

Assessing Cardiac Function from Total-Variation-Regularized 4-D C-arm CT in the Presence of Angular Undersampling

O Taubmann^{1,2}, V Haase³, G Lauritsch⁴, Y Zheng⁵, G Krings⁶,
J Hornegger^{1,2}, A Maier^{1,2}

¹ Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

² Erlangen Graduate School in Advanced Optical Technologies (SAOT), Germany

³ Institute for Signal Processing, University of Luebeck, Germany

⁴ Siemens Healthcare GmbH, Forchheim, Germany

⁵ Siemens Medical Solutions Inc., Princeton, NJ, USA

⁶ University Medical Center Utrecht, Netherlands

E-mail: `oliver.taubmann@fau.de`

Abstract.

Time-resolved tomographic cardiac imaging using an angiographic C-arm device may support clinicians during minimally invasive therapy by enabling a thorough analysis of the heart function directly in the catheter laboratory. However, clinically feasible acquisition protocols entail a highly challenging reconstruction problem which suffers from sparse angular sampling of the trajectory. Compressed sensing theory promises that useful images can be recovered despite massive undersampling by means of sparsity-based regularization. For a multitude of reasons—most notably the desired reduction of scan time, dose and contrast agent required—it is of great interest to know just how little data is actually sufficient for a certain task.

In this work, we apply a convex optimization approach based on primal-dual splitting to 4-D cardiac C-arm computed tomography. We examine how the quality of spatially and temporally total-variation-regularized reconstruction degrades when using as few as 6.9 ± 1.2 projection views per heart phase. First, feasible regularization weights are determined in a numerical phantom study, demonstrating the individual benefits of both regularizers. Secondly, a task-based evaluation is performed in eight clinical patients. Semi-automatic segmentation-based volume measurements of the left ventricular blood pool performed on strongly undersampled images show a correlation of close to 99 % with measurements obtained from less sparsely sampled data.

Keywords: C-arm Computed Tomography, Angular Undersampling, Cardiac Function, 4-D Imaging, Total Variation, Temporal Regularization. Submitted to: *Phys. Med. Biol.*

1. Introduction

Time-resolved volumetric imaging of the heart during minimally invasive interventions has the potential to offer clinicians a comprehensive overview of cardiac function without the need to relocate the patient to a diagnostic scanner. X-ray projections acquired with interventional C-arm systems allow for 3-D imaging by means of rotational angiography (Hetterich et al. 2010). If an electrocardiogram (ECG) is recorded simultaneously, retrospective multi-segment gating enables the 3-D reconstruction of multiple heart phases, i. e. 4-D imaging.

Early approaches used as many as four rotations to collect a sufficient amount of data to deal with angular undersampling caused by the gating process (Lauritsch et al. 2006). However, this leads to a long scan time which entails drawbacks such as the need for prolonged contrast agent injection and breath-hold to avoid respiratory motion artifacts. Therefore, sophisticated few-view reconstruction techniques have been employed in the past to obtain 4-D images from a single sweep.

One such method, originally proposed for this application by (Müller, Maier, Schwemmer, Lauritsch, Buck, Wielandts, Hornegger & Fahrig 2014), performs non-rigid motion estimation and subsequent motion compensation (Schäfer et al. 2006). Although highly effective, its reliance on the ability to robustly estimate motion from preliminary, artifact-degraded gated images inherently limits the amount of possible undersampling. This limitation was substantially softened by later refinements in (Taubmann et al. 2015) and (Taubmann, Maier, Hornegger, Lauritsch & Fahrig 2016). Similarly, (Wielandts et al. 2014, Wielandts et al. 2015) employed streak reduction and registration-based filtering, which is likewise based on deformable image registration using a B-spline model. A potential downside of such algorithms is that the use of smooth motion models causes remaining artifacts in the preliminary images to be interpreted as deformations.

This issue is avoided by another wide-spread approach to deal with undersampled data known as compressed sensing. For dynamic cardiac C-arm computed tomography (CT), its use was first suggested by (Mory et al. 2014) in the form of an algebraic reconstruction technique with spatial and temporal total variation (TV) regularization. Temporal regularization in particular is very useful, serving a similar purpose as motion compensation as it facilitates the exchange of information between the individual gated projection subsets, thereby utilizing the whole scan. (Taubmann, Lauritsch, Krings & Maier 2016) compared different temporal regularizers and confirmed the benefits of temporal TV (tTV) minimization. Previously, (Haase et al. 2016) had adopted the iTV scheme by (Ritschl et al. 2011, Ritschl et al. 2012) in the same framework. All of these studies used conventional gradient descent for TV minimization. Recently, convex optimization schemes involving variable splitting and proximal operators have gained increasing popularity in the field due to their desirable convergence properties. While (Mory & Jacques 2014) reformulated their earlier work as a convex optimization algorithm in accordance with (Chambolle & Pock 2011), they did not show any results obtained using such an implementation.

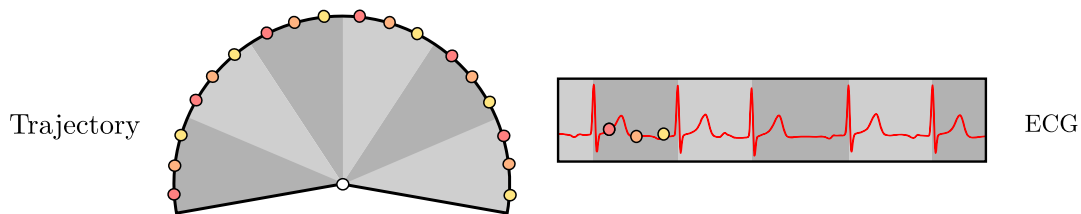


Figure 1. Strict ECG gating uses one view per cardiac phase from each heart cycle.

In this work, we apply a convex optimization approach based on primal-dual splitting to 4-D cardiac C-arm CT. We examine how the quality of spatially and temporally TV-regularized reconstruction degrades when using *very* few projection views (≈ 7 per phase). When only few views are required, the clinical image acquisition protocol becomes more flexible, offering the possibility to reduce scan time, dose and contrast agent burden. After demonstration in a numerical phantom, a task-based evaluation is performed in eight clinical patients. It focuses on the ability to quantify and visualize cardiac function using a semi-automatic model-based segmentation of the left ventricle (LV).

2. Methods

2.1. Strict Electrocardiogram Gating

During C-arm rotation of a few seconds duration, the heart completes several full cycles (shaded areas in Fig. 1). In each cycle, several heart phases (circles in Fig. 1) can be defined based on the relative distance to the two neighboring R-peaks in the ECG. The number of views available for each phase is determined by the number of heart cycles available. To ensure consistency, strict gating chooses exactly one image from each heart cycle, namely the one which is closest in terms of the relative heart phase. This implies that some images may be left out if they are not the best fit for any of the desired phases. Conversely, an image may be used multiple times if it is the best fit for more than one phase.

2.2. Spatially and Temporally TV-regularized Reconstruction

Due to ECG gating, only few projection views are available. Therefore, we formulate 4-D reconstruction as a regularized optimization problem. Below, we define a suitable objective function and describe how to optimize it by means of a proximal algorithm.

2.2.1. Objective Function Let $\mathbf{i} \in \mathbb{R}^{N_{\text{ph}} \cdot N_{\text{vox}}}$ denote the vector obtained by stacking the voxels of all volume images to be reconstructed, with N_{ph} and N_{vox} the numbers of cardiac phases and voxels per volume, respectively. We call the vectorized representation of the measured line integral data, selected and sorted by the gating process, $\mathbf{p} \in$

$\mathbb{R}^{N_{\text{proj}} \cdot N_{\text{pix}}}$, with N_{proj} the number of projection images and N_{pix} the number of detector pixels. The geometry of the tomography problem is described by the system matrix $\mathbf{A} \in \mathbb{R}^{(N_{\text{proj}} \cdot N_{\text{pix}}) \times (N_{\text{ph}} \cdot N_{\text{vox}})}$, which represents the X-ray projection operator associating the 4-D image \mathbf{i} with the measured data \mathbf{p} .

As the data fidelity term, we choose the squared residual $r(\mathbf{i}) = \frac{1}{2} \|\mathbf{A}\mathbf{i} - \mathbf{p}\|_2^2$. The spatial TV norm $\|\mathbf{i}\|_{\text{sTV}} = \|\mathbf{D}_s \mathbf{i}\|_{1,2}$ is defined as the sum (L^1 -norm) of the magnitudes (L^2 -norm) of the spatial image gradient, calculated using the element-wise spatial forward difference operator \mathbf{D}_s ,

$$\mathbf{D}_s : \mathbf{i}_{x,y,z,t} \mapsto \begin{bmatrix} i_{x+1,y,z,t} - i_{x,y,z,t} \\ i_{x,y+1,z,t} - i_{x,y,z,t} \\ i_{x,y,z+1,t} - i_{x,y,z,t} \end{bmatrix}, \quad (1)$$

with zero boundary conditions. $i_{x,y,z,t}$ denotes the element of \mathbf{i} corresponding to the voxel at the spatial indices (x, y, z) into the reconstruction grid of the cardiac phase with index t . Similarly, the temporal TV norm $\|\mathbf{i}\|_{\text{tTV}} = \|\mathbf{D}_t \mathbf{i}\|_1$ is based on the element-wise forward difference $\mathbf{D}_t : i_{x,y,z,t} \mapsto i_{x,y,z,t+1} - i_{x,y,z,t}$ along the temporal dimension, applied with periodic boundary conditions due to the cyclic nature of cardiac motion. The full reconstruction problem with $\lambda_s, \lambda_t \geq 0$ as the regularizer weights then reads,

$$\arg \min_{\mathbf{i}} r(\mathbf{i}) + \lambda_s \cdot \|\mathbf{i}\|_{\text{sTV}} + \lambda_t \cdot \|\mathbf{i}\|_{\text{tTV}} + \iota_{\mathbb{R}_+}(\mathbf{i}), \quad (2)$$

where $\iota_{\mathbb{R}_+}$ is the characteristic function of \mathbb{R}_+ , enforcing non-negativity on the solution.

2.2.2. Convex Primal-Dual Optimization Popular optimization schemes for such problems are based on primal-dual splitting and proximal operators (Chambolle & Pock 2011), specifically soft-thresholding in the case of L^1 minimization. Our objective function consists of both a smooth, i. e. differentiable part $r(\mathbf{i})$ as well as non-smooth yet proximable parts, albeit in linearly transformed domains in case of the L^1 components. This type of function lends itself well for optimization using the proximal splitting algorithm derived in (Condat 2013) and later applied by the same author to TV-regularized inverse imaging problems (Condat 2014). Due to its full-splitting approach, it relies on “simple” operations only and requires no nested loops, i. e. avoids the need to approximate the TV proximal operator, which has no closed-form solution, within each iteration. Adapted to our task with a trivial relaxation factor of 1, one iteration of *Algorithm 1* in (Condat 2014) simplifies to applying the updates given in Algorithm 1 in our paper. For a comprehensive derivation of the algorithm as well as a proof of convergence, we would like to refer the reader to (Condat 2013). A practical description of the steps involved is given in the following paragraph.

Through alternating updates, the algorithm iteratively finds the optimal values of three variables: $\mathbf{i} \in \mathbb{R}^{N_{\text{ph}} \cdot N_{\text{vox}}}$, which converges to the desired image solution (primal domain), as well as $\mathbf{g}_s \in \mathbb{R}^{3 \cdot N_{\text{ph}} \cdot N_{\text{vox}}}$ and $\mathbf{g}_t \in \mathbb{R}^{N_{\text{ph}} \cdot N_{\text{vox}}}$, which correspond to the spatial and temporal image gradient of \mathbf{i} , respectively (dual domain). \mathbf{g}_s is three times as large as \mathbf{i} as it holds a 3-vector for each image voxel. At the start, we initialize the estimates

Algorithm 1 Updates of the proximal reconstruction algorithm.

- 1: $\mathbf{i}_{\text{prev}} \leftarrow \mathbf{i}$
 - 2: $\mathbf{i} \leftarrow (\mathbf{i} - \tau(\nabla_{\mathbf{i}} r(\mathbf{i}) + \mathbf{D}_s^\top \mathbf{g}_s + \mathbf{D}_t^\top \mathbf{g}_t))_+$
 - 3: $\mathbf{g}_s \leftarrow \text{prox}_{\sigma(\lambda_s \|\cdot\|_{1,2})^*}(\mathbf{g}_s + \sigma \mathbf{D}_s(2\mathbf{i} - \mathbf{i}_{\text{prev}}))$
 - 4: $\mathbf{g}_t \leftarrow \text{prox}_{\sigma(\lambda_t \|\cdot\|_1)^*}(\mathbf{g}_t + \sigma \mathbf{D}_t(2\mathbf{i} - \mathbf{i}_{\text{prev}}))$
-

for \mathbf{i} , \mathbf{g}_s and \mathbf{g}_t with zeros. In the first line of Algorithm 1, the primal estimate \mathbf{i} of the previous iteration is temporarily stored. In the second line, a gradient descent step is performed by subtracting $\nabla_{\mathbf{i}} r(\mathbf{i}) = \mathbf{A}^\top(\mathbf{A}\mathbf{i} - \mathbf{p})$, which amounts to a back-projection of the residual error $\mathbf{A}\mathbf{i} - \mathbf{p}$, scaled by the primal step length τ . The current dual estimates \mathbf{g}_s and \mathbf{g}_t are mapped back to the primal domain by the adjoints of the linear operators and are used to modify the descent direction. The adjoint \mathbf{D}_\diamond^\top of a (forward difference) gradient operator \mathbf{D}_\diamond is given as the (backward difference) divergence operator negated. Subsequently, non-negativity is enforced element-wise on the image by the proximal operator of $\iota_{\mathbb{R}_+}$, $(\cdot)_+ : i \mapsto \max\{i, 0\}$. In the third line, the first dual estimate \mathbf{g}_s is updated by adding an extrapolation of the latest primal estimates, brought to the gradient domain by \mathbf{D}_s and scaled with the dual step length σ . The extrapolation approximates a so-called extragradient step in order to improve convergence (Chambolle & Pock 2011). This is followed by the evaluation of the proximal operator of the convex conjugate (Fenchel-Rockafellar conjugate) of the scaled norm $\lambda_s \|\cdot\|_{1,2}$ associated with \mathbf{D}_s in Eq. 2,

$$\text{prox}_{\sigma(\lambda_s \|\cdot\|_{1,2})^*} : \mathbf{v} \mapsto \begin{cases} \mathbf{v} / \max\{\|\mathbf{v}\|_2 / \lambda_s, 1\} & \text{if } \lambda_s > 0, \\ \mathbf{0} & \text{otherwise,} \end{cases} \quad (3)$$

which can be understood as the projection of each component vector \mathbf{v} in \mathbf{g}_s onto a ball with radius λ_s . Note that the proximal operator of the convex conjugate f^* of a function f can be derived from the proximal operator of f by the Moreau identity (Condat 2014),

$$\text{prox}_{\sigma f^*}(x) = x - \sigma \text{prox}_{f/\sigma}(x/\sigma), \text{ for } \sigma > 0. \quad (4)$$

In the fourth line, the same steps are performed analogously for the second dual estimate \mathbf{g}_t , with the corresponding proximal operator,

$$\text{prox}_{\sigma(\lambda_t \|\cdot\|_1)^*} : g \mapsto \begin{cases} g / \max\{|g|/\lambda_t, 1\} & \text{if } \lambda_t > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The primal and dual step lengths τ and σ are chosen to satisfy the convergence criterion $\tau \left(\frac{\beta}{2} + \sigma \|\mathbf{D}_s^\top \mathbf{D}_s + \mathbf{D}_t^\top \mathbf{D}_t\| \right) < 1$, where β denotes the Lipschitz constant of $\nabla_{\mathbf{i}} r(\mathbf{i})$ and $\|\cdot\|$ the operator norm. For our task, an ordered subset scheme replacing the $\nabla_{\mathbf{i}} r(\mathbf{i})$ update, as it is often employed to improve convergence speed, is not crucial as the gating subsets already possess favorable angular distributions (cf. Fig. 1).

2.3. Cardiac Function Analysis

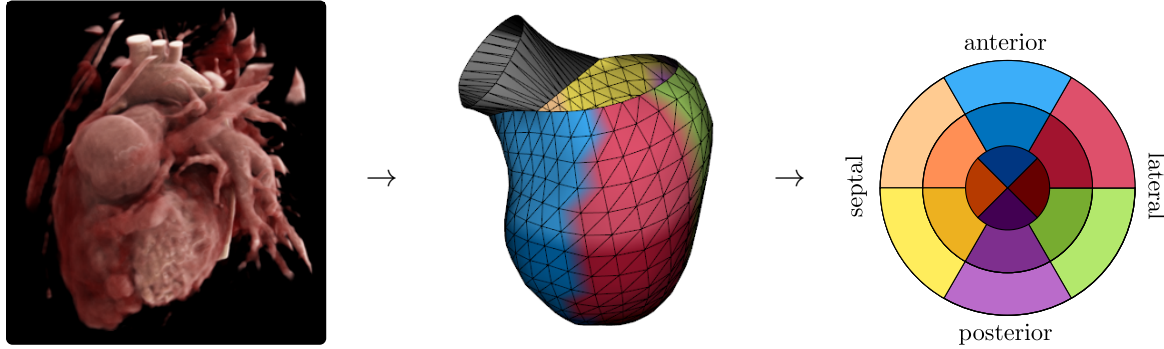


Figure 2. From the reconstructed image (left), the LV endocardium surface is segmented (center) and projected onto a polar map divided into 16 segments as suggested by the American Heart Association (right).

2.3.1. Segmentation Segmentation is performed by fitting a shape model of the LV endocardium to the reconstructed images (Fig. 2). The model consists of 544 vertices arranged in 17 rings distributed along the LV long axis plus a single vertex at the apex. Fitting is a two-step procedure: First, the pose is determined, i.e. a bounding box defining the position, orientation and size of the object. This is done manually as the data available in our study is not sufficient for training an automatic pose detector. The pose estimate chosen for each case is consistent over all volumes—regardless of cardiac phase and undersampling setting—to rule out the influence of manual variability. Secondly, the model mean shape is transformed according to the pose and subsequently deformed by an automatic boundary detector such that it fits the edges present in the image (Zheng et al. 2008). The boundary detector is trained in a leave-one-case-out fashion to ensure the separation of training and test data while still maintaining a high detection quality. As a post-processing step, the surfaces are refined as described in (Zheng et al. 2009) by maximizing their smoothness under the constraint that the blood pool be tightly enclosed.

2.3.2. Volumetry and Motion Visualization From the segmented surfaces, the blood pool volumes in end-diastole and end-systole are calculated. In addition, similar to previous work of (Müller, Maier, Zheng, Wang, Lauritsch, Schwemmer, Rohkohl, Hornegger & Fahrig 2014), a visualization of the contraction pattern is provided by color-coding the displacement of each vertex between end-systole and end-diastole and projecting the resulting mesh onto a polar map according to the 16-segment model suggested by the American Heart Association (Cerqueira et al. 2002) (Fig. 2, right).

2.4. Experiments

2.4.1. Phantom Study The purpose of our phantom study is twofold. First, to demonstrate the achieved image quality of the proposed reconstruction algorithm both visually and quantitatively in a controlled environment. Secondly, to serve as a

reproducible way of determining feasible regularizer weights λ_s , λ_t . Note that while these may not be ideal for clinical data, in order to offer a fair evaluation, we refrain from adapting the parameters to each case individually as this is typically impractical in a realistic scenario as well.

We automatically normalize for differences in the acquisition or reconstruction settings which would alter the relative strength of regularization,

$$\lambda_s = \frac{N_{\text{proj}} \cdot \text{voxel-size}^2}{\text{detector-pixel-size}^2 \cdot N_{\text{ph}}} \cdot \lambda_s^*, \quad \lambda_t = \frac{N_{\text{proj}} \cdot \text{voxel-size}^3}{\text{detector-pixel-size}^2} \cdot \lambda_t^*. \quad (6)$$

Feasible parameters are determined in a small grid search, $(\lambda_s^*, \lambda_t^*)^\top \in \mathbb{G} \times \mathbb{G}$, $\mathbb{G} = \{0, 2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}\} \cdot 10^{-2}$. We choose the optimal parameters $\lambda_{s,\text{opt}}^*$, $\lambda_{t,\text{opt}}^*$ such that the result exhibits the lowest root-mean-square error (RMSE) compared to a reference image reconstructed from static and fully-sampled projection data.

Our data set is based on the XCAT phantom (Segars et al. 2010). Projection images are generated with a polychromatic spectrum using the CONRAD software (Maier et al. 2012, Maier et al. 2013). The spectrum is discretized in energy bins 5 keV wide from 10 keV to 90 keV (peak energy), with a time-current product of 2.5 mAs per X-ray pulse. For bones and bone marrow, material properties match the mass attenuation coefficients found in the NIST X-ray table[‡]. Contrasted blood in the LV and aorta is simulated as iodine-based contrast agent (Ultravist-150, Bayer AG, Leverkusen, Germany) mixed with equal parts water. All other structures are modeled with the absorption behavior of water for modified densities. The simulated acquisition protocol and reconstruction parameters are identical to those in the clinical data sets described below (section 2.4.2). Several projection data sets are generated: A static version in end-diastole for reconstruction of a reference image as mentioned above and dynamic versions with $N_b = 12, 10, 8, 6$ and 4 heart beats to examine the effect of a gradual reduction of the number of views. Assuming a scan time of t seconds, they correspond to heart rates of $\frac{N_b}{t} \cdot 60$ bpm. We reconstruct $N_{\text{ph}} = 8$ cardiac phases, resulting in a total of $(N_{\text{ph}} \cdot N_b)$ projections that are used per data set. No respiratory motion was simulated.

2.4.2. Experiments on Clinical Data In addition to demonstrating that useful images can also be obtained in practice, the purpose of our experiments on clinical patient data is to quantify errors and potential bias when assessing left ventricular function in the reduced setting. The patients suffer from congenital heart disease and undergo catheter-based procedures related to malformations of their cardiovascular system. The scans we use are routinely acquired to obtain 3-D reconstructions for intraprocedural guidance; retrospectively, 4-D reconstruction of the same data is performed offline for research purposes only. Hence, treatment of the patients is not influenced in any way by our study. Eight data sets are acquired with an Artis zee biplane (Siemens Healthcare GmbH, Forchheim, Germany). At approx. 30 Hz, 133 projection images are captured

[‡] <http://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html>

with an angular increment of 1.5° during a 5 s rotation of the C-arm. This results in a short scan with 200° coverage. The isotropic pixel resolution is 0.31 mm/pixel (0.21 mm/pixel in isocenter), the detector size $N_{\text{pix}} = 960^2$ pixels. All images are iteratively reconstructed on a grid of $N_{\text{vox}} = 256^3$ voxels covering a volume of $(25.6 \text{ cm})^3$ over $N_{\text{ph}} = 8$ heart phases. Optimization is run for 200 iterations. In all cases, contrast agent is injected selectively via a catheter to ensure visibility of the LV. In some cases, other parts of the heart or the vascular system, which we do not evaluate, are contrasted as well. A catheter containing electrodes is inserted into one of the cardiac chambers. This pacing catheter acts as a temporary pacemaker lead that electrically excites cardiac tissue, thereby triggering its contraction, in order to control the heart rate, which is known as electrophysiological pacing. It is used to artificially increase the frequency of the heart beat during the scan. The pacing rate and number of projections is listed in Tab. 1 for each data set. All patients are anesthetized during image acquisition. Complete apnea is induced to avoid any respiratory motion.

To simulate a “reduced” setting in the clinical data, we retrospectively leave out the projection images of every second heart beat in the acquisition, thus reducing the number of views by half (Fig. 3, left). This reduction is artificial and does not correspond directly to situations of reduced heart rate or scan time. Overall, the remaining images are no longer distributed equidistantly along the trajectory as retained and skipped angular sections alternate. However, considering any individual phase, the images it is reconstructed from are still distributed equidistantly, now with an angular step that is twice as large as it was originally (arrows in Fig. 3). Please note that in this regard, i. e. the angular distribution *per phase*, this simulated setup is highly similar to a (hypothetical) acquisition at half heart rate (Fig. 3, right). In such an acquisition, for the temporal resolution w.r.t. one heart cycle to remain constant, half of the original frame rate would be sufficient: The same number of frames would be acquired per cycle as each cycle is twice as long. Leaving all other parameters unchanged, this approach thus also simulates a dose reduction by about 50% as dose scales linearly with the number of images. This is the case since the emission of radiation is pulsed, i. e. the X-ray tube is not emitting continuously during the scan. Note that this is an idealized calculation based on a simplified physiological model where the heart cycle is scaled uniformly in time with varying heart rate, which is not entirely realistic.

2.5. Implementation

Reconstruction was implemented in C++ with a focus on both runtime and flexible object-oriented design, allowing for easy modification of data fidelity terms and arbitrary combinations of norms and linear transforms for regularization. All relevant components were optimized for parallel computation on a graphics processing unit (GPU) using CUDA. On a machine equipped with 32 GB RAM, an Intel Xeon E5-1620 v2 CPU (Intel Corporation, Santa Clara, CA) and an Nvidia Quadro K4000 GPU (Nvidia Corporation, Santa Clara, CA), 4-D reconstruction is completed in about half an hour (less than 10 s

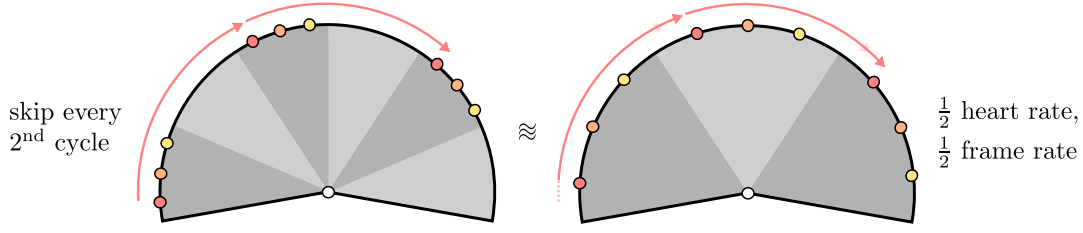


Figure 3. Skipping the data associated with every second heart beat allows an artificial simulation of fewer views. The angular configuration for each individual cardiac phase is somewhat similar to a (hypothetical) scan at half heart rate.

per iteration). Detection and refinement of the LV boundary take 2.8 s per case.

3. Results

3.1. Phantom Study

The results of the grid search are summarized in Fig. 4 and Tab. 2. Both spatial and temporal TV are effective in reducing the error compared to the non-regularized result. Notably, the quantitative improvement is more pronounced for temporal TV. The parameterization $(\lambda_{s,\text{opt}}^*, \lambda_{t,\text{opt}}^*) = (2^{-4}, 2^{-3}) \cdot 10^{-2}$, i.e. a combination of both regularizers, produces an image with few artifacts as well as a fairly accurate representation of the anatomy (Fig. 4, top left) and achieves the lowest overall RMSE (59.7 HU). It is therefore used in all other reconstructions. In the over-regularized regime, the error slowly increases again, albeit not as quickly since only a small fraction of all voxels is affected by the loss of resolution that becomes apparent in

Table 1. Pacing rates and numbers of projections for the clinical patient data sets. For each patient, a total of 133 projections are acquired, from which we reconstruct 8 cardiac phases. In the two rightmost columns, the total number of distinct projections is listed, i.e. projections used in more than one phase are only counted once.

Data set	Heart rate in bpm	Projections per phase		Total unique projections	
		(original)	(reduced)	(original)	(reduced)
Patient 1	240	19	9	125	62
Patient 2	204	17	8	123	60
Patient 3	180	14	7	110	55
Patient 4	168	12	6	95	47
Patient 5	140	12	6	93	45
Patient 6	204	15	7	114	54
Patient 7	180	14	7	109	54
Patient 8	150	11	5	88	40
mean \pm std. dev.	183.3 \pm 32.3	14.3 \pm 2.7	6.9 \pm 1.2	107.1 \pm 13.9	52.1 \pm 7.5

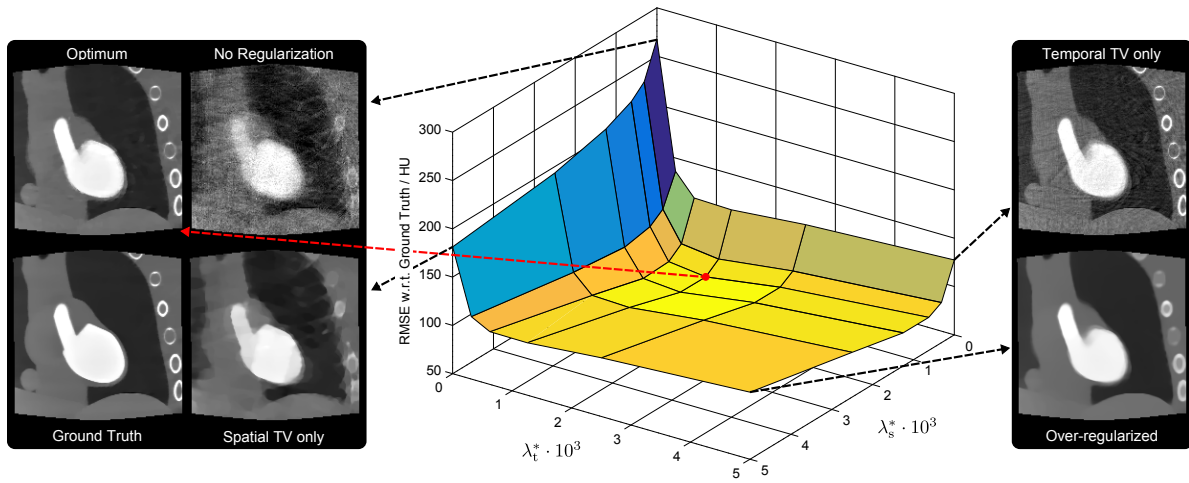


Figure 4. Results of the grid search for optimal regularizer weights. For reference, some non-optimal settings are also displayed, e.g. no or partial regularization only. The same grayscale window is used in all shown image slices. The numeric error values are listed in Tab. 2 for reference.

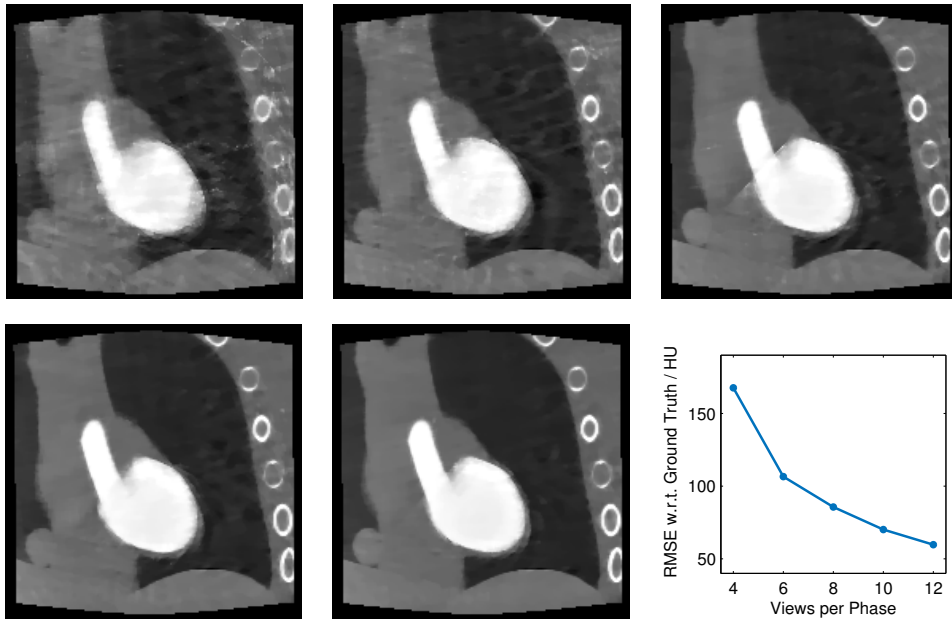


Figure 5. Images reconstructed from the phantom data (top: 4, 6 and 8 views per phase, bottom: 10 and 12 views per phase). The grayscale window is [center, width] = [300, 2600] HU. The plot shows the corresponding root-mean-square errors (RMSE).

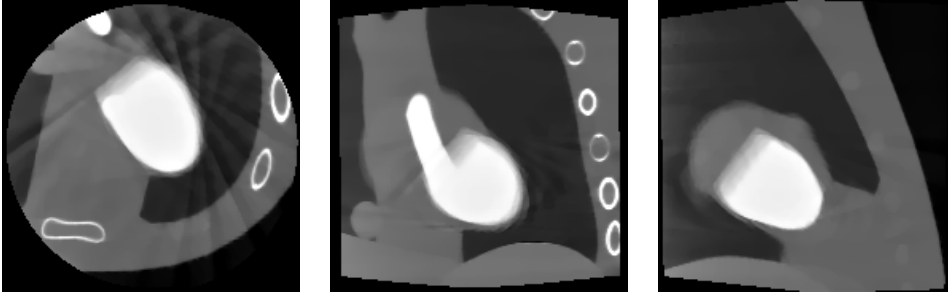


Figure 6. Images obtained by ungated reconstruction of the phantom data set comprising 12 heart cycles from all 133 projection images (axial, coronal and sagittal slices). The grayscale window is [center, width] = [300, 2600] HU.

the corresponding image (Fig. 4, bottom right).

The effect of reducing the number of heart beats and, therefore, the views per phase is illustrated in Fig. 5. Increasingly strong artifacts can be observed, yet most anatomical structures appear reasonably well preserved even when only six views per phase are available. The reconstruction error rises disproportionately, e. g. by 61.2 HU from six to four views, in contrast to an increase of only 46.8 HU from twelve to six views (cf. Fig. 5, bottom right).

For reference, Fig. 6 shows the result obtained by ungated reconstruction from all 133 projection images of the phantom data set comprising twelve heart beats. While static regions above and below the heart are reconstructed well due to sufficient data, inconsistent motion states lead to strong blur and artifacts in the anatomic region of interest, which is also reflected in the RMSE (82.0 HU).

3.2. Clinical Data

In the clinical data shown in Fig. 7, image quality is degraded in the reduced setting as some anatomic structures become blurred and additional streak artifacts are introduced

Table 2. Root-mean-square errors (RMSE) to the ground truth in HU for all tested combinations of regularizer weights. The best value achieved is highlighted. Selected visual results are shown in Fig. 4.

	$\lambda_t^* = 0$	$\lambda_t^* = \frac{2^{-5}}{10^2}$	$\lambda_t^* = \frac{2^{-4}}{10^2}$	$\lambda_t^* = \frac{2^{-3}}{10^2}$	$\lambda_t^* = \frac{2^{-2}}{10^2}$	$\lambda_t^* = \frac{2^{-1}}{10^2}$
$\lambda_s^* = 0$	267.6	136.3	114.2	113.0	119.7	127.5
$\lambda_s^* = \frac{2^{-5}}{10^2}$	231.3	105.2	74.9	70.1	80.0	91.5
$\lambda_s^* = \frac{2^{-4}}{10^2}$	219.1	95.1	64.3	59.7	71.2	84.1
$\lambda_s^* = \frac{2^{-3}}{10^2}$	206.7	89.9	63.1	59.9	71.3	84.3
$\lambda_s^* = \frac{2^{-2}}{10^2}$	193.4	96.1	77.8	76.2	85.3	96.3
$\lambda_s^* = \frac{2^{-1}}{10^2}$	181.3	114.6	104.6	104.2	110.8	119.4

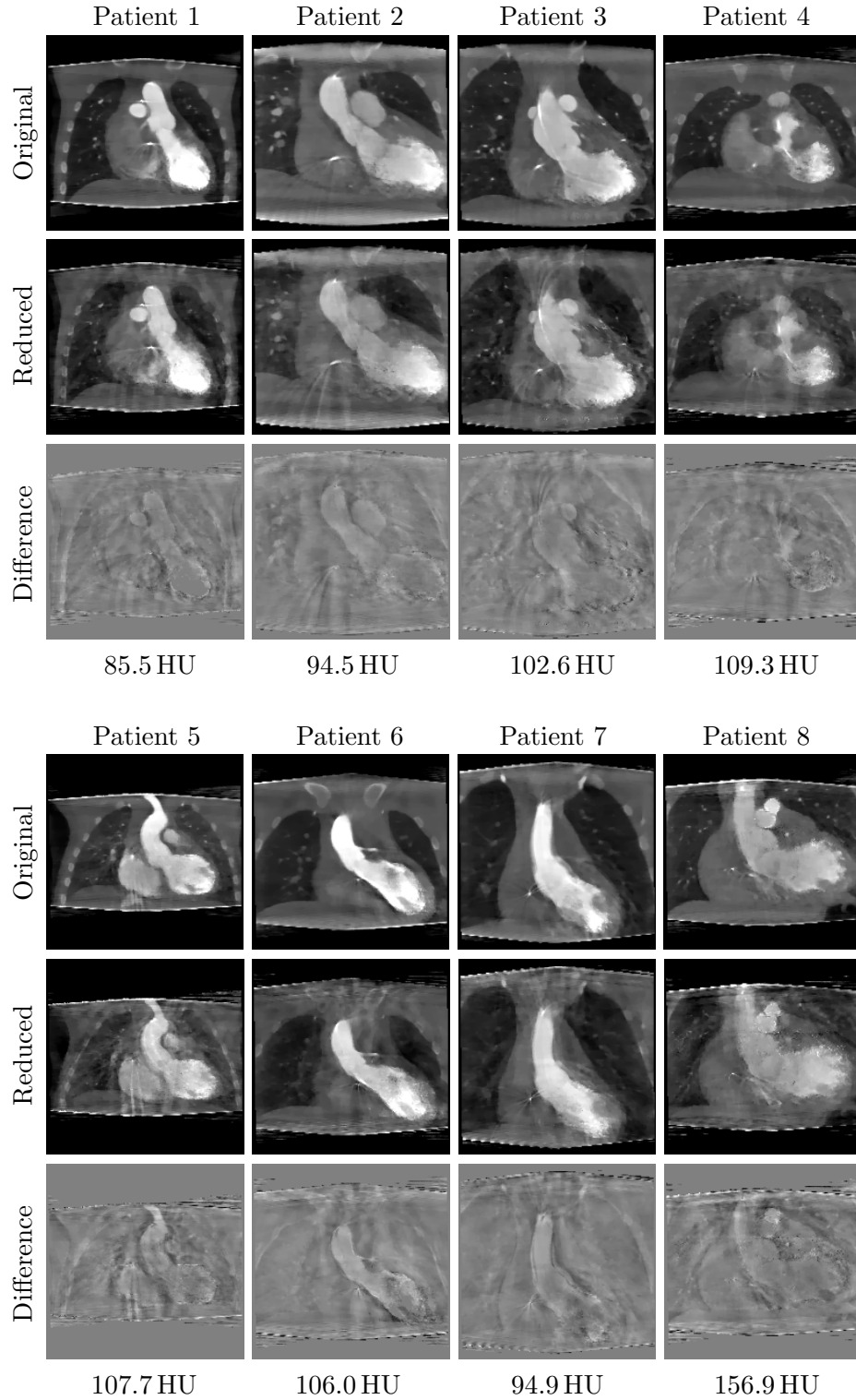


Figure 7. Images reconstructed from the original and reduced data at end-diastole. The grayscale window is $[\text{center}, \text{width}] = [740, 3480]$ HU for the original and reduced images, and $[\text{center}, \text{width}] = [0, 3480]$ HU for the difference images. Below the images, the root-mean-square difference of the volumes in HU is given.

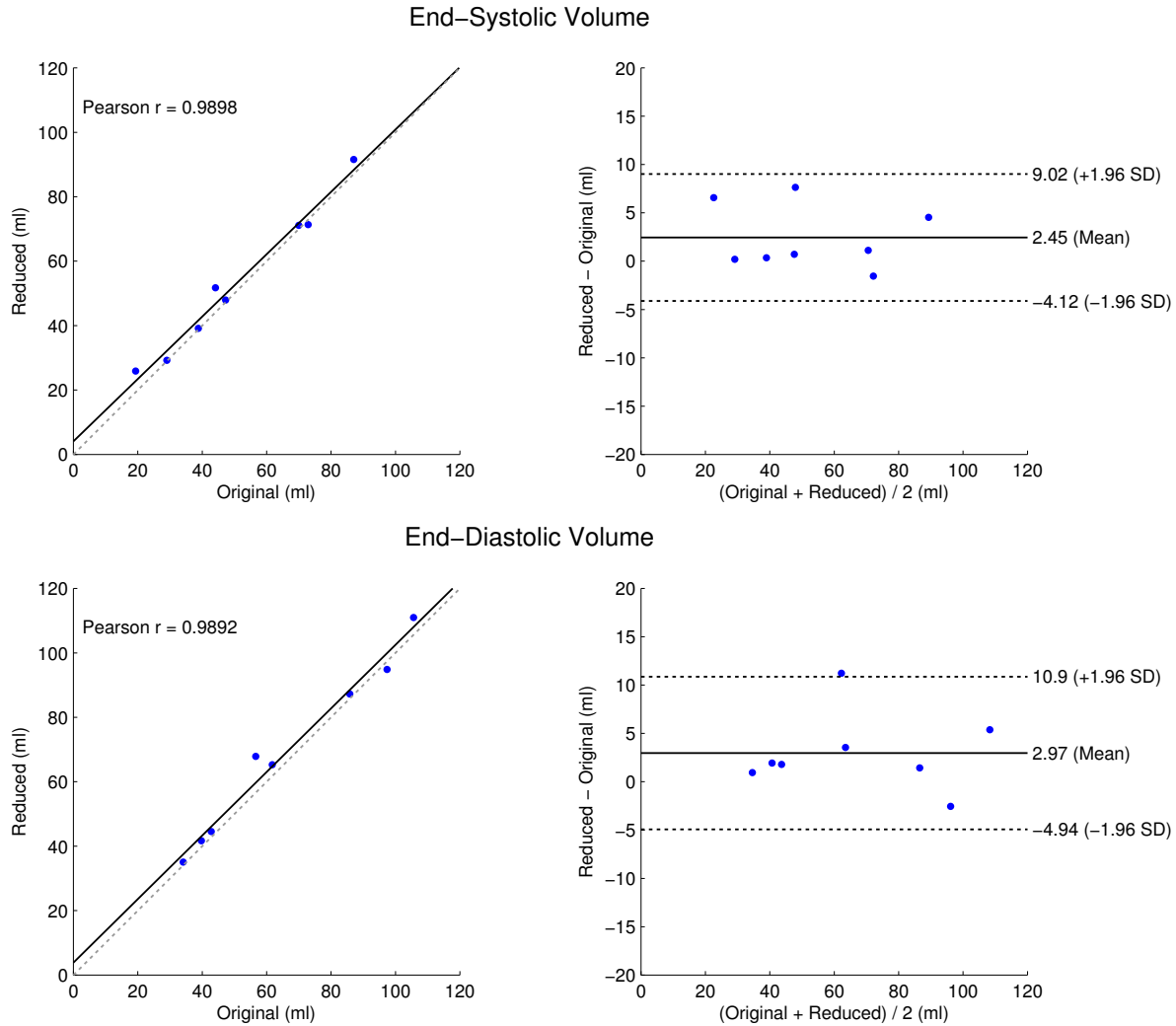


Figure 8. Scatter plots (left) and Bland-Altman plots (right) of end-systolic volumes (top) and end-diastolic volumes (bottom) estimated from original vs. reduced data.

emanating from catheters. The slight deterioration of resolution is due to the fact that with less data being available, optimization needs to rely more on the regularization terms which favor a homogeneous image appearance. Nevertheless, the LV boundary is reconstructed similarly well in most patients, which can be seen from the volumetry results in Fig. 8. Both end-systolic and end-diastolic volumes estimated from the reduced and original data show a Pearson correlation coefficient of nearly 99 % (Fig. 8, left). In each, there appears to be a small bias of less than 3 ml, with the 95 % confidence interval of the difference falling just below 11 ml in end-diastole at its extreme (Fig. 8, right). We perform a two-tailed paired t-test to assess these apparent biases, which reveals that they are not statistically significant. We obtain p-values of 0.0772 and 0.0769 for the end-systolic and end-diastolic volumes, respectively. Therefore, it cannot be concluded that the differences constitute a systematic error.

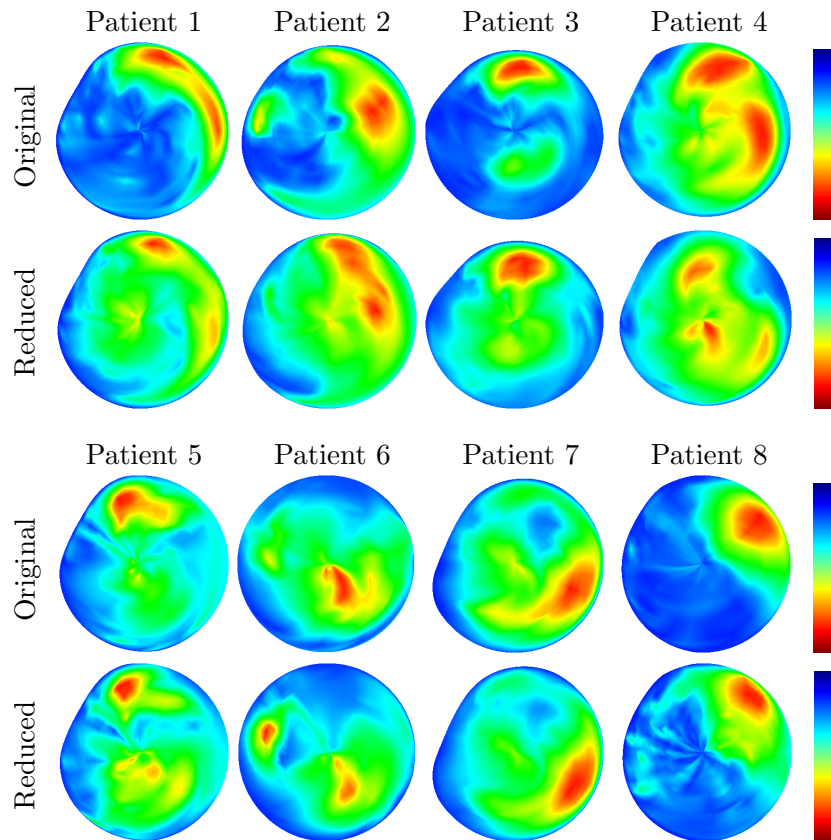


Figure 9. Polar maps of the magnitude of motion between end-diastole and end-systole per surface vertex. The darkest blue and brightest red hues correspond to zero and maximum motion (determined per case and setting), respectively.

In the polar maps displayed in Fig. 9, the regions of strongest contraction match quite well in most patients. In some cases, e.g. for patients 5 and 6, additional such regions appear erroneously when edges are mislocated due to the impaired image quality. Also, regions of comparatively little motion in the original data may be obscured by artifacts in the reduced setting, which is most prominent in patients 1 and 3.

4. Discussion

4.1. Interpretation of the Results

The grid search demonstrates that temporal TV by itself improves reconstruction quality more than spatial TV (cf. Fig. 4), even though it only operates along one dimension of the 4-D image as opposed to three. To understand why this is the case, one has to consider that a very small fraction of all reconstructed image voxels coincide with observably dynamic anatomy such as the endocardial wall (blood pool boundary). The vast majority shows virtually static image content such as the lungs, spine and ribs, or air in the background. Since these parts of the object are captured faithfully in all gating subsets, temporal regularization, which penalizes inconsistent changes over

time and thus substantially restricts the high-dimensional solution space, is so effective. In this way, redundancy in the temporal domain can be leveraged to obtain a more consistent solution. Combined with the denoising capabilities of spatial TV, the achieved image quality exceeds what 3-D reconstruction from only a couple of views can typically provide.

In most clinical cases, while the correlation of the volumetric measurements was nearly perfect, both end-diastolic and end-systolic volumes were overestimated rather than underestimated in the reduced setting. While the difference is not statistically significant, it could be due to the fact that blurring makes the blood pool appear somewhat larger (cf. Fig. 7), causing the boundary detector to locate its edge slightly outwards from its actual position. To some extent, it may be possible to specifically address this in the design of the segmentation algorithm, which is beyond the scope of this paper. The polar maps of Fig. 9 suffer from artifacts in the images reconstructed from reduced data. While the consistently reliable localization of “hot spots” is promising, at this time it must remain unclear whether the achieved accuracy is sufficient for clinically relevant use cases, which have not yet been established as development of dynamic interventional imaging is still in a very early stage. Should it be insufficient, denser angular sampling may yet be required.

4.2. Practical Considerations

Essentially, a key requirement for potential applications of 4-D cardiac C-arm CT is ensuring a sufficient number of heart cycles acquired during the scan. This leads to a trade-off as it can be achieved either by a sufficiently high heart rate, possibly enforced through electrophysiological pacing, or by a sufficiently long scan time. The former can alter the natural contraction patterns and, thereby, the heart motion, limiting the explanatory power of the images. The latter is feasible in principle, but limited by, e. g., the need for a prolonged contrast injection which may be harmful for the patient. In prior work, scan protocols with acquisition durations of about 10 s (Mory et al. 2014) and 15 s (Müller, Maier, Schwemmer, Lauritsch, Buck, Wielandts, Hornegger & Fahrig 2014) were used.

The importance of such considerations motivated us to conduct this study, which constitutes a technical investigation of state-of-the-art compressed sensing reconstruction applied to clinical data as well as its behavior and potential performance under data degradation. However, the presented results are not meant to claim sufficiency of a certain amount of projection data or, correspondingly, recommend a certain dose level for use in the clinical workflow. For specific applications, as well as for other tasks with potentially different requirements, further studies will be required for validation. It also has to be noted that our clinical cases are limited in number and somewhat heterogeneous regarding the pacing and contrasting protocol. Despite this, the achieved results are consistent and stable.

5. Conclusion and Outlook

Spatially and temporally TV-regularized reconstruction is effective in recovering dynamic cardiac images from as few as 6.9 ± 1.2 projection views per heart phase. In clinical patients, segmentation-based volumetry can be performed with very high correlation and only little bias compared to less sparsely sampled data. Additionally, visual exploration of local contraction patterns appears to be feasible to a certain extent, but clinical studies will be mandatory to validate whether this is sufficient for basing potential treatment decisions on the employed visualizations.

In the future, the proposed reconstruction could be enhanced by incorporating more sophisticated data fidelity models, such as a noise model for statistical ray weighting, or improving convergence by, e.g., pre-conditioned optimization, cf. (Xu et al. 2016). Regarding regularization, the image characteristic of TV is helpful for segmentation and in turn volumetry, which was the goal in our study. For other purposes, smoother priors based on, for instance, wavelet transforms may be preferable and could be investigated in the same framework.

Acknowledgments The authors gratefully acknowledge funding by Siemens Healthcare GmbH as well as funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative. We would like to thank Daphne Yu for providing the volume rendering of our data used in Fig. 2. *Disclaimer:* The concepts and information presented in this paper are based on research and are not commercially available.

References

- Cerqueira M D, Weissman N J, Dilsizian V, Jacobs A K, Kaul S, Laskey W K, Pennell D J, Rumberger J A, Ryan T, Verani M S et al. 2002 Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association *Circulation* **105**(4), 539–542.
- Chambolle A & Pock T 2011 A first-order primal-dual algorithm for convex problems with applications to imaging *J. Math. Imaging Vis.* **40**(1), 120–145.
- Condat L 2013 A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms *Journal of Optimization Theory and Applications* **158**(2), 460–479.
- Condat L 2014 A generic proximal algorithm for convex optimization – application to total variation minimization *IEEE Signal Processing Letters* **21**(8), 985–989.
- Haase V, Taubmann O, Huang Y, Krings G, Lauritsch G, Maier A & Mertins A 2016 Make the Most of Time: Temporal Extension of the iTV Algorithm for 4D Cardiac C-Arm CT *in* T Tolxdorff, T Deserno, H Handels & H.-P Meinzer, eds, ‘Bildverarbeitung für die Medizin 2016’ Berlin Heidelberg pp. 170–175.
- Hetterich H, Redel T, Lauritsch G, Rohkohl C & Rieber J 2010 New X-ray imaging modalities and their integration with intravascular imaging and interventions *The International Journal of Cardiovascular Imaging* **26**(7), 797–808.

- Lauritsch G, Boese J, Wigstrom L, Kemeth H & Fahrig R 2006 Towards cardiac C-arm computed tomography *IEEE Trans. Med. Imag.* **25**(7), 922–934.
- Maier A, Hofmann H, Berger M, Fischer P, Schwemmer C, Wu H, Müller K, Hornegger J, Choi J H, Riess C, Keil A & Fahrig R 2013 CONRAD - A software framework for cone-beam imaging in radiology *Med. Phys.* **40**(11).
- Maier A, Hofmann H, Schwemmer C, Hornegger J, Keil A & Fahrig R 2012 Fast Simulation of X-ray Projections of Spline-based Surfaces using an Append Buffer *Phys. Med. Biol.* **57**(19), 6193–6210.
- Mory C, Auvray V, Zhang B, Grass M, Schäfer D, Chen S, Carroll J, Rit S, Peyrin F, Douek P & Boussel L 2014 Cardiac C-arm computed tomography using a 3D + time ROI reconstruction method with spatial and temporal regularization *Med. Phys.* **41**, 021903.
- Mory C & Jacques L 2014 A modified 4D ROOSTER method using the Chambolle-Pock algorithm in ‘Proc. 3rd Intl. Conf. on Image Formation in X-ray CT’ pp. 191–193.
- Müller K, Maier A, Schwemmer C, Lauritsch G, Buck S D, Wielandts J Y, Hornegger J & Fahrig R 2014 Image artefact propagation in motion estimation and reconstruction in interventional cardiac C-arm CT *Phys. Med. Biol.* **59**(12), 3121–3138.
- Müller K, Maier A, Zheng Y, Wang Y, Lauritsch G, Schwemmer C, Rohkohl C, Hornegger J & Fahrig R 2014 Interventional heart wall motion analysis with cardiac C-arm CT systems *Physics in Medicine and Biology* **59**(9), 2265–2294.
- Ritschl L, Bergner F, Fleischmann C & Kachelrieß M 2011 Improved total variation-based CT image reconstruction applied to clinical data *Physics in Medicine and Biology* **56**(6), 1545.
- Ritschl L, Sawall S, Knaup M, Hess A & Kachelrieß M 2012 Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior *Physics in Medicine and Biology* **57**(6), 1517.
- Schäfer D, Borgert J, Rasche V & Grass M 2006 Motion-compensated and gated cone beam filtered back-projection for 3-D rotational X-ray angiography *IEEE Trans. Med. Imag.* **25**, 898–906.
- Segars W P, Sturgeon G, Mendonca S, Grimes J & Tsui B M W 2010 4D XCAT phantom for multimodality imaging research *Med. Phys.* **37**, 4902–4915.
- Taubmann O, Lauritsch G, Krings G & Maier A 2016 Convex temporal regularizers in cardiac C-arm CT in ‘Proc. 4th Intl. Conf. on Image Formation in X-ray CT’ pp. 545–548.
- Taubmann O, Lauritsch G, Maier A, Fahrig R & Hornegger J 2015 Estimate, Compensate, Iterate: Joint Motion Estimation and Compensation in 4-D Cardiac C-arm Computed Tomography in N Navab, J Hornegger, W. M Wells & A. F Frangi, eds, ‘Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)’ Vol. 9350 of *Lecture Notes in Computer Science* pp. 579–586.
- Taubmann O, Maier A, Hornegger J, Lauritsch G & Fahrig R 2016 Coping with Real World Data: Artifact Reduction and Denoising for Motion-Compensated Cardiac C-arm CT *Medical Physics* **43**(2), 883–893.
- Wielandts J Y, De Buck S, Ector J, Nuyens D, Maes F & Heidebuechel H 2014 Registration-based filtering: An acceptable tool for noise reduction in left ventricular dynamic rotational angiography images? *Proc. SPIE* **9036**, 903628–903628–7.
- Wielandts J Y, De Buck S, Michielsen K, Louw R, Garweg C, Nuyts J, Ector J, Maes F & Heidebuechel H 2015 Multi-phase rotational angiography of the left ventricle to assist ablations: feasibility and accuracy of novel imaging *European Heart Journal – Cardiovascular Imaging*.
- Xu Q, Yang D, Tan J, Sawatzky A & Anastasio M A 2016 Accelerated fast iterative shrinkage thresholding algorithms for sparsity-regularized cone-beam CT image reconstruction *Medical Physics* **43**(4), 1849–1872.
- Zheng Y, Barbu A, Georgescu B, Scheuering M & Comaniciu D 2008 Four-Chamber Heart Modeling and Automatic Segmentation for 3-D Cardiac CT Volumes Using Marginal Space Learning and Steerable Features *IEEE Transactions on Medical Imaging* **27**(11), 1668–1681.
- Zheng Y, Georgescu B, Vega-Higuera F & Comaniciu D 2009 Left ventricle endocardium segmentation

for cardiac CT volumes using an optimal smooth surface *Proc. SPIE* **7259**, 72593V–72593V–11.