Motion correction for 4D CBCT reconstruction with TV regularization

Rob Heylen, Johan Nuyts

***Abstract*—Recently, we proposed a 4D CBCT reconstruction algorithm with spatial and temporal total variation regularization, and employed this algorithm for dual-energy dynamic angiography of the brain on a software phantom. This approach is very sensitive to movement, as misalignment between time frames will interfere with the temporal regularization. In this work, we integrate an iterative motion correction algorithm with the 4D CBCT reconstruction algorithm. Several strategies can be followed, and we assess which strategy is most successful. Experiments are performed on a dynamical software phantom.**

I. Introduction

With 4D cone-beam CT (CBCT) one can reconstruct multiple 3D time frames from a single CBCT acquisition. As the angular interval for each time frame is relatively small, prior information is typically employed to reduce artefacts and improve image quality. Many authors have shown that total variation (TV) regularization in both the time and spatial domain is very useful for 4D CBCT reconstruction, with applications in cardiac, pulmonary and brain imaging [1], [2]. Computationally efficient implementations based on the primal-dual optimization algorithm of Chambolle and Pock [3] have recently been presented as well [4], [5], [6].

Furthermore, dual-energy (DE) CBCT systems are being deployed more frequently, which increases the diagnostic capabilities of CBCT examinations. A popular application is material decomposition, where the dual-energy image is transformed into two material images. This decomposition can be performed in image space, but also in data space when the DE views are matched. An interesting decomposition basis for dynamic angiography of the brain would be a water and iodine component, where the iodine image visualizes the injected contrast. Applying 4D CBCT reconstruction to such an iodine image can generate several time frames of the contrast flow, corresponding to different phases of the bolus traversal, and can provide time-resolved dynamic collateral imaging with a single CBCT acquisition.

A TV prior penalizes differences between neighbouring voxels, either in the spatial or temporal direction. Such a prior is sensitive to movement, as voxels should stay registered between different temporal frames. As the delay between these time frames is in the order of several seconds, one cannot assume that these frames stay registered with sub-voxel accuracy in practical use, and a registration method or motion correction needs to be applied.

In this work, we investigate the influence of significant movement on 4DTV CBCT reconstruction methods, and the ability of a motion correction method to alleviate the resulting issues. Experiments are performed on a dynamical software phantom, so that ground truth is always available.

Manuscript received January 17, 2020.

R.H. and J.N. are with the Department of Imaging and Pathology, Division of Nuclear Medicine, KULeuven, Belgium. This work was done under the NEXIS project, that has received funding from the European Union's Horizon 2020 Research and Innovations Programme (Grant Agreement #780026).

II. 4D CBCT reconstruction

The algorithm for 4D CBCT reconstruction combines several data fidelity and prior information terms into a single optimization equation, which is solved via the Chambolle-Pock primal-dual optimization algorithm [3]. The algorithm is very similar to the ROOSTER algorithm proposed in [4], but has slight differences in implementation and interpretation.

We consider a single circular DE CBCT scan, typically over 360 degrees and with an acquisition time of around 25s. We focus on reconstructing a 4D reconstruction of an iodine sinogram, obtained by material decomposition of a cranial DE CBCT scan. We assume that the image intensities change along linear curves between different time frames, and that these changes are due to changes of the iodine contrast concentrations. See also [7] for a more detailed explanation of the algorithm.

We consider several terms in the optimization equation:

- A data fidelity term that penalizes differences between the forward projection of the reconstruction and the iodine sinogram. The penalty is implemented as the L2 norm of the difference. The forward and backward projectors employ linear interpolation between the different time frames.
- A binary mask indicating static areas, such as the skull. The masked voxels are averaged over all time frames. This significantly improves performance, especially when highly attenuating static areas are masked. A heuristic method to obtain a skull mask is to threshold a static water image, obtained by FDK reconstruction of the water sinogram.
- Non-negativity, as the amount of iodine in a voxel cannot be negative.
- TV regularization, both in the spatial direction (3D) and the temporal direction (1D). Both regularizers can be weighted differently, and we choose the temporal TV regularizer 10 to 100 times stronger than the spatial regularizer.

This leads to the optimization equation

$$\underset{x\in X}{\operatorname{argmin}} \left(\|Ax - s\|_2^2 + \alpha \|x\|_{\text{TV4}} + \mathrm{I}_{\mathbb{R}^{\mathbb{N}+}}(x) \right) \tag{1}$$

with

$$\|x\|_{\text{TV4}} = \sum_{ijkt} |(\Delta x)_{ijkt}| \tag{2}$$

$$|(\Delta x)_{ijkt}| \quad = \quad \left(\sum_{d=1}^{3} (\Delta_d x)_{ijkt}^2 + (\gamma \Delta_4 x)_{ijkt}^2 \right)^{1/2} \tag{3}$$

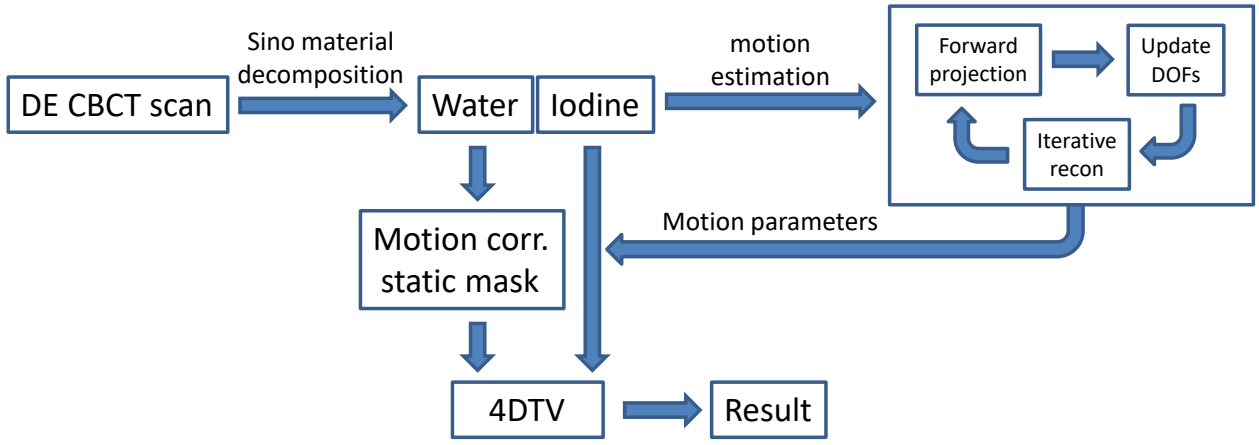

Fig. 1: Diagram representing the motion-corrected DE CBCT data processing chain for dynamic collateral imaging.

and $\mathrm{I}_{\mathbb{R}^{N+}}(.)$ the characteristic function of the positive orthant. The γ factor in the last term weighs the spatial and temporal directions differently. This equation can be written in terms of the Chambolle and Pock framework, yielding algorithm 1 to solve (1). See [3], [4], [6], [7] for more information.

Algorithm 1: The 4DTV algorithm

Initialize all variables, choose τ, $\sigma > 0$, $\theta \in [0, 1]$;
for $n \geq 0$ **do**
  $y_1^{n+1} = \frac{2}{2+\sigma}\left(y_1^n + \sigma(A\overline{x} - s)\right)$;
  $x_1^{n+1} = A^* y_1^{n+1}$;
  $y_2^{n+1} = \mathrm{proj}_{\alpha P}\left(y_2^n + \sigma\Delta\overline{x}\right)$;
  $x_2^{n+1} = -\mathrm{div}\left(y_2^{n+1}\right)$;
  $x^{n+1} = \mathrm{proj}_{\mathbb{R}^{N+}}\left(x^n - \tau x_1^{n+1} - \tau x_2^{n+1}\right)$;
  $\overline{x}^{n+1} = x^{n+1} + \theta(x^{n+1} - x^n)$;

III. Rigid motion correction and 4DTV

We assume that rigid motion might be present during the CBCT acquisition. This rigid motion can be described by 6 degrees of freedom (DOF), corresponding to the Euler angles and a 3D translation, for each view/angle in the sinogram. We propose a data-driven method for estimating this motion, adapted from the recently proposed method for rigid motion correction in helical CT [8], [9].

The rigid motion correction method is an iterative procedure, consisting of the following steps. See [8], [9] for more details:

1) Determine an initial static reconstruction, e.g. with the FDK algorithm. Set $k = 1$.
2) Downscale/Upscale the sinogram and the current reconstruction by a factor D_k. Set $it = 1$.
3) For each DOF:
   a) Calculate the numerical derivative of the L2 reconstruction error wrt. this DOF with finite differences.
   b) Update this DOF for each view/angle by minimizing the L2 error. We employ a linear extrapolation method.
4) Use an iterative reconstruction method to improve the current reconstruction employing the updated DOF estimates.
5) Increase it. If $it \geq N_k$ increase k and go to 2, else go to 3.

The employed iterative reconstruction method is the proposed 4DTV algorithm, with motion-aware projectors and backprojectors. Note that a static version of this algorithm can be easily obtained by choosing the number of time frames equal to one, and that the TV terms can be deactivated as well, leading to a classical iterative reconstruction algorithm.

Calculation of the numerical derivative in step 3(a) only requires an additional forward projection for 3 of the 6 DOFs: Translations parallel to the sensor plane and axial rotations do not require an additional forward projection, as these can be easily calculated as 2D geometrical operations in the sensor plane. We also noted experimentally that repeating step 3 a couple of times before moving to step 4 improves the convergence significantly, as each updated DOF influences the others in turn. Most of the computational time is spent in the iterative reconstruction in step 4.

By employing multiple scales, as implemented in step 2 of the algorithm, the convergence can be sped up in a major way: significant motion can be easily estimated on downsampled data sets, and the computational resources required decrease fast. We typically employed $D = [4, 2, 1]$ downsampling in our experiments, with $N = [5, 3, 1]$ iterations for each scale. The iterative reconstruction algorithm was run for 10 iterations each time.

We tested two strategies for combining the 4DTV reconstruction algorithm with the motion estimation: simultaneous estimation of rigid motion and multiple time frames by using the full implementation of 4DTV in step 4 of the algorithm, or cascade the two algorithms, where first the rigid motion is estimated (assuming a static reconstruction), followed by 4DTV reconstruction using the estimated motion and the static reconstruction as initialization. We found that the second strategy produces better results than the first, and has a lower computational cost.

A schematic representation of the entire data processing chain is given in Fig. 1

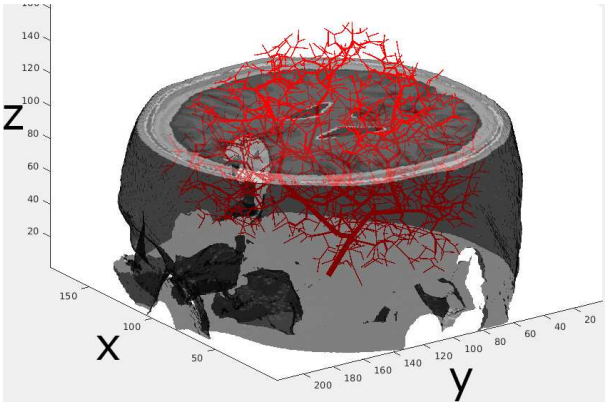

Fig. 2: The simulated vasculature inside the brainweb phantom.

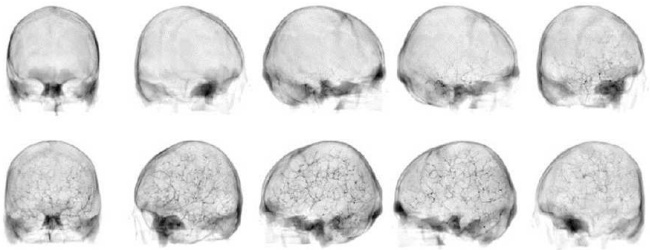

Fig. 3: 10 views of the iodine sinogram, evenly spread over 360 degrees, showing the evolution of the contrast through the vascular tree.

IV. Experiments

A. The dynamical phantom

We created a dynamic phantom for testing the capabilities of the presented algorithm on simulated DE CBCT scans (see also [7] for more information). This phantom is based on the brainweb phantom [10], where we added an additional dynamical iodine class consisting of a simulated arterial tree [11] inside the grey/white matter with contrast dynamics. The dynamics was obtained by considering the binary tree as a network of interconnected tubes, and calculating the hydrodynamics of such a system under laminar flow [12]. An input function was obtained from a voxel in the carotid area of a perfusion CT examination. The diffusion coefficient, the single free parameter in the dynamics, was modified such that the time for the bolus to reach the terminal points was comparable to the time delay between the arterial and venous phase in the perfusion data set.

This data set was forward projected with a DE CBCT simulator, and decomposed into water and iodine components via Newton optimization of the spectral detector model. A visualization of the arterial tree is shown in Fig. 2, and several views of the iodine sinogram in Fig. 3, showing the dynamics as the contrast bolus passes through the network. The total scan time was 25 seconds, and we simulated a full 360 rotation with 200 views of 300×300 pixels of size $(1\text{mm})^2$, with a focus height of 1000mm and a detector distance of 150mm. The reconstructed volume corresponds with the size of the brainweb phantom, and has size $(181, 217, 181)$, with $(1\text{mm})^3$ voxels.

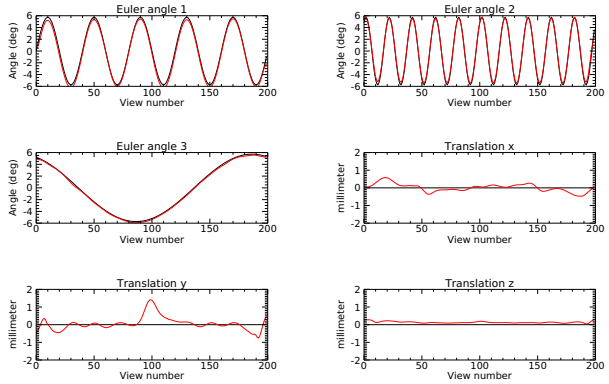

Fig. 4: The true movement (black) and the calculated movement by the rigid motion correction (red), for all 6 DOFs.

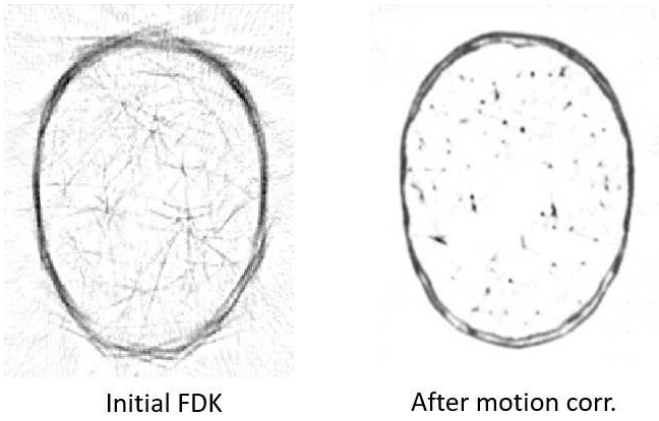

Fig. 5: Central slice of static reconstructions before (left) and after (right) motion correction.

B. Rigid motion and estimation

To add rigid motion to the data set, we modified each DOF as a function of view/angle as the CBCT acquisition takes place. These movement functions are displayed in Fig. 4, along with the estimated movement after rigid motion correction. We changed only rotational DOFs, keeping the translational DOFs equal to zero. Note that although these motion functions are artificial, they represent significant motion in all rotational directions.

The rigid motion estimation procedure succeeds in estimating most DOFs relatively well, with only some translational deviations for some views. Note that there are several invariances allowed in the estimated DOFs, as the estimation procedure might result in a translated or rotated coordinate frame which is also valid for reconstruction. This hampers quantitative assessment strategies based on voxel-wise comparison.

In Fig. 5, the input and output images of the motion correction procedure are shown. The input is an FDK reconstruction of the iodine sinogram including movement. It is clear that the motion corrected image is of much better quality.

C. 4DTV reconstruction

After motion correction, the estimated motion parameters are used in the 4DTV algorithm for reconstructing multiple time frames. Also a static motion-corrected version of the

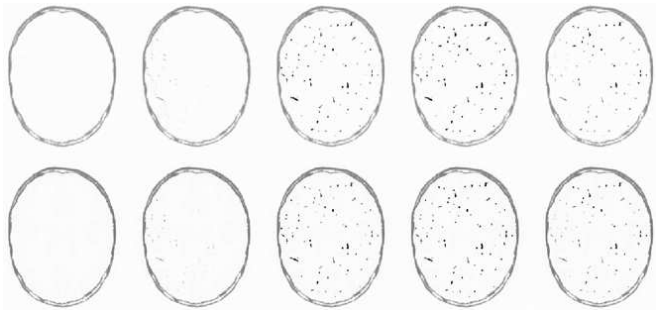

Fig. 6: Central slice of 5 dynamically reconstructed frames after motion correction (top row), and the corresponding ground truth (bottom row).

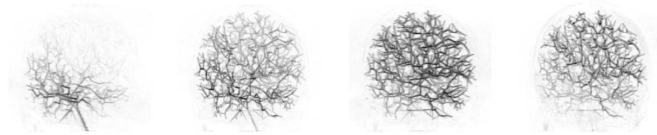

Fig. 7: MIPs of subtraction images with the first frame.

water image has to be reconstructed to determine the static skull mask.

We reconstructed 5 time frames with 100 iterations of the 4DTV algorithm. The central slice of these reconstructions is shown in Fig. 6, along with the ground truth obtained from the dynamical software phantom.

Because the first reconstructed time frame does not yet show any contrast dynamics, we can employ this frame to create digital subtraction images. Maximum intensity projections (MIP) of these subtraction images are shown in Fig. 7.

To stress the influence of the motion correction step in the reconstruction procedure, we have also recreated all results with the same 4DTV parameters, but without added motion or motion correction. The second MIP from Fig. 7 is repeated in Fig. 8, along with the version without motion and the MIP obtained from the ground truth. While the motion-corrected data set has more blur, it is clear that these results are qualitatively comparable.

V. Conclusions

We have shown that motion correction can be combined with 4DTV, significantly improving reconstruction results, and that using motion correction and 4D reconstruction successively is superior to integrating 4DTV and iterative motion correction into a single algorithm. We performed experiments on an artificial dynamical data set, obtained with a software phantom including contrast dynamics, and show that the effects of motion can be alleviated almost completely. In future work, we will apply these techniques to realistic Monte Carlo simulations and real data sets, aiming to obtain dynamic collateral imaging with DE CBCT devices.

Fig. 8: The motion-corrected 4DTV result of the data set including motion (left), the 4DTV result on the original data set without motion (middle), and the ground truth (right).

References

[1] D.C.Hansen and T. Sorensen, "Fast 4d cone-beam ct from 60 s acquisitions," *Phys. Im. in Rad. Oncology*, vol. 5, pp. 69–75, 2018.

[2] L. Ritschl, S. Sawall, M. Knaup, A. Hess, and M. Kachelriess, "Iterative 4d cardiac micro-ct image reconstruction using an adaptive spatio-temporal sparsity prior," *Phys. Med. Biol.*, vol. 57, pp. 1517–1525, 2012.

[3] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, May 2011.

[4] C. Mory and L. Jacques, "A modified 4d rooster method using the chambolle–pock algorithm," *Proc. 3rd Intl. Conf. on Image Formation in X-Ray CT*, pp. 191–193, 2014.

[5] O. Taubmann, V. Haase, G. Lauritsch, Y. Zheng, G. Krings, J. Hornegger, and A. Maier, "Assessing cardiac function from totalvariation-regularized 4d c-arm ct in the presence of angular undersampling," *Phys. Med. Biol.*, vol. 62, p. 2762–2777, 2017.

[6] V. Nikitin, M. Carlsson, F. Andersson, and R. Mokso, "Four-dimensional tomographic reconstruction by time domain decomposition," *IEEE Tran Comp. Im.*, pp. 1–1, Feb 2019.

[7] R. Heylen, G. Schramm, P. Suetens, and J. Nuyts, "4d cbct reconstruction with tv regularization on a dynamic software phantom," *IEEE Nuclear science symposium and medical imaging conference*, pp. 1–3, 2019.

[8] T. Sun, J.-H. Kim, R. Fulton, and J. Nuyts, "An iterative projection-based motion estimation and compensation scheme for head x-ray ct," *Medical Physics*, vol. 43, no. 10, pp. 5705–5716, 2016.

[9] J.-H. Kim, T. Sun, A. Alcheikh, Z. Kuncic, J. Nuyts, and R. Fulton, "Correction for human head motion in helical x-ray ct," *Phys Med Biol*, vol. 61, no. 4, pp. 1416–1438, 2016.

[10] C. Cocosco, V. Kollokian, R.-S. Kwan, and A. Evans, "Brainweb: Online interface to a 3d mri simulated brain database," *NeuroImage*, vol. 5, no. 4, pp. part 2/4, S425, 1997.

[11] R. Karch, F. Neumann, M. Neumann, and W. Schreiner, "A three-dimensional model for arterial tree representation, generated by constrained constructive optimization," *Computers in Biology and Medicine*, vol. 29, no. 1, pp. 19 – 38, 1999.

[12] G. I. Taylor, "Dispersion of soluble matter in solvent flowing slowly through a tube," *Proc. Royal Soc. of London. Ser. A. Math. Phys. Sci.*, vol. 219, no. 1137, pp. 186–203, 1953.