Regularized 4D-CT reconstruction from a single dataset with a spatio-temporal prior

Fabien Momey, Eric Thiébaut, Catherine Burnier-Mennessier, Loïc Denis, Jean-Marie Becker, Laurent Desbat

Regularized reconstruction for dynamic X-ray CT without motion compensation

Fabien Momey*, Éric Thiébaut, Catherine Burnier, Loïc Denis, Jean-Marie Becker, and Laurent Desbat

***Abstract*—X-ray Computerized Tomography (CT) reconstructions can be severely impaired by the patient's respiratory motion and cardiac beating. Motion must thus be recovered in addition to the 3D reconstruction problem. The approach generally followed to reconstruct dynamic volumes consists of largely increasing the number of projections so that independent reconstructions be possible using only subsets of projections from the same phase of the cyclic movement. Other methods are based on motion compensation (MC) using a deformation model estimated beforehand on an other dynamic CT data set.**

Our work takes a different path; it uses dynamic reconstruction, based on inverse problems approach, without any additional measurements, nor explicit knowledge of the motion. The dynamic sequence is reconstructed out of a single data set, only assuming the motion's continuity and periodicity. This inverse problem is solved by the minimization of the sum of a data-fidelity term, consistent with the dynamic nature of the data, and a regularization term which implements an efficient spatio-temporal version of the total variation (TV). We demonstrate the strength of our approach and its practical feasibility on 2D and 3D+t reconstructions of a mechanical phantom and patient data.

***Index Terms*—Dynamic tomography, Reconstruction, Inverse Problems, Regularization, Signal processing.**

I. Introduction

IN X-ray dynamic CT, the motion induced by the breathing and the heart beating of the patient implies that the acquired projections, *i.e.* the data, are not related to the same "static object". Reconstructing the patient's anatomy as if there was no motion causes dramatical artifacts [2], [34], [8]. In other words, the inverse problem of reconstruction has to be addressed in 4D. More specifically in 3D+t, where the motion and deformation of anatomical structures as a function of time are recovered. In radiotherapy, this is a critical matter because the planning of treatment by the medical physicist requires a precise localization of the lung tumor, in order to preserve at best healthy tissues during their irradiation. As a result, the "object" to be reconstructed is modelized as a spatio-temporal signal $f(\boldsymbol{x}, t)$, defined for spatial coordinates $\boldsymbol{x} \in \mathbb{R}^3$ and time $t \in \mathbb{R}$.

Dynamic CT reconstruction has benefited from active research for the past twenty years. Some investigations were made on the scanning protocol itself [14], [27], [39], [40]. Ritchie *et al.* [27] established early that even ultrafast scanning was not sufficient to avoid artifacts. Such an observation still holds twenty years later. A rather frequent ad hoc solution is based on a correlation between the 3D CT data and a 1D temporal record of the patient's pseudo-periodical movements. The acquired projections are then sorted, according to the phase of the cycle to which they are related. Data subsets are extracted, and independent static reconstructions are then performed using each of these subsets. Such a method, called *gated CT* or *4D-CT* [32], [21], [13], [37], [19], [3], [35], [24], requires a sufficient number of projections for each reconstructed phase, increasing the amount of X-ray dose delivered to the patient for the reconstruction of the whole sequence.

Another family of methods recover the motion in a first step, then use the obtained deformation model in a second step to compensate for the motion. The motion is generally estimated in the form of a deformation vector field $\Gamma_t(\boldsymbol{x})$ that maps the volume at any time t back to its state at a reference time t_0:

$$f(\boldsymbol{x}, t) = f\left(\Gamma_t(\boldsymbol{x}), t_0\right) . \tag{1}$$

Such a motion model can be incorporated into the tomographic projector in order to use all the available projections to reconstruct the patient's anatomy at this reference state: a process called *motion compensation* (MC). These approaches heavily rely on the quality of the estimation of the motion, a challenging inverse problem, especially because of the complex modeling of deformations such as breathing or cardiac beating. Moreover, in most cases, for example in dynamic Cone-Beam CT (CBCT), motion estimation is done on an additional 4D-CT reconstruction [6], [5], [20], [25], [26], requiring more measurements and therefore higher X-ray doses.

Our approach proposes to use only the current data set for the reconstruction of the 3D+t object, which suppresses the

This work was supported by the MiTiV project (Méthodes Inverses pour le Traitement en Imagerie du Vivant), funded by the French ANR (N° ANR-09-EMER-008).

Fabien Momey* was with the Université de Lyon, F-42023, Saint-Etienne, France; CNRS, UMR5516, Laboratoire Hubert Curien, F-42000, Saint-Etienne, France; Université de Saint-Etienne, Jean Monnet, F-42000, Saint-Etienne, France. He was also with the Université de Lyon, Lyon, F-69003, France; Université Lyon 1, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval, F-69230, France; CNRS, UMR 5574, Centre de Recherche Astrophysique de Lyon; École Normale Supérieure de Lyon, Lyon, F-69007, France. He is now with the Université Grenoble Alpes, F-38000 Grenoble, France; CEA, LETI, MINATEC Campus, F-38054 Grenoble, France (e-mail: fabien.momey@cea.fr).

Éric Thiébaut is with the Université de Lyon, Lyon, F-69003, France; Université Lyon 1, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval, F-69230, France; CNRS, UMR 5574, Centre de Recherche Astrophysique de Lyon; École Normale Supérieure de Lyon, Lyon, F-69007, France (e-mail: eric.thiebaut@univ-lyon1.fr).

Loïc Denis, Catherine Burnier and Jean-Marie Becker are with the Université de Lyon, F-42023, Saint-Etienne, France; CNRS, UMR5516, Laboratoire Hubert Curien, F-42000, Saint-Etienne, France; Université de Saint-Etienne, Jean Monnet, F-42000, Saint-Etienne, France (e-mail: {loic.denis,catherine.burnier,jean-marie.becker}@univ-st-etienne.fr).

Laurent Desbat is with the Université Grenoble Alpes, TIMC-IMAG, F-38000 Grenoble, France; CNRS, TIMC-IMAG, F-38000 Grenoble, France; CHU de Grenoble,TIMC-IMAG, F-38000 Grenoble, France (e-mail: laurent.desbat@imag.fr).

need for over-numerous projections. No explicit knowledge of the motion is required since we directly reconstruct the whole 3D+t sequence, globally addressing the inverse problem of reconstruction from dynamic projections as the minimization of a data-fidelity term and a regularization term, as will be seen later (3). We adapt the classical tomographic projection model to deal with the fourth temporal dimension and to calculate the model of the projections at a given time t. To account for the specific continuity of the spatio-temporal object, we implemented a 3D+t edge-preserving smoothness regularization based on Rudin et al. [28] total variation (TV).

In this introduction, we have pointed out the drawbacks of motion-compensated methods and summarized our new approach. As explained in Section II, the demonstration of our method will be based on the cone beam CT (CBCT) modality. Our approach, fully detailed in Section III, is however very general and can readily be applied to other modalities. In Section IV, we demonstrate its strength on 2D+t reconstructions from numerically simulated dynamic data and then on 2D+t and 3D+t reconstructions of a mechanical phantom and real patient data.

II. Cone-Beam CT: demonstration framework

A. *Motion is critical in Cone-Beam CT*

CBCT is an especially interesting modality for the study of dynamic respiration CT. Indeed, CBCT scanners have been implanted on linear accelerators for radiotherapy in the early 2000's [15], [16], [17]. The particularities of this type of system is the slow period of rotation of its gantry (about 2 minutes), and that all the projections (about 600-700) are acquired on a single rotation of its flat detector. With this modality, 4D-CT reconstructions are difficult due to the lack of projections, even if the slow rotation speed allows several respiratory cycles to be completed during the acquisition, giving a satisfying angular coverage for each reconstructed phase.

As a result, many motion-compensated reconstruction methods have been explored specifically for this type of data [20], [41], [25], [26]. Li *et al.* [20] have incorporated the deformation field in the analytic reconstruction algorithm FDK [12]. Rit *et al.* [25], [26] have compared the Li *et al.*'s approach with an iterative reconstruction method based on the SART algorithm [1], inserting the deformation field into the tomographic projection model. These approaches suppose that motion is unchanged between the acquisition on which motion is estimated and the CBCT acquisition for which motion is compensated. Zeng *et al.* [41] have proposed to avoid a preliminary 4D-CT reconstruction step by a direct estimation of the deformation model on the CBCT projections, using the complementary information given by an a priori 3D static reconstruction of the patient. However, the 3D anatomical model still has to be obtained from a previous reconstruction step, generally from breath-hold acquired data, a constraining step that generates additional irradiation.

B. *"2D CBCT" dynamic data simulation*

For testing purposes, we simulated a dynamic 2D phantom based on the well known Shepp-Logan model [33],

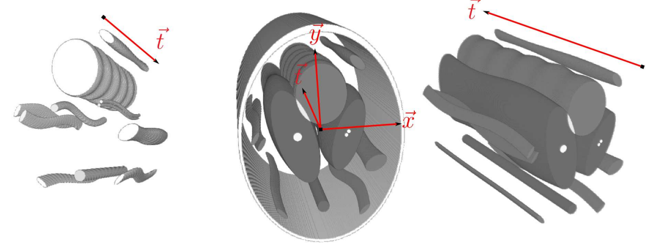

Fig. 1. Segmented 3D views of one motion cycle of our dynamic 2D Shepp-Logan phantom. The depth corresponds to the temporal dimension.

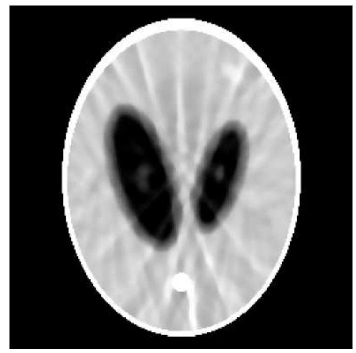

Fig. 2. Static reconstruction from the set of dynamic projections generated from the simulated dynamic Shepp-Logan phantom.

[18]. We let the shape parameters of its composing ellipses vary periodically over time, following a temporal sinusoidal signal with a given period. As a result, these ellipses can be translated, rotated or distorted, simulating the anatomical variations induced by the respiratory motion at any instant t. Fig. 1 illustrates, on 3D views, the anatomical evolution of one motion cycle of our phantom. The observed waves in the temporal direction show the anatomical variations of the ellipses.

Fan beam projections are computed in close form under the hypothesis of an instantaneous measurement. This yields CBCT-like dynamic data in 2D. We simulated 600 such projections for a 360° angle coverage corresponding to a total acquisition time of 120 seconds. The period of the "respiratory" cycle is set to 5 seconds. As a result, 24 complete motion cycles are done during the acquisition, with 25 projections per cycle.

Figure 2 displays a static reconstruction from the generated set of dynamic projections. The strong motion artifacts that result from the oversimplified "static" hypothesis highlight the need for a 2D+t sequence reconstruction.

III. Method

A. *Reconstruction criterion*

We solve the inverse problem [36] of 3D+t reconstruction by minimizing the joint criterion:

$$\boldsymbol{f}^{+} = \arg\min_{\boldsymbol{f}} \left\{ \mathcal{J}_{\mathsf{data}}(\boldsymbol{f}) + \mathcal{J}_{\mathsf{prior}}(\boldsymbol{f}) \right\} \,, \tag{2}$$

where $\mathcal{J}_{\mathsf{data}}$ and $\mathcal{J}_{\mathsf{prior}}$ are respectively the data-fidelity and regularization terms (detailed in the following paragraphs)

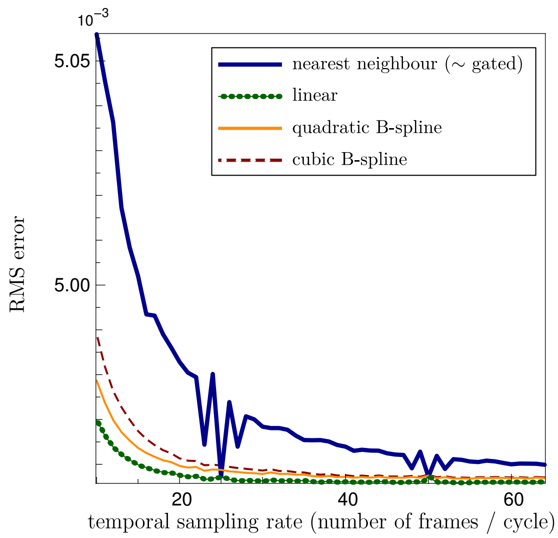

Fig. 3. RMS errors between simulated data of our dynamic Shepp-Logan phantom and the output of our dynamic projection model, $\mathbf{R}^\theta \cdot \mathbf{S}^{t_\theta}$, as a function of the temporal sampling rate. Each curve is for a given temporal interpolation method implemented by the operator $\mathbf{S}^{t_\theta}$.

and $\boldsymbol{f}$ denotes the parameters of the 3D+t sequence to be reconstructed, *i.e.* a 4D image of "spatio-temporal" voxels, corresponding to one period of motion.

B. Data fidelity and direct model

Assuming independent projections and Gaussian noise, the data fidelity writes:

$$\mathcal{J}_{\mathsf{data}}(\boldsymbol{f}) = \sum_{\theta \in \Theta} \left\| \boldsymbol{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{S}^{t_\theta} \cdot \boldsymbol{f} \right\|^2_{\mathbf{W}_\theta} , \qquad (3)$$

where $\boldsymbol{y}^\theta$ denotes the measured projections at orientation θ of the detector and time t_θ, Θ is the set of projection angles. The term $\mathbf{R}^\theta \cdot \mathbf{S}^{t_\theta} \cdot \boldsymbol{f}$ is the model of the data, operator $\mathbf{S}^{t_\theta}$ interpolates its arguments at time t_θ while operator $\mathbf{R}^\theta$ performs the projection for orientation θ (as further detailed below). The discrepancy between the data and its model is measured by the squared Mahalanobis distance:

$$\|\boldsymbol{u}\|^2_{\mathbf{W}_\theta} = \boldsymbol{u}^\top \cdot \mathbf{W}_\theta \cdot \boldsymbol{u}$$

with $\mathbf{W}_\theta$ a weighting matrix equal to the inverse of the covariance of the noise.

As already mentioned, we assume that the acquisition of a projection is instantaneous. This hypothesis motivates the decomposition of the dynamic projection of the 3D+t object $\boldsymbol{f}$ in two successive operations:

1) The operator $\mathbf{S}^{t_\theta}$ performs a temporal interpolation of the 3D+t voxels to extract the 3D object $f(\boldsymbol{x}, t_\theta)$ at the projection time t_θ. Our choice is justified by the fact that the temporal variation of the object is a continuous phenomenon. Hence it is possible to approximate the anatomical state of the object at any date by interpolating the frames sampled by the 4D voxels. Of course, the quality of the interpolation depends on the temporal sampling rate, as well as on the type of interpolation used.
Figure 3 shows the evolution of the root mean square (RMS) error between the data simulated in Section II-B and the output of our dynamic projection model as a function of the temporal sampling rate, for different interpolation methods. As expected, the error decreases when the number of frames increases. In 4D-CT, the selection of projections can be compared to using a nearest neighbor interpolation, which generates much more modeling errors than linear or B-spline-based interpolations as can be seen on the figure. Oscillations in the RMS error curves (notably for the nearest neighbor interpolation) correspond to the sampling rates where some frames are exactly synchronized with the actual projections. In these particular cases, $\mathbf{S}^{t_\theta}$ is just the identity operator and the remaining errors are due to the spatial sampling of $\boldsymbol{f}$ and to the approximations made by the projector $\mathbf{R}^\theta$. In our implementation, we use linear temporal interpolation as it gives the least approximation errors (see Fig. 3).
The coefficients of $\mathbf{S}^{t_\theta}$ are determined from a 1D temporal signal giving the periodic evolution of the motion, similar to the signals used for 4D-CT reconstructions. It is easy from this signal to extract the period of a cycle, and to identify the phase of each temporal frame. Then the projection dates t_θ give the locations of the interpolation points on the temporal axis, from which the coefficients of $\mathbf{S}^{t_\theta}$ are deduced. Since the period of the respiratory or cardiac motion is likely to vary from one cycle to another, the sequence $\boldsymbol{f}$ has to be normalized to a mean period. Hence the cycles in the 1D signal are virtually dilated or contracted to coincide with the normalized cycle and achieve a finer and more realistic calculation of the interpolation coefficients. A similar temporal registration method is used by Blondel *et al.* [4], [5] for cardiac SPECT. In our tests, we made the same simplification as Blondel *et al.* [4], [5], *i.e.* we have neglected the variation of the spatial amplitude of the motion from one cycle to another. Our results on empirical data discussed in sections IV-B and IV-C support the use of this simplifying hypothesis.
2) The operator $\mathbf{R}^\theta$ is the tomographic projector which performs a static projection of the 3D object interpolated at t_θ. $\mathbf{R}^\theta$ can be any numerical model implemented for iterative CT reconstruction. As already mentioned, projections must be separated according to the motion phase in order to reconstruct the temporal sequence. This implies a drastic reduction of the number of projections available per reconstructed temporal frame. In this context of reconstruction from a few projections, we have shown [22] that the accuracy of the projector is a critical issue to make the best of available data. In order to cope with such cases, we have developed an accurate and fast model. It exploits B-spline basis functions to represent the 3D object and to approximate their projections. This model will be designated henceforth under the name *spline driven* [22]. When reconstructing objects from a limited number of projections, *spline driven* has proven to be more accurate than classical projectors such as *distance driven* [10], with a satisfactory computional time. To better exploit dynamic data, we therefore used the *spline driven* projector to

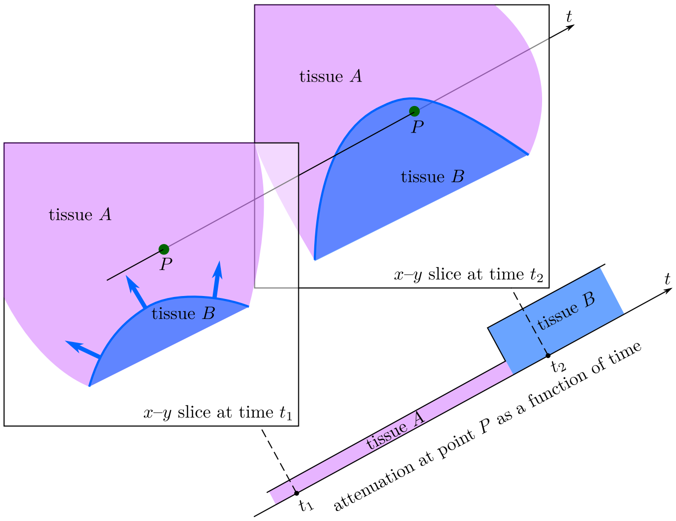

Fig. 4. A spatio-temporal representation of a situation where a same voxel P is "seen" at different times in two different neigbouring tissues : due to the motion of the interface between tissues A and B, point P is in tissue A at time t_1, and in tissue B at time t_2. The attenuation at point P then quickly changes between t_1 and t_2 when the boundary reaches P.

implement our dynamic reconstruction approach.

C. Spatio-temporal regularization

The term $\mathcal{J}_{\mathsf{prior}}$ in equation (2) is essential to regularize the otherwise ill-conditioned inverse problem of dynamic reconstruction. In contrast to gated CT reconstructions that are performed independently, the regularization $\mathcal{J}_{\mathsf{prior}}$ can enforce a temporal coherence in the reconstructions. The spatial regularization must be chosen so as to favor smooth areas while preserving discontinuities at the boundary of anatomical structures. Total variation [28] is typically used in this context. From the 3D representation of our 2D+t simulated phantom shown in Fig. 1, we observe that the variations of shapes are exclusively noticeable at the interfaces between the different parts of the object. For such an object, the spatial gradients with a high magnitude, identifying the separation between different tissues, are therefore intrinsically correlated to the temporal gradients out of which moving boundaries can be traced back. Figure 4 illustrates the case where a given spatial position sees two different tissues between time t_1 and time t_2. The temporal profile of attenuation sketched in Fig. 4 corresponds to a large motion where the change of the intensity value at position P is sharp. On the contrary, for sub-voxel motions, the sharp boundary will progressively cross the voxel and thus yield a smoother transition. Cardiac and respiratory movements also cause smooth spatial and temporal changes because they induce fluctuations of the absorption, particularly at the interfaces between different tissues, *e.g.* the lung walls or the myocardium's tissues. Consequently, we advocate the use of an $\ell_2 - \ell_1$ edge preserving regularization [7], [11], [9], [38] which is known to favor smoothness for small amplitude changes while preserving larger discontinuities. Being concerned with spatio-temporal (dis)continuity, we propose to use the following generalized form of relaxed TV:

$$\mathcal{J}_{\mathsf{prior}}(\boldsymbol{f}) = \sum\nolimits_{\boldsymbol{k},\ell}\Big(\epsilon^2 + \mu_{\mathsf{space}}^2 \left\|\nabla_{\boldsymbol{k},\ell}^{\mathsf{space}} \cdot \boldsymbol{f}\right\|^2 + \mu_{\mathsf{time}}^2 \left(\partial_{\boldsymbol{k},\ell}^{\mathsf{time}} \cdot \boldsymbol{f}\right)^2\Big)^{1/2}. \tag{4}$$

Here operators $\nabla_{\boldsymbol{k},\ell}^{\mathsf{space}}$ et $\partial_{\boldsymbol{k},\ell}^{\mathsf{time}}$ are respectively the 3D spatial gradient and the temporal derivative of the 3D+t object at the voxel $(\boldsymbol{k},\ell)$ with $\boldsymbol{k}$ and ℓ the spatial index and temporal index (i.e., the frame) of the voxel. In order to tune the strength of the regularization (relatively to the data fidelity term) and to account for the heterogeneity of the physical dimensions of the spatial gradient and of the temporal derivative, we introduce two different hyperparameters $\mu_{\mathsf{space}} \geq 0$ and $\mu_{\mathsf{time}} \geq 0$ in Eq. (4). Note that the ratio of these two hyperparameters is homogeneous to a velocity. In Eq. (4), the parameter $\epsilon > 0$ is not to be mistaken with a relaxation parameter introduced to avoid a singularity: it has to be appropriately tuned to set the trade-off between sharp and smooth changes.

Having a convex differentiable criterion (2), we perform the numerical minimization with the VMLM algorithm [23] which is a limited memory quasi-Newton optimization method with BFGS updates.

IV. Results

A. Numerical data

To test our approach, we reconstructed our dynamic Shepp-Logan phantom presented in section II-B, from a simulated set of 600 projections. We chose to sample the 2D+t sequence with 25 frames. In this configuration, the frames are exactly synchronized with the projections which, as explained below, lets us mimic a gated reconstruction. Figure 5 displays some frames of two different reconstructions from this dataset:

1) The sequence in the second column was reconstructed while imposing only the spatial regularization for each frame. This was achieved by setting $\mu_{\mathsf{time}} = 0$ in Eq. (4). Thanks to the perfect synchronisation of the frames with the projections, the result corresponds to an ideal *gated*-like reconstruction. The spatial hyperparameter $\mu_{\mathsf{space}} > 0$ was tuned to achieve the best visual quality for the result.
2) The sequence in the third column was reconstructed using the spatio-temporal regularization in Eq. (4) with $\mu_{\mathsf{time}} > 0$ and $\mu_{\mathsf{space}} > 0$ both tuned to improve the visual quality of the result.

To visualize the temporal continuity between the reconstructed frames, Fig. 6 displays spatio-temporal slices of these reconstructions (the green lines indicate the positions of the slices on the topmost panel of the figure).

In the first case, each frame has been reconstructed independently from the others because there was no temporal correlation imposed between frames in this configuration. A set of only 24 projections is available per frame, which explains the rather poor quality of the reconstruction. This is particularly noticeable on the fine moving structures which are perturbed by strong artifacts. We can observe on Fig. 6 the *decorrelation* between the frames, identified by small fluctuations of contrast from one frame to another in the uniform parts of the object.

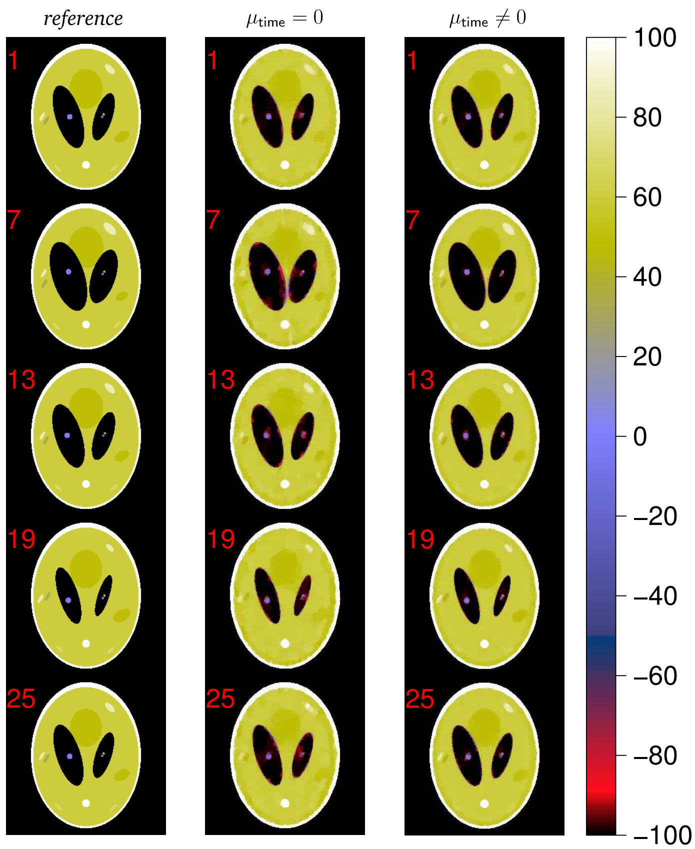

Fig. 5. 2D+t reconstructions (25 frames) of the dynamic Shepp-Logan phantom (values are given in Hounsfield unit). 24 projections are available per frame of the reconstructed cycle, with which they are synchronized (perfect *gated*-like case). The first column shows some reference frames identified by their number. The second column shows the reconstruction where no temporal regularization is applied ($\mu_{\text{time}} = 0$), *i.e.* the frames are independently reconstructed with a spatial regularization. The third column shows the reconstruction with a 3D (2D+t) spatio-temporal regularization ($\mu_{\text{time}} \neq 0$).

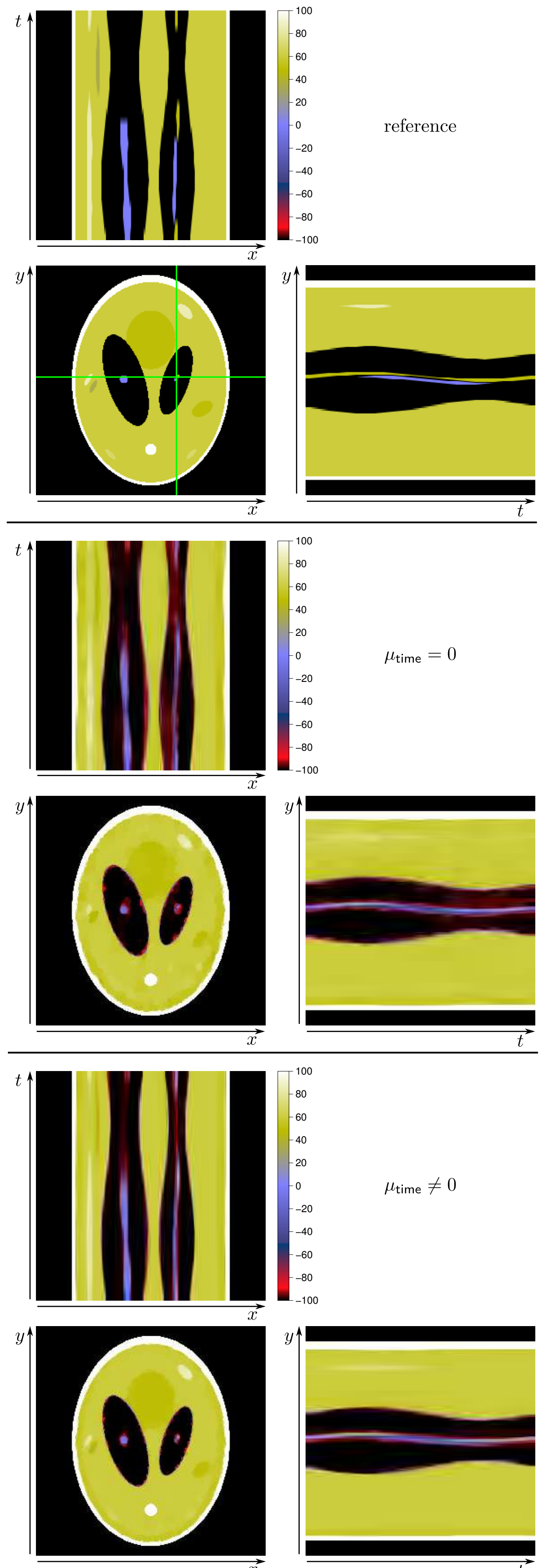

Fig. 6. Spatio-temporal slices of the 2D+t reconstructions of Fig. 5.

The second case, $\mu_{\text{time}} \neq 0$, shows the substantial gain brought by the temporal regularization of the 2D+t sequence in the quality of the reconstruction. The imposed continuity between frames helps to reduce artifacts due to the lack of available projections per frame. As a result, structures are better recovered, as well as their motion, in a non-ambiguous way. The gain is particularly clear on the spatio-temporal slices of Fig. 6 where the transitions between the frames are much smoother than in the *gated*-like case. Thanks to this improvement, fine nearby structures can be more easily discriminated without ambiguity.

B. Mechanical phantom data

We have reconstructed a mechanical phantom from a dataset acquired at the *Centre Léon Bérard*, on a scanner *Elekta Synergy Cone-Beam CT*. The flat panel is composed of a grid of 512×512 detector pixels of size $0.08 \times 0.08\,\text{cm}^2$. The acquisition process is similar to that described for CBCT scanners in section II-A.

The phantom has the features of a human thorax. Several zones of various densities mimic the lungs, the muscles and the spine. A small sphere of 2 cm in diameter is inserted

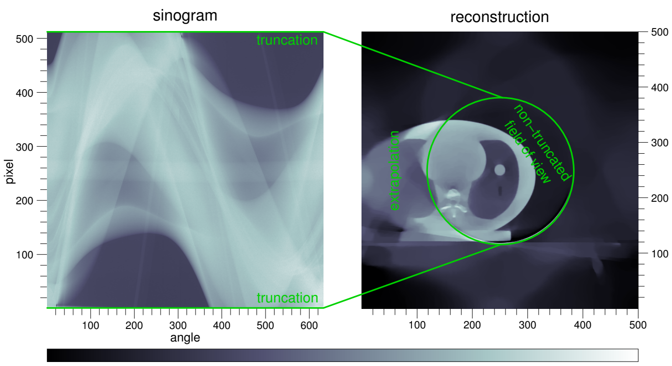

Fig. 7. 2D+t reconstruction of a mechanical phantom. *Left:* A sinogram extracted from the CB projections of the mechanical phantom. *Right:* The corresponding reconstruction. The disk's interior corresponds to the part of the object that is projected, hence "visible", on all the frames. A remarkable feature of our inverse approach is that it is able to reconstruct, in a rather satisfactory way, an "uncomplete" part of the data.

in the right lung. It can be mechanically animated with a periodic circular motion in the transversal plane, and a periodic translation in the cranio-caudal direction. The two motions can be combined, reproducing the trajectory of a tumor during the respiratory cycle.

1) 2D+t reconstructions: A set of 630 projections has been acquired regularly in time and angle on 360° during 116 seconds. The spherical insert was animated with a periodical motion in the transversal plane. The period of this "respiratory" cycle was 4 seconds. Thus, about 29 cycles are done during the acquisition, corresponding to about 22 projections per cycle.

The phantom was positioned so that the isocenter of the scanner is contained in the same plane as the trajectory of the spherical insert. Extracting the median detector lines corresponding to this plane from the CB projections, we obtained a fan beam sinogram from which we reconstructed a 2D+t sequence of the corresponding phantom's slice. A temporal frame of this reconstruction is shown in Fig. 7. The projections are clearly truncated by the detector. With iterative reconstruction methods, this problem can be partially addressed by reconstructing a larger volume of interest in order to encompass the whole object, the regularization helping to extrapolate regions with missing data. The static reconstruction displayed by Fig. 7 shows that, despite the truncation of the field of view, the object support is correctly restored even though there are some artifacts and a loss of details outside the fully seen region. However note that these defects have no impact on the quality of the reconstruction inside the fully seen region.

Our dynamic 2D+t reconstructions are presented on Fig. 8. As for the dynamic Shepp-Logan phantom, we have performed two reconstructions: a plain spatially regularized one (Fig. 8(e)), and a fully spatial and time regularized one (Fig. 8(f)). We chose to reconstruct a 2D+t sequence of 22 frames. The operator $\mathbf{S}^{t_\theta}$ of our direct model (see Section III-A) is implemented by a linear temporal interpolation. Hence, as the direct model links each object frame to two projections, about $29{\times}2$ projections per frame are available for the reconstruction. For comparisons, Fig. 8 also displays static reconstructions from a varying number of projections with the mechanical insert at still (a–c) and, to illustrate motion artifacts, from the set of dynamic projections (d). Thanks to the spatio-temporal regularization, the reconstructed frames of the dynamic reconstruction in Fig. 8(f) have the same visual quality as the reconstruction from 630 projections in Fig. 8(a), and are much better than the static reconstruction from 64 projections shown in Fig. 8(b) or from the *gated*-like reconstruction in Fig. 8(e). The motion of the insert is correctly recovered and the anatomic structures better reconstructed, in particular the spine.

2) 3D+t reconstructions: Figure 9 displays a 3D+t reconstruction of the phantom from another set of CB projections, acquired in the same conditions as previously. This time, the spherical insert was periodically animated with two kinds of motion: circular in the transversal plane and translational in the cranio-caudal direction. Hence the insert had a realistic 3D movement. We reconstructed a sequence of 22 frames made of $225{\times}225{\times}75$ voxels of size $2{\times}2{\times}2\,\mathrm{mm}^3$. To better visualize the dynamic aspect of our 3D+t reconstruction, we have uploaded a video clip of the sliced frames in the supplementary files (*phantom-dynamic.mp4*). Outside the field of view, the reconstruction appears corrupted by motion and truncation artifacts. However, inside the field of view, the quality of the reconstruction is comparable to that of our 2D+t reconstructions, thanks to the spatio-temporal regularization. To focus on the quality of the restored images, Fig. 9 only shows the non-truncated field of view. Transition from 2D+t to 3D+t involves no additional difficulties, but a bigger computional burden (many more variables, computations of 4D instead of 3D gradients in the regularization process).

C. Patient data

We have processed the same CBCT dataset of a patient's thorax as the one used in the article of Rit *et al.* [25]. The CB projections were acquired at the Centre Léon Bérard, (Lyon, France), on a scanner *Elekta Synergy Cone-Beam CT*. Thus the acquisition process is the same as for the mechanical phantom in section IV-B. In [25], the reconstructions were performed with a motion-compensated (MC) method developed by the authors and with a *gated* method. These reconstructions have been made available to us, in order to make a comparison with our method. We also obtained the recorded 1D temporal signal giving the periodic linear phase of the respiratory cycles, from which we have calibrated the temporal interpolator $\mathbf{S}^{t_\theta}$. The mean period of a cycle was evaluated to 2.4 s. We only kept the full cycles, reducing the number of required CB projections to 625. The duration of acquisition was 116 seconds, involving about 48 cycles, and 13 projections per cycle. We chose to reconstruct a 3D+t sequence of 13 frames. Using again linear interpolation for $\mathbf{S}^{t_\theta}$, about $48{\times}2$ projections were available per reconstructed frame.

Figure 10 displays the full 3D+t reconstruction. Each frame consists in $275{\times}200{\times}135$ voxels of $2{\times}2{\times}2\,\mathrm{mm}^3$. Inside the field of view, the thorax is correctly reconstructed, and one can identify the different anatomical structures: the right lung

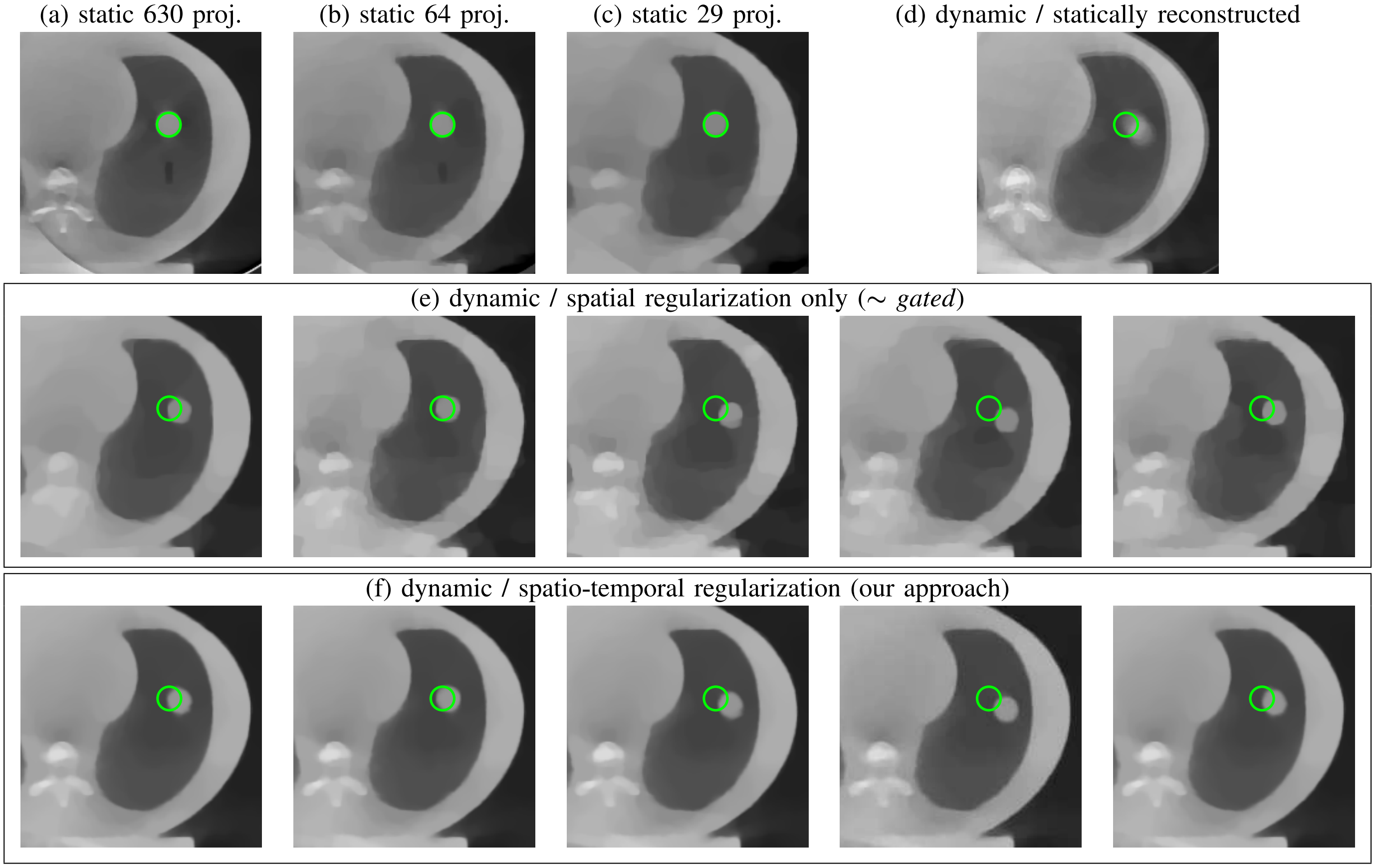

Fig. 8. 2D+t reconstructions of the median transversal slice of the mechanical phantom. Each restored sequence consists in 22 temporal frames with 500×500 pixels of size $1\times1\,\mathrm{mm}^2$ (the region of interest is zoomed in). (a,b,c): Static reconstructions from respectively 630, 64 and 29 projections of the static phantom (*i.e.*, the insert was motionless). (d) Static reconstruction of the phantom from the dynamic dataset. (e): Only spatially regularized reconstruction. (f): Spatio-temporally regularized reconstruction. The reference position of the spherical insert is identified on each frame by a small circle.

containing a tumor, its bronchi, and also the heart and the ribs. The respiratory motion is well recovered, particularly that of the diaphragm and of the tumor, without ambiguities nor noticeable artifacts. As for the mechanical phantom, a video clip of this 3D+t reconstruction has been uploaded in the supplementary files (*patient-dynamic.mp4*).

For comparison, Fig. 11 shows side by side the MC and gated reconstructions by Rit *et al.* in [25] and the result with our method. The end-exhale phase reconstructed by the MC method has a size of $262\times261\times132$, *i.e.* a voxel's size of $0.98\times0.98\times2\,\mathrm{mm}^3$. The *gated* 3D+$t$ sequence was reconstructed with the same parameters. Because only one frame is reconstructed by the MC method (the 3D+t sequence can be obtained by applying the motion model to the reference frame), we compare the reconstructed frame corresponding to the end-exhale phase. We have re-interpolated our reconstruction on the grid of finer voxels used by Rit *et al.*, using a cubic B-spline interpolation kernel.

Regarding the *gated* reconstruction, our method has eliminated the motion artifacts and better recovered the anatomical structures. Indeed *gated* methods have to be applied on extended datasets (*4D-CT*) to be efficient, which is not the case for CBCT and explains the poor quality of the *gated* reconstruction. It evidences the gain of our approach which achieves a good quality of reconstruction using only the CBCT dataset.

The MC reconstruction reveals a finer resolution of recovered structures than ours, which looks smoother. Note that MC requires an accurate motion model computed from a high resolution 4D-CT data set. Our method only needs the current set of CB projections to reconstruct the patient's thorax and the respiratory motion, recovering the motion as well as the static structures without ambiguity or motion artifact. As a result, lower X-ray doses are necessary with our method than with the MC method.

V. Conclusion

This paper described a new approach for dynamic CT reconstruction without motion compensation. Our motivation was to perform 3D+t reconstructions from a limited number of projections, such as provided by a CBCT system, without resorting to additional 4D-CT data. Rather than building an explicit model of the motion of human tissues, we addressed the full 3D+t reconstruction problem, i.e., we reconstruct sequences of 3D images. Our results on CBCT data show that satisfying reconstructions can be obtained from limited data sets using a simple prior model based on spatial and temporal coherence.

Reconstruction of a 3D+t sequence from limited data is possible by considering jointly all projections in a global

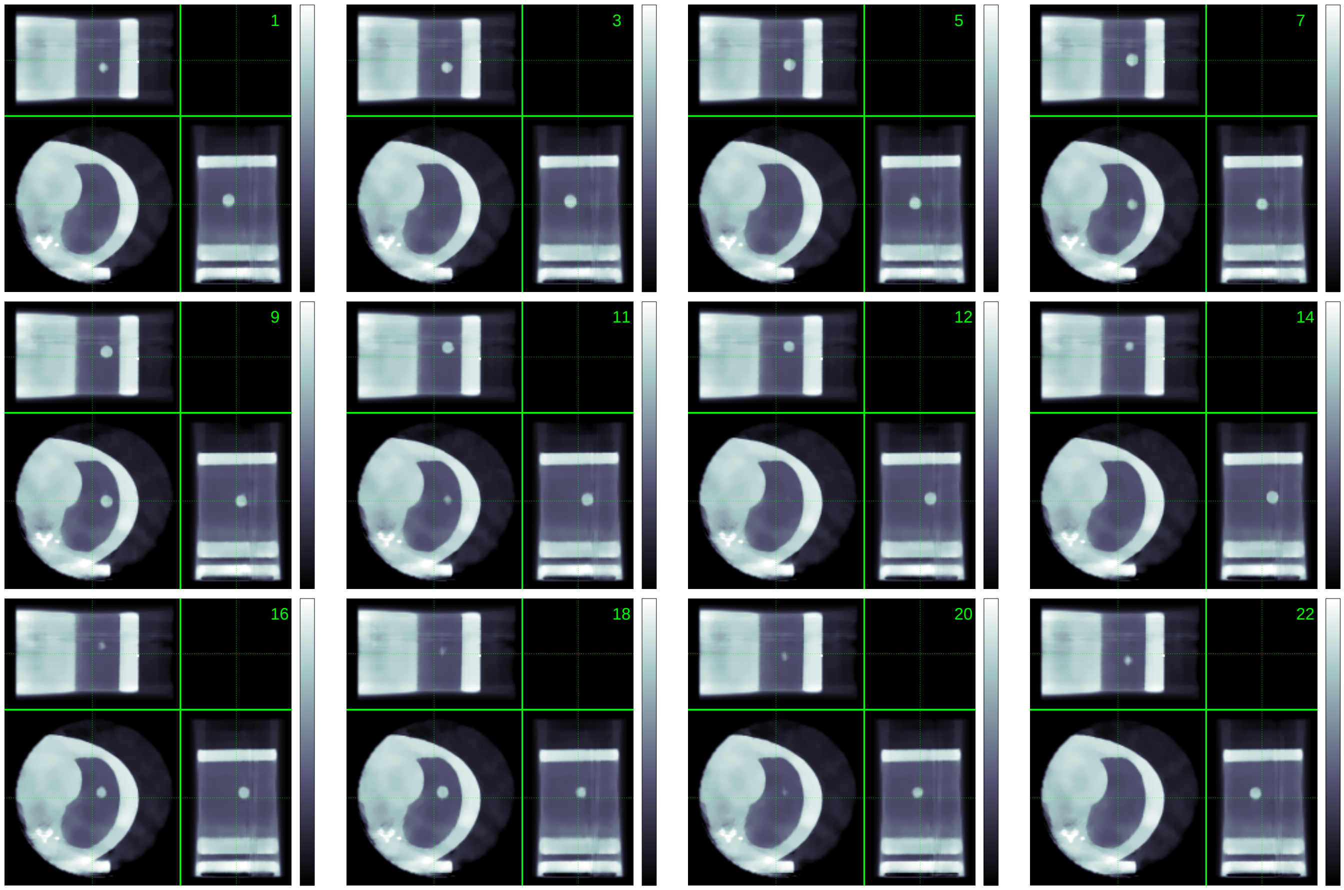

Fig. 9. 3D+t reconstruction of the mechanical phantom from the CB projections. A sliced view of each frame in the axial (center), sagittal (right), and coronal (top) directions is displayed. Only the non-truncated field of view is shown. A MPEG-4 video clip of these slices of the reconstructed frames is available as a supplementary file (*phantom-dynamic.mp4*).

inverse problem. We carefully designed the model of projections, based on spatial and temporal interpolation. Images are represented on quasi-isotropic B-spline functions that can be accurately and efficiently projected. Temporal frames are linearly interpolated to match phases of the cyclic movement under reconstruction. A joint spatio-temporal regularization exploits spatial and temporal coherence of the dynamic volume to provide images with smooth regions and motion. To preserve sharp boundaries and prevent from blurring motion, we proposed a relaxed 3D+t total-variation regularization.

Our method has been applied to dynamic CBCT reconstruction problems. Reconstructions of a dynamic phantom and of patient data both confirmed that our method can produce 3D+t reconstructions using only CBCT data, with a quality significantly improved compared to gated reconstructions. The motion of a lung tumor can be clearly identified in our reconstruction, which is confirmed by comparison with a reconstruction obtained using a state-of-the-art motion compensation method based on an accurate estimation of motion from high-resolution 4D-CT images. The reconstruction quality does not match that achieved using accurate motion models but may be very useful when high-resolution dynamic data sets are not available.

At present, we are working on refined registration to account for variability of the motion between cycles (differences in motion amplitude); in the future, one of our main goals is to achieve better final reconstruction of the spatio- temporal image, by alternating reconstruction and registration.

By avoiding the use of additional 4D-CT data in dynamic CBCT reconstructions, we believe that the approach we followed will help in the future to reduce the X-ray dose delivered to the patient.

Acknowledgments

The empirical data were kindly provided by Simon Rit from the *Centre Léon Bérard* (Lyon, France). Our algorithms were developped in Yorick [http://yorick.sourceforge.net/index.php], a freely available data processing language provided by David Munro. Computations for the 3D+t reconstructions were carried out on the Horizon cluster [http://www.projet-horizon.fr].

References

[1] A. Andersen and A. Kak. *Simultaneous Algebraic Reconstruction Technique (SART): a superior implementation of the ART algorithm*. Ultrasonic Imaging, 1984. 2

[2] J. Balter, R. Ten Haken, T. Lawrence, K. Lam and J. Robertson. *Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing*. International Journal of Radiation Oncology Biology Physics, 1996. 1

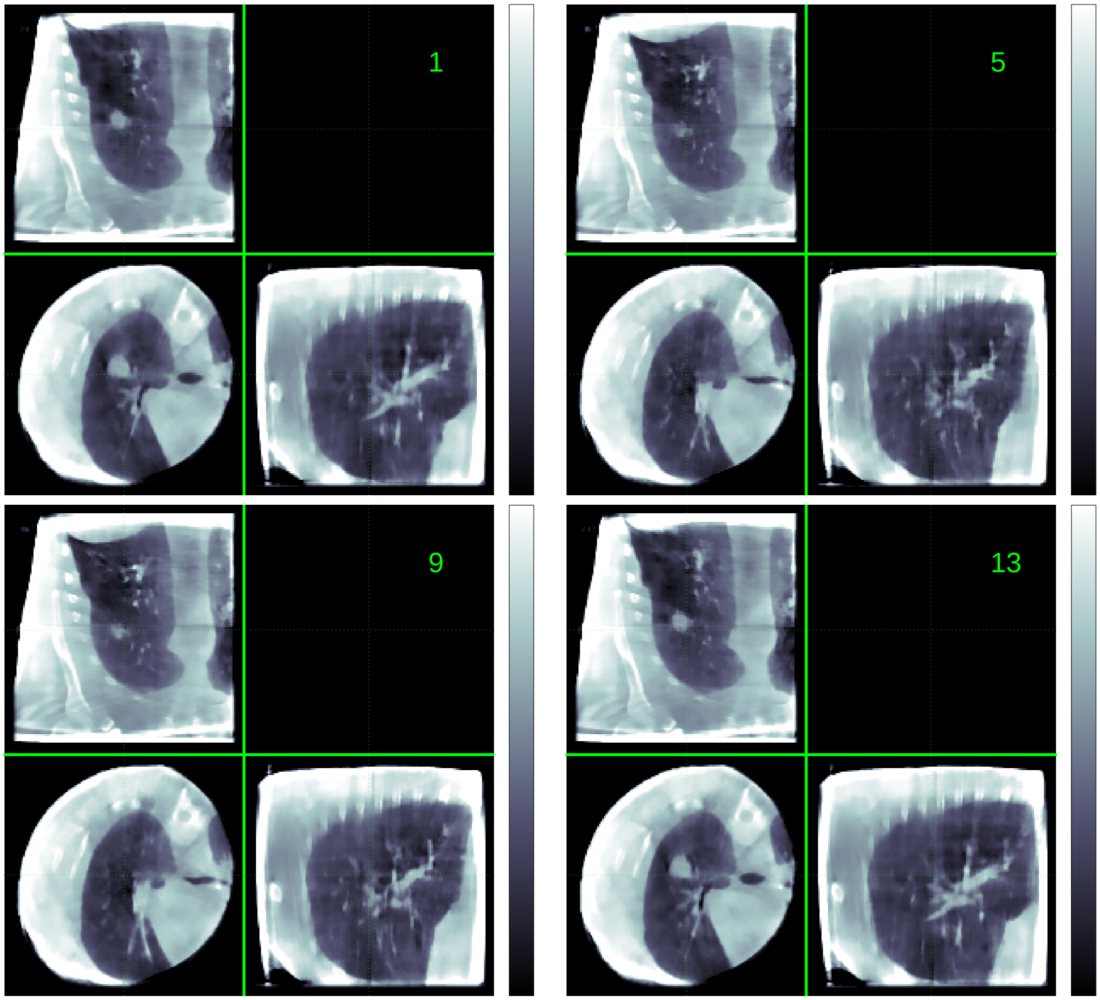

Fig. 10. Our 3D+t reconstruction of the thorax of a patient (*Elekta Synergy Cone-Beam CT* data from [25], used by courtesy of Centre Léon Bérard) A sliced view of each frame in the axial (center), sagittal (right), and coronal (top) directions is shown. A MPEG-4 video clip of these slices of the reconstructed frames is available as a supplementary file (*patient-dynamic.mp4*).

(a) Rit *et al.* *gated* (from a **single** scan) (b) Rit *et al.* MC (from **two** scans) (c) Our method (from a **single** scan)

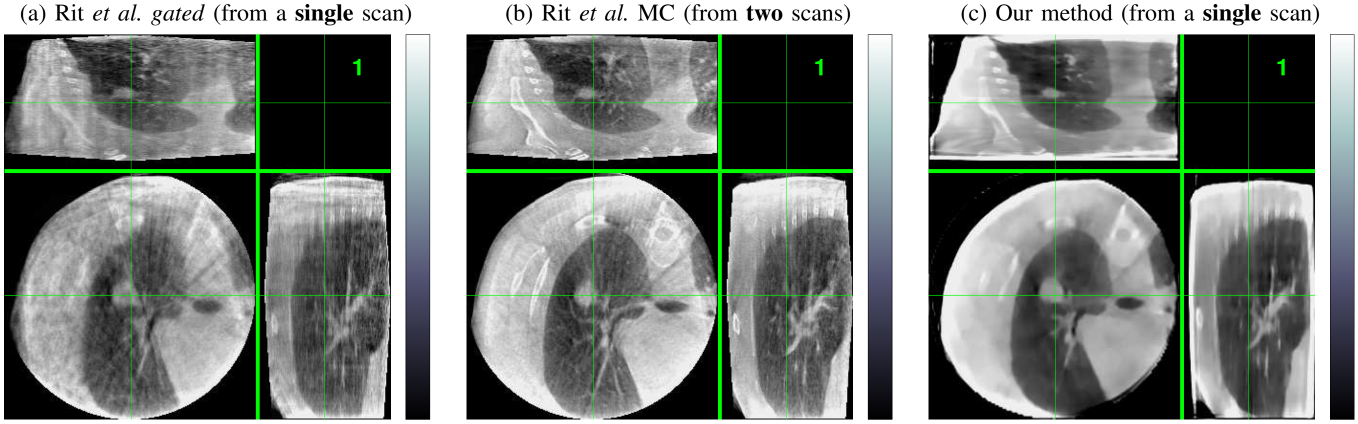

Fig. 11. Comparison of the reconstructions from the patient's dataset performed by a *gated method* (a) and the MC method of Rit *et al.* (b), with the reconstruction by our spatio-temporally regularized method (c). The end-exhale phase is shown. For better readability, only the fully seen field of view is displayed.

[3] R. Berbeco, H. Mostafavi, G. Sharp and S. Jiang. *Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers*. Physics in Medicine and Biology, 2005. 1

[4] C. Blondel, R. Vaillant, G. Malandain and N. Ayache. *4D deformation field of coronary arteries from monoplane rotational X-ray angiography*. International Congress Series, 2003. 3

[5] C. Blondel, R. Vaillant, G. Malandain and N. Ayache. *3D tomographic reconstruction of coronary arteries using a precomputed 4D motion field*. Physics in Medicine and Biology, 2004. 1, 3

[6] J. Brankov, Y. Yang, M. Narayanan and M. Wermck. *Motion-compensated 4D processing of gated SPECT perfusion studies*. IEEE Nuclear Science Symposium Conference Record, 2002. 1

[7] P. Charbonnier, L. Blanc-Féraud, G. Aubert and M. Barlaud. *Deterministic edge-preserving regularization in computed imaging*. IEEE Transactions on Image Processing, 1997. 4

[8] G. Chen, J. Kung and K. Beaudette. *Artifacts in computed tomography scanning of moving objects*. Seminars in Radiation Oncology, 2004. 1

[9] W. Chlewicki, F. Hermansen and S. Hansen. *Noise reduction and convergence of Bayesian algorithms with blobs based on the Huber function and median root prior*. Physics in Medicine and Biology, 2004. 4

[10] B. DeMan and S. Basu. *Distance driven projection and backprojection in three dimensions*. Physics in Medicine and Biology, 2004. 3

[11] I. Elbakri and J. Fessler. *Statistical image reconstruction for polyenergetic X-ray computed tomography*. EEE Transactions on Medical Imaging, 2002. 4

[12] L. Feldkamp, L. Davis and J. Kress. *Practical cone-beam algorithm*. Journal of the Optical Society of America A, 1984. 2

[13] E. Ford, G. Mageras, E. Yorke and C. Ling. *Respiration-correlated spiral CT: a method of measuring respiratory-induced anatomic motion for radiation treatment planning*. Medical Physics, 2003. 1

[14] H. Goldberg, R. Gould, I. Feuerstein, J. Sigeti and M. Lipton. *Evaluation of ultrafast CT scanning of the adult abdomen*. Investigative Radiology, 1989. 1

[15] D. Jaffray, D. Drake, M. Moreau, A. Martinez and J. Wong. *A radiographic and tomographic imaging system integrated into a medical linear accelerator for localization of bone and soft-tissue targets*. International Journal of Radiation Oncology Biology Physics, 1999. 2

[16] D. Jaffray and J. Siewerdsen. *Cone-beam computed tomography with a flat-panel imager: initial performance characterization*. Medical Physics, 2000. 2

[17] D.A. Jaffray, J.H. Siewerdsen, J.W. Wong and A.A. Martinez. *Flat-panel cone-beam computed tomography for image-guided radiation therapy*. International journal of Radiation Oncology Biology Physics, 2002. 2

[18] A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, 1988. 2

[19] P. Keall, G. Starkschall, H. Shukla, K. Forster, V. Ortiz, C. Stevens, S. Vedam, R. George, T. Guerrero and R. Mohan. *Acquiring 4D thoracic CT scans using a multislice helical method*. Physics in Medicine and Biology, 2004. 1

[20] T. Li, E. Schreibmann, Y. Yang and L. Xing. *Motion correction for improved target localization with on-board cone-beam computed tomography*. Physics in Medicine and Biology, 2006. 1, 2

[21] D. Low, M. Nystrom, E. Kalinin, P. Parikh, J. Dempsey, J. Bradley, S. Mutic, S. Wahab, T. Islam and G. Christensen. *A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing*. Medical Physics, 2003. 1

[22] F. Momey, L. Denis, C. Burnier, É. Thiébaut, J.M. Becker and L. Desbat. *Spline driven: high accuracy projectors for tomographic reconstruction*. Manuscript in preparation, 2013 [http://hal.archives-ouvertes.fr/hal-00990015]. 3

[23] J. Nocedal. *Updating quasi-Newton matrices with limited storage*. Mathematics of computation, 1980. 4

[24] E. Rietzel, T. Pan and G. Chen. *Four-dimensional computed tomography: image formation and clinical protocol*. Medical Physics, 2005. 1

[25] S. Rit, D. Sarrut and L. Desbat. *Comparison of analytic and algebraic methods for motion-compensated cone-beam CT reconstruction of the thorax*. IEEE Transactions on Medical Imaging, 2009. 1, 2, 6, 7, 9

[26] S. Rit, J. Wolthaus, M. Van Herk and J. Sonke. *On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion*. Medical Physics, 2009. 1, 2

[27] C. Ritchie, J. Godwin, C. Crawford, W. Stanford, H. Anno, and K. Yongmin. *Minimum scan speeds for suppression of motion artifacts in CT*. Radiology, Radiological Society of North America, 1992. 1

[28] L.I. Rudin, S. Osher and E. Fatemi. *Nonlinear total variation based noise removal algorithms*. Physica D: Nonlinear Phenomena, 1992. 2, 4

[29] W. Segars, D. Lalush and B. Tsui. *A realistic spline-based dynamic heart phantom*. IEEE Transactions on Nuclear Science, 1999.

[30] W. Segars, D. Lalush and B. Tsui. *Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms*. IEEE Transactions on Nuclear Science, 2001.

[31] W. Segars, D. Lalush and B. Tsui. *4D XCAT phantom for multimodality imaging research*. Medical physics, 2010.

[32] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. Van Herk, J. Lebesque and K. Miyasaka. *Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy*. International Journal of Radiation Oncology Biology Physics, 2002. 1

[33] L. Shepp and B. Logan. *The Fourier reconstruction of a head section*. IEEE Transactions on Nuclear Science, 1974. 2

[34] S. Shimizu, H. Shirato, K. Hagei, T. Nishioka, X. Bo, H. Dosaka-Akita, S. Hashimoto, H. Aoyama, K. Tsuchiya and K. Miyasaka. *Impact of respiratory movement on the computed tomographic images of small lung tumors in three-dimensional (3D) radiotherapy*. International Journal of Radiation Oncology Biology Physics, 2000. 1

[35] J.J. Sonke, L. Zijp, P. Remeijer and M. Van Herk. *Respiratory correlated cone beam CT*. Medical Physics, 2005. 1

[36] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM Pub., 2005. 2

[37] S. Vedam, P. Keall, V. Kini, H. Mostafavi, H. Shukla and R. Mohan. *Acquiring a four-dimensional computed tomography dataset using an external respiratory signal*. Physics in Medicine and Biology, 2003. 1

[38] J. Wang, T. Li and L. Xing. *Iterative image reconstruction for CBCT using edge-preserving prior*. Medical Physics, 2009. 4

[39] P. Willis and Y. Bresler. *Optimal scan for time-varying tomography. I. Theoretical analysis and fundamental limitations*. IEEE Transactions on Image Processing, 1995. 1

[40] P. Willis and Y. Bresler. *Optimal scan for time-varying tomography. II. Efficient design and experimental validation*. IEEE Transactions on Image Processing, 1995. 1

[41] R. Zeng, J. Fessler and J. Balter. *Estimating 3-D respiratory motion from orbiting views by tomographic image registration*. IEEE Transactions on Medical Imaging, 2007. 2