior (e) Ear

(f) NoCorr (g) FCC (h) CorrGT (i) Anterior (j) Ear

Figure 7: Axial reconstructions for low-frequency motion data (*LF-1* and *LF-2*). Arrangement of results identical to Figure 6. (W: 697 HU, C: 105 HU).

High-frequency motion artifacts are clearly visible in the sinogram domain, especially for *Oscil* (cf. Figure 8g) and *Rect* (cf. Figure 8s), due to their higher motion amplitudes. Also the low-amplitude *Chirp* motion can be well identified in the zoomed sinogram in Figure 8o. Compared to the spectrum of the motion-free reference in Figure 8d, the triangular regions for all high-frequency motion patterns show a substantial increase of energy. Interestingly, *Chirp* motion exhibits a wider redistribution of spectral energies into the triangular regions as the *Oscil* motion. This indicates that the redistribution of energies is more sensitive to the motion's frequency rather than its amplitude. Our correction method was able to restore the distorted sinogram outline which can be clearly seen in the zoomed sinogram regions of Figure 8(i,o,u). Additionally, large portions of the spectral energies could be moved back to the object's support region, leading to a clearly visible restoration of the triangular regions.

When applying our motion correction to the motion-free data we notice very little difference in the sinogram (cf. Figure 8a and 8b) as well as the spectral domain (cf. Figure 8d and 8e).

Due to the low-frequency object motion, only little effect can be seen in the sinogram domain for the *LF-1* and *LF-2* motion types. Similarly, the difference between corrected and motion-corrupted sinograms is only visible from a slight shift and deformation of the sinogram boundary in Figure 9c and 9i. The changes to the triangular regions are substantially smaller than for the high-frequency motion patterns. The vertical bar in the center of the spectrum for the *LF-1* motion type corresponds to the angular

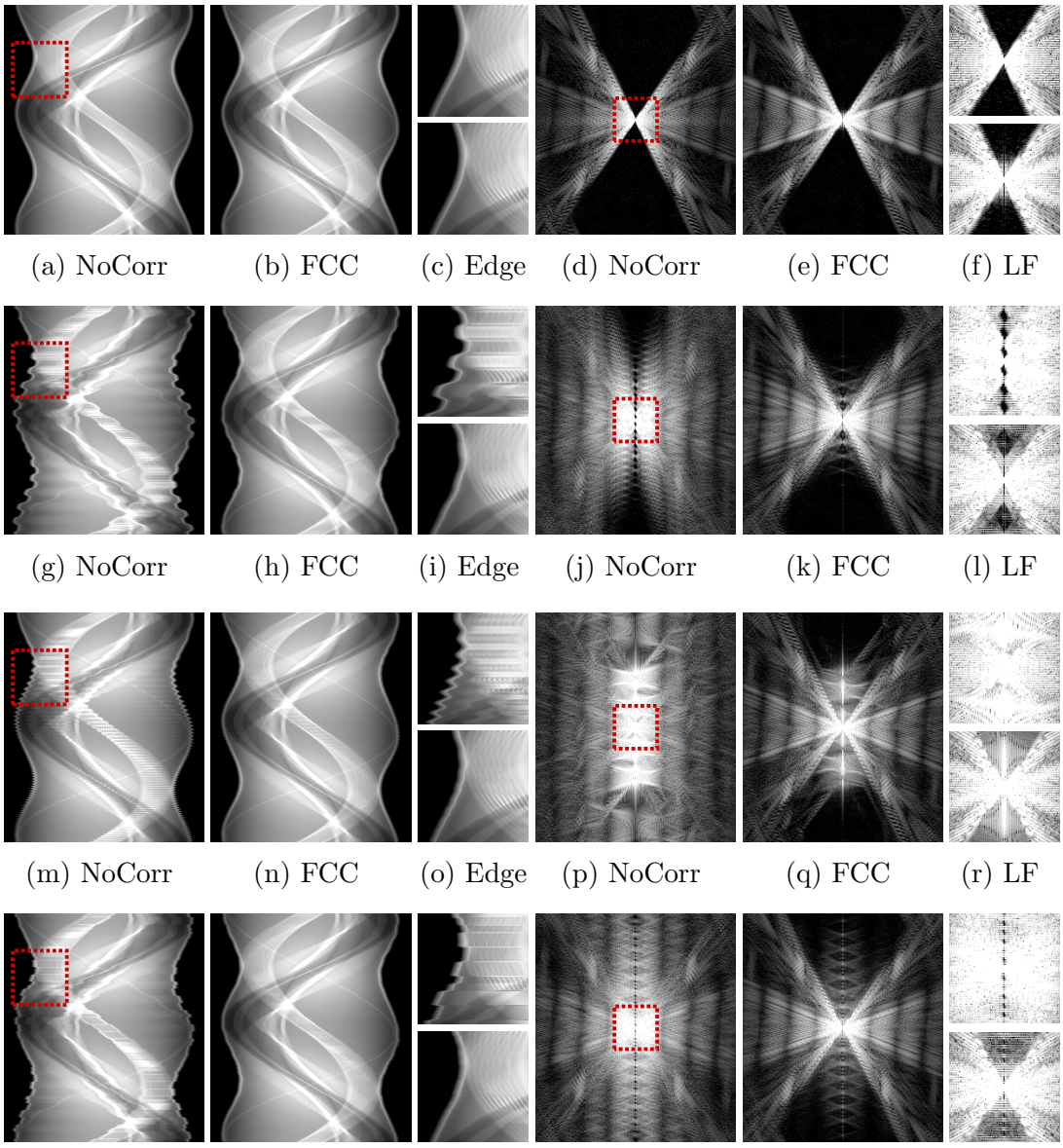
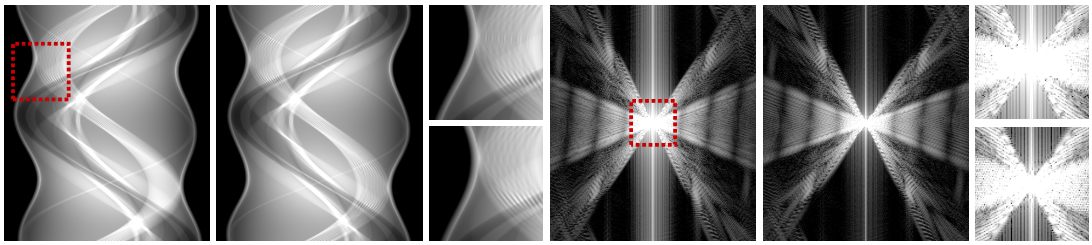
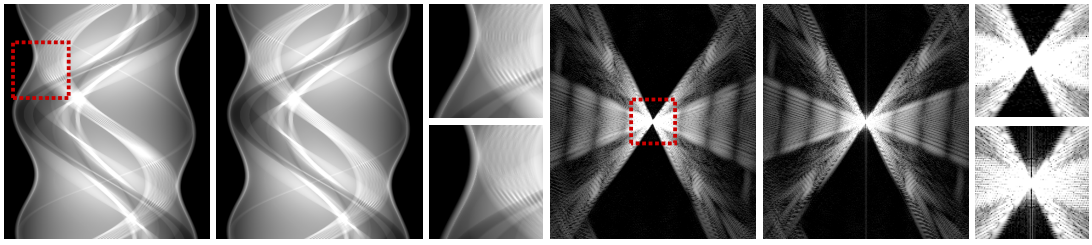


Figure 8: Sinograms and their log-spectra for the central slice of the projection stack. Top to bottom: Motion-free (*None*) and high-frequency motion data (*Oscil*, *Chirp* and *Rect*). Left to right: Sinogram without correction (NoCorr), corrected using FCC, two zooms at the sinograms' edge (Edge), the log-spectrum of the uncorrected and corrected sinogram and a zoom to the lower frequencies of the spectra (LF). Visualization window for sinograms and log-spectra were $[1.5, 5]$ and $[5, 10]$, respectively.

discontinuity due to varying start and end positions of the object. In contrast corrections for motion type *LF-2* in Figure 9j are almost identical to the motion-free spectrum in Figure 8d. Our correction method was able to visually reduce the energy inside the spectral regions belonging to the *LF-1* motion. However, the energy spread of the *LF-2*



(a) NoCorr (b) FCC (c) Edge (d) NoCorr (e) FCC (f) LF



(g) NoCorr (h) FCC (i) Edge (j) NoCorr (k) FCC (l) LF

Figure 9: Sinograms and log-spectra for low-frequency motion data (*LF-1* and *LF-2*). Arrangement of results and visualization windows identical to Figure 8.

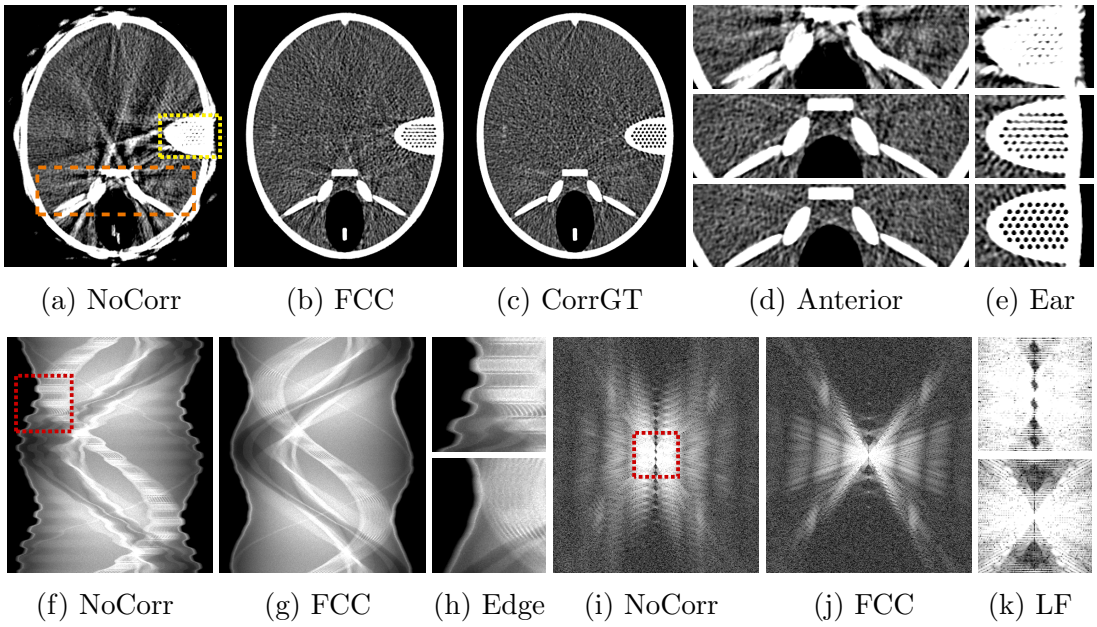
motion in Figure 9l into the triangular regions even appears to increase.

In Figure 10 we show examples of reconstructions, sinograms, and spectral domain visualizations of a noise-prone case. In general, our correction method performed similarly when applied to data with a considerable amount of noise. Despite the reduction in image quality due to the noise itself, we could not observe a lower accuracy of the motion estimation when compared to the noise-free examples. This is further supported by the quantitative evaluation of the motion estimates in Table 3. Even though Figure 10i and 10j clearly show the noise in the spectral domain, we observed only minor differences between the motion estimates of noisy and noise-free data.

4.2. Quantitative Results for the FORBILD Cases

4.2.1. Image-Based Measures The SSIM and RMSE results are given in Table 2a and Table 2b, respectively. The first three rows of the tables show the results for the ideal, noise-free data, whereas the last three rows correspond to the noisy data. We evaluated reconstructions without correction (NoCorr), the corrected versions using the proposed method (FCC) and the corrections using the 2-D ground-truth motion (CorrGT). For presentation SSIM and RMSE values are scaled by a factor 100. Thus, the maximum SSIM value is 100, which corresponds to the minimum RMSE of 0, i.e., both volumes would be identical. Note that the CorrGT values build an upper bound for the FCC results, as they merely reflect loss of image quality due to the motion model itself.

For all motion-corrupted datasets, the FCC led to a substantial increase of the SSIM value and a clear reduction of the RMSE when compared to the NoCorr values. The RMSE values of the FCC results are in a comparable range for all motion cases.



(a) NoCorr (b) FCC (c) CorrGT (d) Anterior (e) Ear

(f) NoCorr (g) FCC (h) Edge (i) NoCorr (j) FCC (k) LF

Figure 10: Top: Axial reconstructions for the noisy *Oscil* data. Bottom: Sinograms and spectra for central slice of projection stack. The motion estimation worked equally well as for the ideal data, despite the high amount of noise. Visualization windows are equal to Figure 6 and Figure 8.

In contrast, the improvements in SSIM are not as consistent. Cases with very high-frequency motion, i.e., *Chirp* and *Rect*, showed the highest difference between CorrGT and FCC, whereas, for other motion types FCC yielded measures close to the CorrGT results. From the SSIM results of CorrGT we can see that the high-frequency motions, i.e., *Oscil*, *Chirp*, and *Rect*, were most difficult to reconstruct, which may correspond to the increased level of streaking artifacts due to fast variations in the motion. These motion types also had the lowest SSIM if no correction was applied (NoCorr). When applying the FCC to a motion-free dataset (*None*), we notice a decrease of the SSIM by 1.6 and 1.7 for the ideal and noisy data, respectively. Similarly, Table 2b shows an increase of the RMSE by 2.6 for the ideal and 1.5 for the noisy data. Noise did not substantially change the RMSE value in case of the motion-corrupted data, yet the uncorrected, motion-free reconstruction increased from 0 to 1.4. The SSIM appears to be more sensitive to the noise with a reduction from 100 to 85.3 in case of the motion-free reference reconstruction. Despite the general reduction of SSIM values in case of noise, the relative changes of the FCC compared to NoCorr results are similar to those of the noise free cases, showing the methods robustness against noise.

4.2.2. Motion Estimation Accuracy Figure 11 shows the estimated motion s_k and t_k in blue for the high-frequency and the low-frequency motion, and the projected 2-D ground-truth motion $\boldsymbol{g}_k$ in red on noise-free data. Additionally, Table 3 shows the MAD and

		None	*Oscil*	*Chirp*	*Rect*	*LF-1*	*LF-2*
Ideal	NoCorr	100.0	61.7	73.6	69.2	85.1	92.7
	FCC	98.4	87.6	84.0	88.5	96.2	97.4
	CorrGT	100.0	90.7	95.4	94.2	97.6	99.1
Noisy	NoCorr	85.3	53.8	64.2	60.6	72.4	78.5
	FCC	83.6	74.9	72.3	75.7	81.9	82.9
	CorrGT	85.3	77.5	81.2	80.3	83.2	84.6

(a) SSIM results

		None	*Oscil*	*Chirp*	*Rect*	*LF-1*	*LF-2*
Ideal	NoCorr	0.0	11.4	4.8	8.0	13.7	8.2
	FCC	2.6	3.6	3.3	3.3	3.4	3.4
	CorrGT	0.0	2.1	0.8	1.7	2.6	2.0
Noisy	NoCorr	1.4	11.5	5.1	8.1	13.7	8.3
	FCC	2.9	3.9	3.6	3.6	3.7	3.7
	CorrGT	1.4	2.6	1.7	2.2	3.0	2.4

(b) RMSE results

Table 2: Quantiative results for reconstructions without correction (NoCorr), corrected using the FCC method, and corrected using the 2-D ground-truth motion (CorrGT). Evaluation was performed for all motion types (see columns) with ideal and noisy data.

its SD between 2-D ground-truth and estimated motion for the ideal and noisy data. Note, that the geometric effect at the center of the reconstruction is approximately scaled by the ratio of source-isocenter- to source-detector-distance $\frac{L}{L+D}$, which in our case evaluates to $\frac{1}{2}$.

The motion in vertical detector direction t_k is estimated very well in all cases and can even follow the sharp transitions of the *Chirp* and *Rect* motion (cf. Figure 11c and 11d). Also, on motion-free data (*None*) only very little fluctuation occurs. Quantitative results for t_k in Table 3 support these findings, where the highest MAD values evaluate to 708 µm for *Chirp*, 415 µm for *Oscil* and 293 µm for *Rect*. After scaling by $\frac{1}{2}$ to estimate potential effects on the central reconstruction domain, all MAD values for t_k result in values below the reconstruction voxel size of 500 µm.

Somewhat larger deviations are observed for the horizontal detector direction s_k. As can be seen in Figure 11 all estimated motion curves show a low-frequency offset to the ground-truth. Yet, the general trend of the motions' variation is well covered, particularly for high-frequency motion in Figures 11(b,c,d). The reduced accuracy is also reflected in the MAD and SD values in Table 3, where the MAD values range between 814 µm for the motion-free data up to 1184 µm for the *Chirp* motion. The SD values are generally higher than those of t_k with the highest values for *Chirp* and *Rect*.

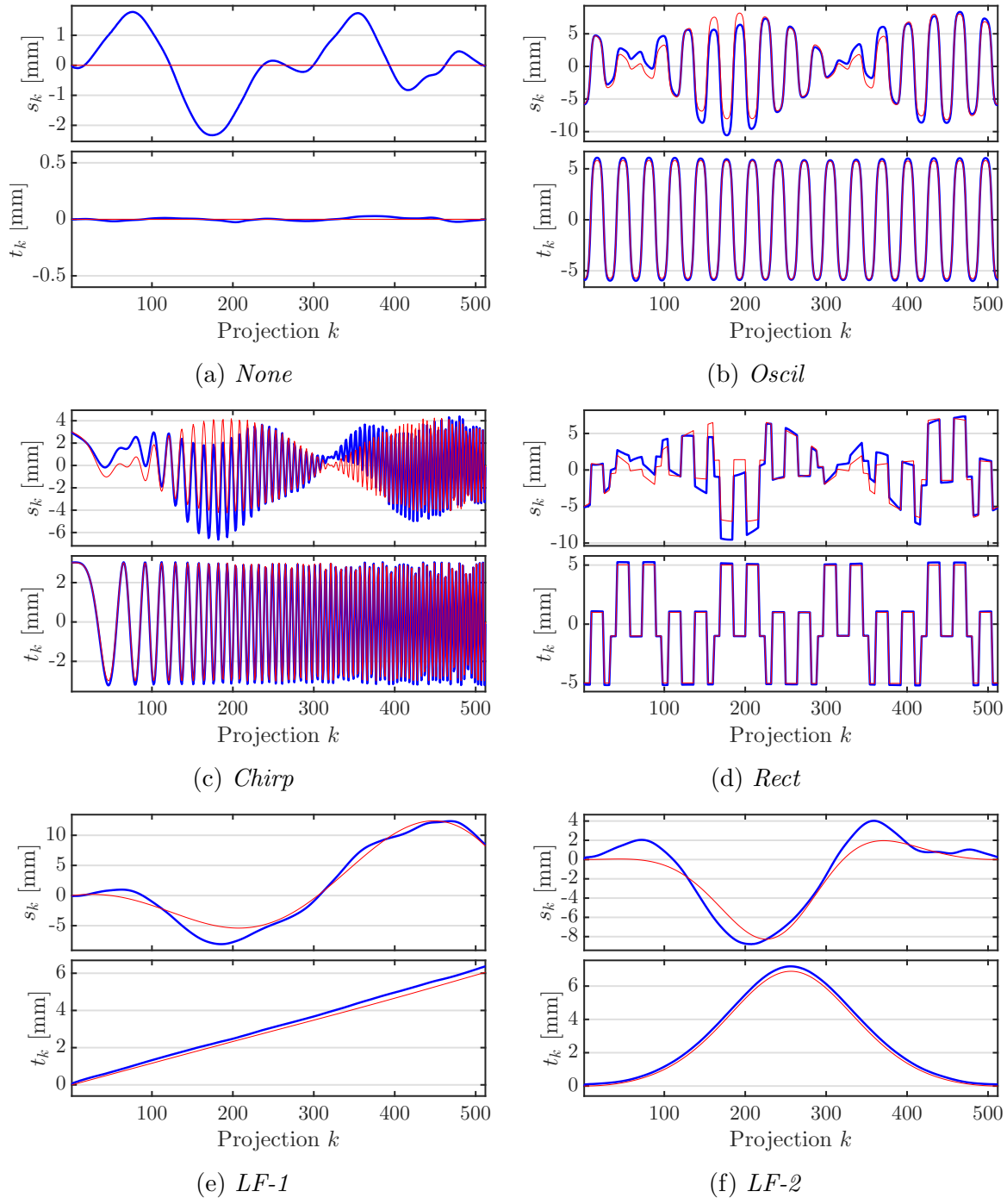
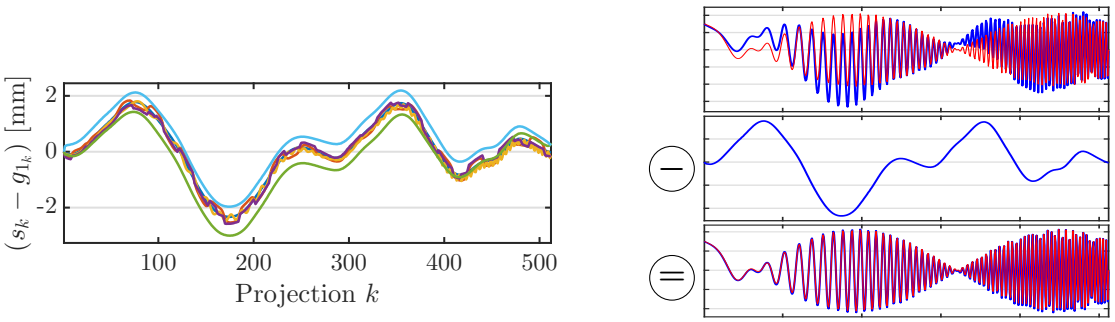


Figure 11: Plots of estimated motion parameters s_k and t_k (solid, blue line) and the 2-D ground-truth motion (dotted, red line).

We made an interesting observation when having a closer look at the difference of estimated and ground-truth motion for the horizontal detector direction, i.e., $s_k - g_{1_k}$. In Figure 12a we plotted this difference for all cases. It shows that the error is dominated by the motion estimated on the motion-free data. Thus, the estimated motion s_k always consists of an additive mixture of the actual motion signal and the error estimated on

		None	*Oscil*	*Chirp*	*Rect*	LF1	LF2
Ideal	s	814 (472)	934 (548)	1184 (681)	913 (632)	909 (631)	960 (447)
	t	11 (0)	415 (121)	708 (295)	293 (228)	217 (5)	194 (4)
Noisy	s	800 (457)	923 (541)	1171 (673)	899 (620)	881 (543)	946 (431)
	t	11 (0)	418 (123)	708 (295)	293 (228)	125 (5)	195 (4)

Table 3: Mean absolute difference and standard deviation (in brackets) between estimated motion parameters and 2-D ground-truth motion given in µm.



(a) Horizontal motion error

(b) Additive model

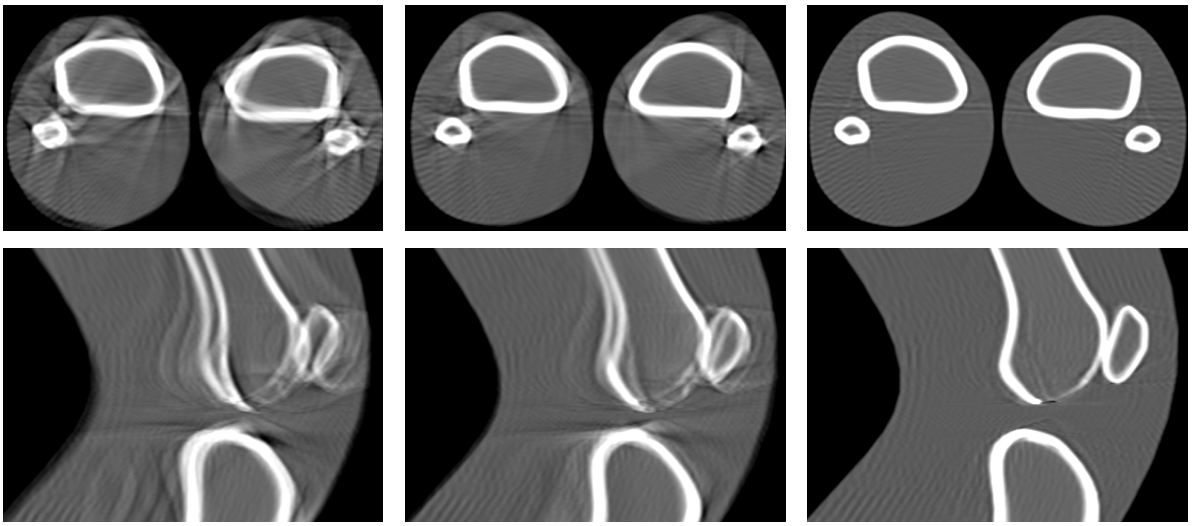
Figure 12: (a) Differences between estimated and ground-truth motion show very similar deviations for horizontal direction. (b) Subtracting the motion, estimated on the motion-free reference data, from the estimated motion in case of motion-corrupted data, reveals a very accurate motion estimation.

the motion-free data. In Figure 12b we subtract the motion estimated on the motion-free data (cf. Figure 11a) from the estimated *Chirp* motion in Figure 11c resulting in a very accurate estimation of the *Chirp* motion. We noticed the same effect for the other motion patterns. Note, that the approach shown in Figure 12b is for visualization purposes only and was not applied to obtain reconstruction results.

The deviation for the motion-free data has an MAD of 814 µm and is in the range of −2.33 mm to 1.78 mm. The approximate motion effect in at the isocenter will be scaled by a factor of $\frac{1}{2}$ (see Section 4.2.2) resulting in 407 µm which is still below the reconstruction voxel size of 500 µm.

4.3. XCAT Knee Phantom Under Weight-Bearing Conditions

Fig. 13 shows reconstructed images of the XCAT phantom. The images show axial slices through tibia and fibula (top row) and sagittal slices of right-sided femur, patella, and tibia (bottom row). From left to right, we can see reconstructions without motion correction (NoCorr) in Fig. 13a, after correction with FCC in Fig. 13b, and the motion-free data (GT) in Fig. 13c. Severe motion artifacts can be seen in case no correction is applied, manifested by streaking and blurring. In addition, the outlines of femur and



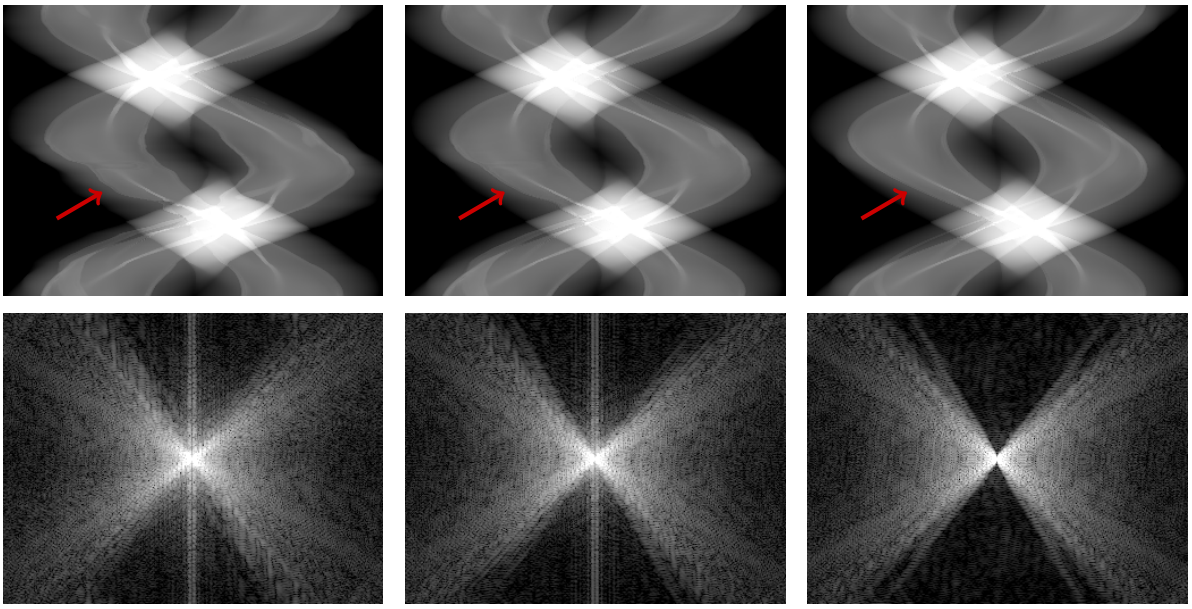
(a) NoCorr (b) FCC (c) GT

Figure 13: Reconstructions of the XCAT squat phantom. Top row: Axial slices inferior to the knee joint. Bottom row: Sagittal slices through the right leg. The colmuns correspond to reconstructions without correction (NoCorr), after correction using FCC, and of the motion-free projection images (GT). (W: 1815 HU, C: 246 HU).

patella are hardly visible in the sagittal slice of Fig. 13a. Our method was able to restore most of the structure at left and right tibia and shows a clearer visibility of the fibulas. The structure of femur and patella improved by a large amount, yet, the difference to the ground-truth reconstruction is still substantial. Overall, the skin surface and soft tissue appears sharper after correction and contains less streaking artifacts.

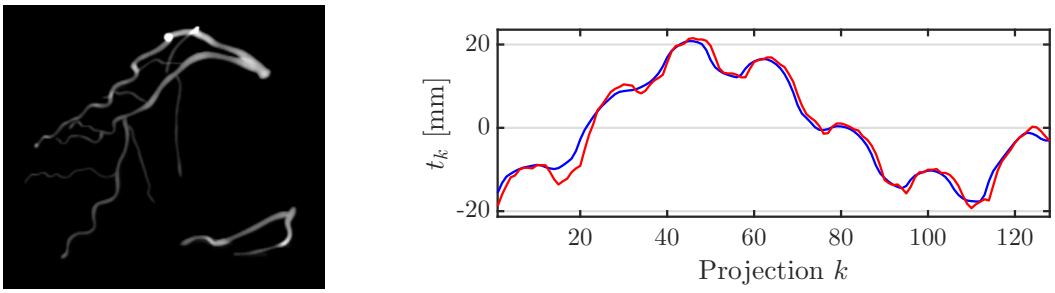
The same quantitative evaluation as used for the FORBILD phantoms was performed. The resulting values support the visual impression in Fig. 13, where NoCorr yielded an SSIM of 69.7 and an RMSE of 5.69. After applying the proposed motion correction, image quality improved, resulting in an SSIM of 79.2 and an RMSE of 5.19.

Sinograms of the XCAT knee phantom are shown in the top row of Fig. 14. Irregular trajectories of object boundaries clearly depict the motion artifacts for the uncorrected sinogram in Fig. 14a. A location with a very apparent motion artifact is indicated by a red arrow. After correction with FCC, the sinogram appears smoother and a clear improvement can be seen at the location of severe motion. Yet, we still observe several differences when comparing FCC and GT (cf. Figures 14b and 14c). The spectra, shown with a logarithmic scaling at the bottom row of Fig. 14, depict the triangular regions for a vertical frequency of $\psi = 0\,\mathrm{mm}^{-1}$. The proposed motion correction could remove a large amount of energy from the triangular regions compared to NoCorr, yielding a sharper outline of the zero regions. However, after correction we still notice a higher amount of energies located within the zero regions, compared to the ground truth in Fig. 14c. In general, the spectral energies are in line with their assigned sinograms and



(a) NoCorr (b) FCC (c) GT

Figure 14: Sinograms and spectra for the XCAT squat phantom. Top row: Sinograms extracted at a central slice. Bottom row: Spectra of the central slice, i.e., for $\psi = 0\,\mathrm{mm}^{-1}$ after logarithmic scaling. Visualization windows of sinograms and spectra are $[0.05, 0.5]$ and $[4, 12]$, respectively.



(a) Projection of DSA (b) Respiratory motion estimates

Figure 15: (a) Angiographic image after virtual DSA. (b) Comparison of estimated (blue) and ground-truth (red) detector translations for respiratory motion estimation.

the achieved image quality.

4.4. Respiratory Motion Estimation in Rotational Angiography

We conducted a feasibility study for respiratory motion in rotational angiography. The breathing motion is estimated by vertical detector translations only. Ground-truth translations were extracted by tracking two bifurcation points at a proximal and distal point within the coronary tree. Finally, the ground-truth is given by the average displacement of the tracked point locations.

Fig. 15a shows a typical projection image after application of virtual DSA, with the background being well removed. The respiratory motion estimated by FCC is depicted as blue curve in Fig. 15b. We can clearly see that the estimation accurately follows the extracted ground truth, shown as red curve. Even the local maxima, which correspond to craniocaudal displacement during the cardiac cycle, are well covered by FCC. Even though FCC are not yet applicable for non-rigid deformations, they may be used to improve initialization for a non-rigid cardiac reconstruction.

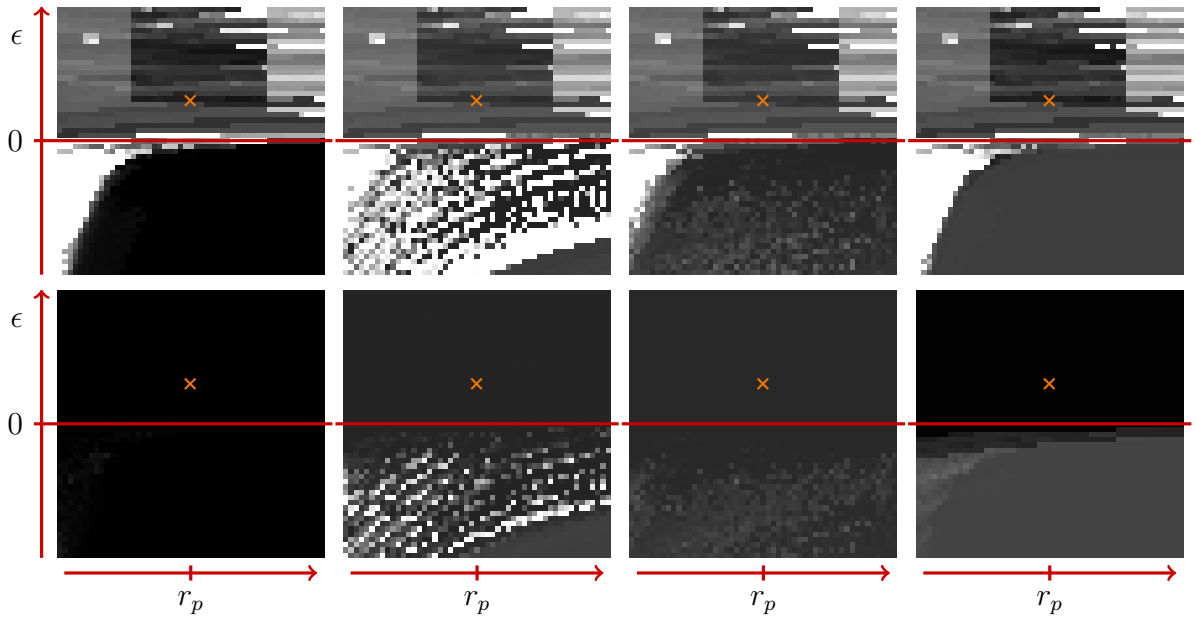
4.5. Mask Sensitivity Analysis

The goal of this evaluation is to verify the effect of the mask's parameters. Additionally, we aim to obtain a deeper insight in the systematic deviation of motion estimated in detector u-direction. According to our definition in Eqn. 3 and the adjustment from Section 2.4, the shape of the triangular regions depends on the estimated object radius r_p and the adjustment parameter ϵ. To get a better understanding of their non-intuitive effect on the estimation accuracy, we performed a 2-D grid-search over r_p and ϵ values. The MAD has been computed for each parameter configuration and for each FORBILD case. Note, that due the computational complexity (see Section 4.6), the projection volume was downsampled by a factor of four, which also led to generally worse results of the MAD.

We varied the object radius r_p from its estimated value (cf. Eqn. 25) in the range of $\pm 50\,\mathrm{mm}$ with a sampling distance of $2\,\mathrm{mm}$. Accordingly, the extension parameter ϵ has been adjusted in the range of $\epsilon \in [-0.01, 0.01]$ with a sampling distance of 4×10^{-4}. Note, that a negative ϵ actually decreases the triangular regions, similar to a morphological erosion, whereas, a positive value increases the region's coverage as proposed in Fig. 2.

Fig. 16 shows the MAD results for selected FORBILD cases. The images hold MAD errors in its gray values whereas the image x- and y-axis encode r_p and ϵ, respectively. The top row corresponds to $MAD(\boldsymbol{\alpha})_1$ and the bottom row to $MAD(\boldsymbol{\alpha})_2$, i.e., to detector u- and v-direction, respectively. Interestingly, the regions where $\epsilon > 0$, i.e., the mask has been extended, are reproducible over all cases. However, for $\epsilon < 0$, i.e., for a reduced mask size, the results seem to depend strongly on the type of simulated motion. For detector v-direction we observe generally better results than for the horizontal translations. With an $\epsilon > 0$ consistently good results were achieved, relatively independent of variations in r_p, which indicates a robust method for estimation of vertical detector translations.

The orange cross represents the parameter settings used for our evaluation, thus, they correspond to the additive bias presented in Fig. 12. For certain motion types we were able to find parameter settings with an $\epsilon < 0$ (smaller mask) that led to improved motion estimates in u-direction. For *None*, this is the case within the black area in the lower right of the top image in Fig. 16a. For *Oscil* these beneficial configurations are scattered within the streaky pattern, meaning that small variations could cause



(a) *None* (b) *Oscil* (c) *Chirp* (d) *LF-2*

Figure 16: Sensitivity analysis of mask parameters. Intensities correspond to MAD, evaluated after each optimization. Top row: Results for detector u-direction. Bottom row: Results for detector v-direction. Parameter $\epsilon \in [-0.01, 0.01]$ and r_p is deviating from its initial estimate by $\pm 50\,\text{mm}$. Note, the reproducibility over all cases for $\epsilon > 0$ (larger mask), whereas results for $\epsilon <= 0$ (smaller mask) depend stronger on the motion type. Visualization window: 0 mm to 10 mm.

large changes in the accuracy of the estimated motion. None of these settings could be identified for low-frequency motions *LF-1* and *LF-2*.

In conclusion, an $\epsilon > 0$ led to higher robustness and stable results over all motion types. Yet, depending on the motion, the error might be reducible with a reduction of the mask size ($\epsilon < 0$). However, the outcome may be less predictable and parameter tuning becomes more difficult.

Generally, we noticed that u-direction motion estimates with a larger error, i.e., corresponding to the white areas of the images in Fig. 16, are always dominated by a relatively harmonic, low-frequency signal. This is well visualized in Fig. 17, where we show 100 estimated motion curves for randomly selected parameter configurations of the *Oscil* case. For all estimations we see that the actual oscillating motion is well estimated, yet, it is added to the clearly visible erroneous low-frequency perturbation.

One reason for the presented low-frequency bias could be the heuristic extension of 2-D to 3-D FCC introduced in Section 2.1. To verify the validity of our 3-D extension, we performed an identical sensitivity analysis for mask parameters on 2-D fan-beam data. We repeated the same numerical FORBILD simulations in 2-D, including the rendering of sinograms. Therefore, only the central slice of the detector is used and the axial direction of the individual motion types was set to $\Delta z_k = 0$. Due to the computational

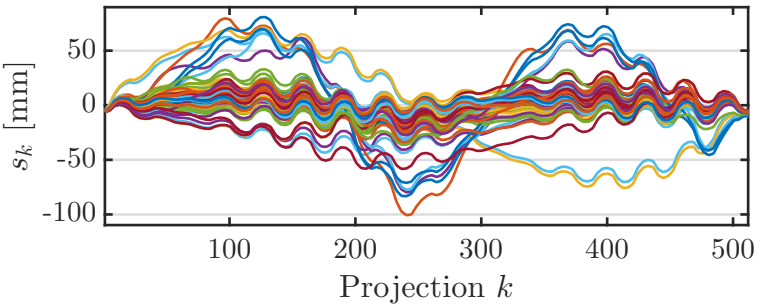
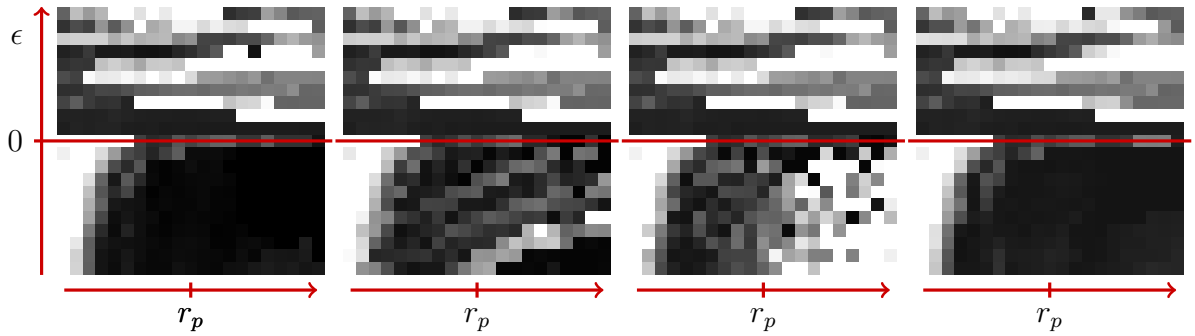


Figure 17: Random selection of estimated u-direction translations for *Oscil* motion for various combinations of r_p and ϵ.



(a) *None* (b) *Oscil* (c) *Chirp* (d) *LF-2*

Figure 18: Sensitivity analysis of mask parameters for 2-D fan-beam datasets. The parameter ranges for r_p and ϵ are identical to the 3-D evaluation, yet, a lower a sampling rate has been applied. Behavior, of optimization method follows same trend as shown for the 3-D results (cf. Fig. 16). Visualization window: 0 mm to 30 mm.

complexity, we used a lower sampling rate, yet, the parameter ranges were identical to the 3-D experiment.

The resulting sensitivity maps for the 2-D fan-beam case are shown in Fig. 18. The motion refers only to detector shifts in u-direction, as for the fan-beam case no v-direction is present. We also observe reproducible results for $\epsilon > 0$, whereas the distributions of MAD values for $\epsilon < 0$ are very similar to the 3-D evaluation. We also observed very similar motion patterns than were estimated for the 3-D cases. This includes the low-frequency bias shown in 17. Overall, the 3-D extension of FCC performed very similar to the already more established 2-D FCC. In conclusion, the results indicate that the systematic low-frequency bias does not originate from the approximations made when extending FCC to 3-D cone-beam CT.

4.6. Robustness and Runtime

To evaluate the computational effort of the proposed algorithm, we performed runtime measurements for all simulations involving the FORBILD phantom. Thus, a total of

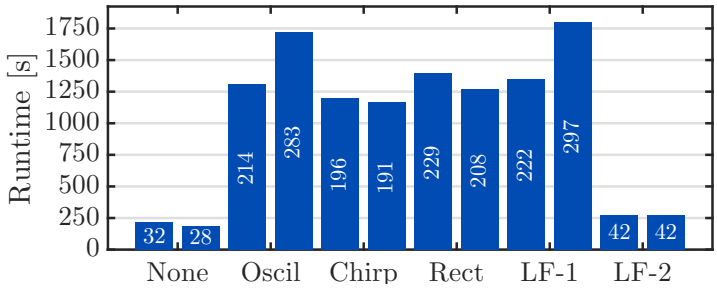
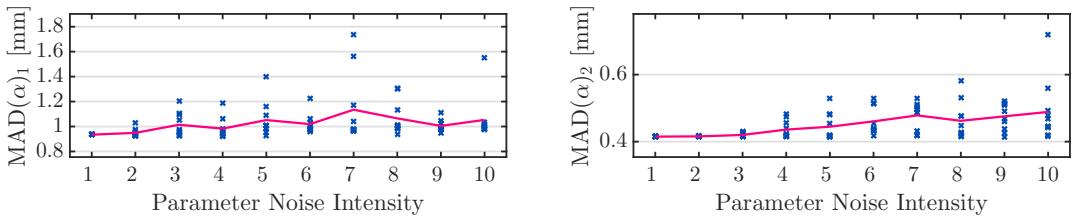


Figure 19: Total runtime of FCC motion estimation for all FORBILD cases. Depending on the number of iterations (printed in white), the computation times lie between 3 min to 30 min.

12 cases have been analyzed, including different motion patterns and noisy as well as noise-free data. All performance measurements were collected on a desktop computer (Windows 7 64-Bit, Intel Xeon CPU E5-1650 v3, 3.5 GHz, 32 GB RAM, Nvidia Quadro K5200, 8 GB Graphics Memory). The maximum number of iterations was never reached by the optimizer, thus, the optimization was either terminated by too small values within the function gradient or too little variations of the motion parameters.

Fig. 19 depicts the total runtime for all FORBILD cases, where the left and right bar of each case correspond to noise-free and noisy data, respectively. The number of iterations used by the optimizer is reported within the bars. The longest computation time of 1802 s was measured for the noisy *LF-1* data and the lowest time of 189 s for the noisy, motion-free data. The overall runtime appears to scale linearly with the number of iterations. Runtimes of individual components are independent of the iteration count and provide a fairer comparison. Overall, 1984 iterations were performed, using the same number of gradient computations. The median gradient computation took 5734 $[5725, 5743]$ ms, where the brackets report lower and upper quartiles. Each gradient computation consisted of 512 computation steps for partial derivatives with respect to s_k and t_k, having a median computation time of 11.2 $[11.15, 11.24]$ ms. The number of cost-function evaluations was 2125, hence, higher than the iteration count due to the optimizer's line-search approach. The median evaluation time was 87.14 $[87.58, 88.27]$ ms.

During our evaluation we initialized the estimated motion parameters with a null vector. To evaluate the robustness of the optimization process with respect to larger deviations in the initial motion estimates, we repeatedly initialized the motion parameters with detector translations drawn from a 2-D Gaussian distribution. A maximum intensity was determined by setting the standard deviations in u- and v-direction to $\sigma_u = \frac{131\,\mathrm{mm}}{3}$ and $\sigma_v = \frac{25\,\mathrm{mm}}{3}$, i.e., to a third of the non-radiated image borders. Translations larger than the border were truncated to the border width. Thus, no unrealistic periodic reappearance of the image content was ensured. The standard deviations were varied ten times by a factor equally distributed within $[0, 1]$. For each factor, we repeatedly applied the FCC algorithm ten times, yielding a total of 100



(a) Detector u-direction (b) Detector v-direction

Figure 20: MAD for the *Oscil* dataset. Parameters were initialized with increasing intensity of Gaussian noise. Blue crosses show ten repetitions per intensity level, while the red curve depicts the average error.

runs. Fig. 20 shows that the estimation error increases only moderately with increasing noise level. This supports that the parameter estimation is reproducible and robust to variances in parameter initialization.

5. Discussion and Conclusion

We extended the Fourier Consistency Conditions (FCC) from fan-beam (Berger et al.; 2014) to cone-beam CT. It is applied directly in the Fourier-space of the projection image stack. Our experiments show that it substantially reduces artifacts due to 3-D translational object motion, with the ability to detect even very high-frequency motion. A clear advantage of our method is that we do not require any temporal assumptions for the motion: for example Yu et al. (2006) limited their estimation to a polynomial model along time resulting in only 6 parameters. In contrast, we were able to successfully estimate $2K - 2 = 1022$ detector shifts independently for each projection, which allows compensation to a great variety of translational motion shapes. An extension to the full set of affine or even non-linear deformations is possible using the more general cost function in Eqn. 2. These deformations would need to be implemented in the spatial domain which prohibits the use of precomputed 2-D FFTs on the projection images and requires a 3-D FFT for each partial derivative and cost function evaluation.

Large improvements in image quality were achieved by the FCC correction method on the FORBILD phantom for all tested motion patterns, with SSIM and RMSE values close to that of the CorrGT reconstructions in most of the cases. In general, the high-frequency motions, i.e., *Oscil*, *Chirp*, and *Rect*, had the lowest SSIM when reconstructing with the ground-truth motion CorrGT. This may correspond to the increased level of streaking artifacts caused by the fast variations within these motion patterns. The visually observed energy within the triangular regions after applying the proposed method is in agreement with the observed residual reconstruction errors, i.e., higher for high-frequency motion and lower for low-frequency motion. Certain motion types, such as *LF-2*, appear to have only limited effect to the triangular regions. Nevertheless, reconstruction results improved also for the *LF-2* motion and the MAD

is comparable to other motion patterns.

The visual impression of the reconstructions is in our case better reflected by SSIM than RMSE. We believe that this is due to the SSIM's sensitivity to the edge alignment of the object which is very indicative for quantifying motion artifacts. The SSIM also allows the verification of the applicability of a plain shifting motion model for different motion cases using the CorrGT reconstructions. The reasons for the reduction in SSIM in case of the CorrGT reconstructions are two-fold. First, the SSIM is influenced by inaccuracies of the approximative motion compensation algorithm (Schaefer et al.; 2006) especially for higher motion amplitudes. Second, a plain shifting motion model can only approximate 3-D translational motion, as the shape of the projection as well as its intensities may vary based on the magnification factor as well as the cone-angle of the CT system.

Our results show that the FCC cost function is robust to even severe Poisson distributed noise artifacts. This can be clearly seen by comparing the results of Figure 10 to the corresponding noise-free results in Figure 6(f-j) and Figure 8(g-l). Further, the relative improvement of SSIM and RMSE in Table 2 and for MAD and SD in Table 3, are not much affected by noise.

The proposed method showed very accurate results when estimating axial motion using the vertical detector direction. This makes it a promising candidate for correction of respiratory motion, where often only vertical detector translations are estimated, e.g., in the field of radiotherapy (Rit et al.; 2013). In addition to the FORBILD cases we could show that FCC are also applicable to more realistic cases. We show that respiratory motion can be accurately tracked when applied to an XCAT-based rotational angiography of the coronary arteries. Additionally, we were able to successfully apply FCC to a knee phantom simulated under weight-bearing conditions, which was based on real patient motion. This also allowed testing of the truncation robust extension of FCC using apodization windows, as introduced in Section 2.6.

A limitation of our method is that it is invariant to constant detector offsets, i.e., a general misalignment of the detector cannot be corrected. For our experiments we decided to alleviate this problem by fixing the motion parameters of the first projection image to the ground-truth motion. This was helpful in multiple ways. First, we made sure that an undefined value does not reduce image quality which allows a fair qualitative and quantitative comparison. Second, adjusting the initial parameters to the ground-truth motion implicitly aligns the reconstructions' coordinate systems for all tested motion cases which is essential for a meaningful SSIM and RMSE computation. However, we are aware that the ground-truth motion is not available in a real scenario. We expect that a moderate offset in v-direction will manifest in a slight increase of cone-beam artifacts and an object shift in z-direction. In contrast a constant shift in u-direction has more severe implications on the reconstruction result. Thus, estimating only one parameter may be sufficient to obtain a reasonable image quality. A straightforward estimation method is to create reconstructions for various u-direction offsets and either manually select the visually most accurate reconstruction

or automatically select the best reconstruction based on image quality measures such as Entropy or Positivity (Rohkohl et al.; 2013).

An important result of our study is the systematic low-frequency deviation of the motion in horizontal detector direction as outlined in Figures 12a and 12b. The key finding is that the erroneously estimated motion in case of the motion-free data can also be found as additive bias in the residual error of the motion estimates for the motion-corrupted data. Figure 17 shows that the low-frequency deviation is often reflected by sinusoidal harmonics, depending on the parameter configuration of the mask. Yet, we were not able to specify its source exactly. One reason could be the mask's discretization, as shown in Section 2.4. We propose a slight increase of the mask size to avoid undefined spectral regions. Our results indicate that this extension is able to stabilize the optimization and reduces the deviation to a much smaller amplitude. Additionally, we were able to show that the same deviations are also present in 2-D fan-beam geometries, proving that they do not originate from approximations during our extension to 3-D cone-beam CT.

For future work we plan an extension of the method to also lateral truncations. Additionally, we will aim to get a better understanding of the sources for the reported deviations in horizontal detector direction, e.g., by a comparison with the parallel beam geometry, where FCC have been initially derived for. The goal is to reduce the residual error for the horizontal detector motion to a similar level than for the vertical direction, yielding very accurate motion estimates and improved image quality.

6. Conclusion

In summary, this work presents a novel approach for motion estimation and compensation in 3-D cone-beam CT. The method is based on FCC that are defined in the Fourier domain of the acquired projection image stack. The method works directly in the projection domain and estimates 2-D detector shifts for each projection image in order to optimize an FCC-based cost function. Even though detector shifts may be limited to simple 3-D motion, the presented method does not rely on temporal assumptions and can therefore be considered for many types of motion. Resulting motion parameters can be easily incorporated during the final backprojection step of the image reconstruction, thus, no iterative or intermediate reconstruction is necessary. We do not require any prior reconstructions and also no additional signals such as externally attached fiducial markers or an ECG. Our extension of FCC to 3-D cone-beam CT showed promising results, when applied to a variety of motion patterns on a challenging numerical phantom. Furthermore, we could successfully apply FCC to a realistic phantom of knees under weight-bearing conditions and for respiratory motion estimation in rotational angiography of the coronaries. Particularly good results were achieved for motion along the vertical detector directions, i.e., collinear to the CT's rotation axis. Even high-frequency motion could be accurately identified and estimated such that we noticed a great reduction of motion-induced streaking artifacts.

Acknowledgments

The authors gratefully acknowledge funding of the Research Training Group 1773 Heterogeneous Image Systems and of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative. Further, we acknowledge funding support from the NIH Shared Instrument Grant S10 RR026714 supporting the zeego@StanfordLab, the NIH Weight-bearing imaging Grant R01 AR065248, and Siemens AT.

References

Berger, M. (2016). *Motion-Corrected Reconstruction in Cone-Beam Computed Tomography for Knees under Weight-Bearing Condition*, PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

Berger, M., Maier, A., Xia, Y., Hornegger, J. and Fahrig, R. (2014). Motion Compensated Fan-Beam CT by Enforcing Fourier Properties of the Sinogram, *in* F. Noo (ed.), *Proceedings of the third international conference on image formation in x-ray computed tomography*, pp. 329–332.

Berger, M., Müller, K., Aichert, A., Unberath, M., Thies, J., Choi, J.-H., Fahrig, R. and Maier, A. (2016). Marker-free motion correction in weight-bearing cone-beam ct of the knee joint, *Medical Physics* **43**(3): 1235–1248.

Brokish, J. and Bresler, Y. (2006). Sampling Requirements for Circular Cone Beam Tomography, *Nuclear Science Symposium Conference Record, 2006. IEEE*, Vol. 5, pp. 2882–2884.

Choi, J.-H., Fahrig, R., Keil, A., Besier, T. F., Pal, S., McWalter, E. J., Beaupré, G. S. and Maier, A. (2013). Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization., *Medical Physics* **40**(9): 091905.

Choi, J.-H., Maier, A., Keil, A., Pal, S., McWalter, E. J., Beaupré, G. S., Gold, G. E. and Fahrig, R. (2014). Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. II. Experiment, *Medical Physics* **41**(6): 061902.

Clackdoyle, R. and Desbat, L. (2013). Full data consistency conditions for cone-beam projections with sources on a plane, *Physics in Medicine and Biology* **58**(23): 8437.

Clackdoyle, R. and Desbat, L. (2015). Data consistency conditions for truncated fanbeam and parallel projections, *Medical Physics* **42**(2): 831–845.

Edholm, P. R., Lewitt, R. M. and Lindholm, B. (1986). Novel Properties Of The Fourier Decomposition Of The Sinogram, *in* T. F. Budinger, Z.-H. Cho and O. Nalcioglu (eds), *Proc. SPIE 0671, Physics and Engineering of Computerized Multidimensional Imaging and Processing*, International Society for Optics and Photonics, pp. 8–18.

Ens, S., Ulrici, J., Hell, E. and Buzug, T. M. (2010). Automatic detection of patient

motion in cone-beam computed tomography, *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, pp. 1257–1260.

Feldkamp, L. A., Davis, L. C. and Kress, J. W. (1984). Practical cone-beam algorithm, *Journal of the Optical Society of America A* **1**(6): 612–619.

Gendrin, C., Furtado, H., Weber, C., Bloch, C., Figl, M., Pawiro, S. A., Bergmann, H., Stock, M., Fichtinger, G., Georg, D. and Birkfellner, W. (2012). Monitoring tumor motion by real time 2D/3D registration during radiotherapy., *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* **102**(2): 274–80.

Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete fourier transform, *Proceedings of the IEEE* **66**(1): 51–83.

Hartley, R. I. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, 2 edn, Cambridge University Press, Cambridge, Great Britain.

Hawkins, W., Leichner, P. and Yang, N.-C. (1988). The circular harmonic transform for spect reconstruction and boundary conditions on the fourier transform of the sinogram, *Medical Imaging, IEEE Transactions on* **7**(2): 135–138.

Helgason, S. (1980). *The Radon Transform*, 1 edn, Birkhauser Boston, Boston, MA, USA.

John, F. (1938). The ultrahyperbolic differential equation with four independent variables, *Duke Mathematical Journal* **4**(2): 300–322.

Lauritsch, G. and Bruder, H. (2001). Forbild head phantom, Website. last checked: 2016/01/21.

Leng, S., Nett, B., Speidel, M. and Chen, G.-H. (2007). Motion artifact reduction in fan-beam and cone-beam computed tomography via the fan-beam data consistency condition (FDCC), *Proc. SPIE*, Vol. 6510, pp. 65101W–65101W–13.

Li, T., Xing, L., Munro, P., McGuinness, C., Chao, M., Yang, Y., Loo, B. and Koong, A. (2006). Four-dimensional cone-beam computed tomography using an on-board imager, *Medical Physics* **33**(10): 3825–3833.

Ludwig, D. (1966). The radon transform on euclidean space, *Communications on Pure and Applied Mathematics* **19**(1): 49–81.

Maier, A. and Berger, M. (2010). Jpop: Java parallel optimization package, Website. https://www5.cs.fau.de/research/software/java-parallel-optimization-package/, last checked: 2017/05/23.

Maier, A., Hofmann, H., Schwemmer, C., Hornegger, J., Keil, A. and Fahrig, R. (2012). Fast Simulation of X-ray Projections of Spline-based Surfaces using an Append Buffer, *Physics in Medicine and Biology* **57**(19): 6193–6210.

Mazin, S. R. and Pelc, N. J. (2010). Fourier properties of the fan-beam sinogram., *Medical physics* **37**(4): 1674–80.

Müller, K., Maier, A., Zheng, Y., Wang, Y., Lauritsch, G., Schwemmer, C., Rohkohl, C., Hornegger, J. and Fahrig, R. (2014). Interventional heart wall motion analysis with cardiac C-arm CT systems, *Physics in Medicine and Biology* **59**(9): 2265–2294.

Natterer, F. (1986). *The mathematics of computerized tomography*, Springer.

Natterer, F. (1993). Sampling in fan beam tomography, *SIAM Journal on Applied Mathematics* **53**(2): 358–380.

Ouadah, S., Stayman, J. W., Gang, G. J., Ehtiati, T. and Siewerdsen, J. H. (2016). Self-calibration of cone-beam CT geometry using 3D-2D image registration, *Physics in Medicine and Biology* **61**(7): 2613–2632.

Patch, S. K. (2002a). Computation of unmeasured third-generation VCT views from measured views, *Medical Imaging, IEEE Transactions on* **21**(7): 801–813.

Patch, S. K. (2002b). Consistency conditions upon 3D CT data and the wave equation, *Physics in Medicine and Biology* **47**(15): 2637.

Rit, S., Sarrut, D. and Desbat, L. (2009). Comparison of analytic and algebraic methods for motion-compensated cone-beam ct reconstruction of the thorax, *IEEE Transactions on Medical Imaging* **28**(10): 1513–1525.

Rit, S., Sarrut, D. and Sonke, J.-J. (2013). Respiratory Motion Correction in Cone-Beam CT for Image-Guided Radiotherapy, *in* J. Ehrhardt and C. Lorenz (eds), *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*, Biological and Medical Physics, Biomedical Engineering, Springer Berlin Heidelberg, pp. 319–334.

Rohkohl, C., Bruder, H., Stierstorfer, K. and Flohr, T. (2013). Improving best-phase image quality in cardiac CT by motion correction with MAM optimization., *Medical physics* **40**(3): 031901.

Schaefer, D., Borgert, J., Rasche, V. and Grass, M. (2006). Motion-compensated and gated cone beam filtered back-projection for 3-D rotational X-ray angiography, *Medical Imaging, IEEE Transactions on* **25**(7): 898–906.

Schwemmer, C., Rohkohl, C., Lauritsch, G., Müller, K. and Hornegger, J. (2013). Residual motion compensation in ECG-gated interventional cardiac vasculature reconstruction, *Physics in Medicine and Biology* **58**(11): 3717.

Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J. and Tsui, B. M. W. (2010). 4D XCAT phantom for multimodality imaging research, *Medical Physics* **37**(9): 4902–4915.

Shirato, H., Harada, T., Harabayashi, T., Hida, K., Endo, H., Kitamura, K., Onimaru, R., Yamazaki, K., Kurauchi, N., Shimizu, T., Shinohara, N., Matsushita, M., Dosaka-Akita, H. and Miyasaka, K. (2003). Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy, *International Journal of Radiation Oncology*Biology*Physics* **56**(1): 240–247.

Unberath, M., Aichert, A., Achenbach, S. and Maier, A. (2016). Virtual Single-frame

Subtraction Imaging, *in* M. Kachelrieß (ed.), *Proceedings of the 4th CT Meeting*, pp. 89–92.

Unberath, M., Berger, M., Aichert, A. and Maier, A. (2017). Fourier Consistency-Based Motion Estimation in Rotational Angiography, *in* DKFZ (ed.), *Proc. Bildverarbeitung für die Medizin (BVM) 2017*, Heidelberg, pp. 110–116.

Verrill, S. (2005). An unconstrained nonlinear optimization solver: A users's guide, Website. https://www1.fpl.fs.fed.us/uncmin_userguide.pdf, last checked: 2017/05/23.

Wang, Z., Bovik, A., Sheikh, H. and Simoncelli, E. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity, *IEEE Transactions on Image Processing* **13**(4): 600–612.

Welch, A., Campbell, C., Clackdoyle, R., Natterer, F., Hudson, M., Bromiley, A., Mikecz, P., Chillcot, F., Dodd, M., Hopwood, P., Craib, S., Gullberg, G. T. and Sharp, P. (1998). Attenuation correction in PET using consistency information, *Nuclear Science, IEEE Transactions on* **45**(6): 3134–3141.

Yu, H. and Wang, G. (2007). Data Consistency Based Rigid Motion Artifact Reduction in Fan-Beam CT, *Medical Imaging, IEEE Transactions on* **26**(2): 249–260.

Yu, H., Wei, Y., Hsieh, J. and Wang, G. (2006). Data consistency based translational motion artifact reduction in fan-beam CT, *IEEE Transactions on Medical Imaging* **25**(6): 792–803.

Zeng, R., Fessler, J. A. and Balter, J. M. (2005). Respiratory motion estimation from slowly rotating x-ray projections: Theory and simulation, *Medical Physics* **32**(4): 984.