

# Shearlet-based regularization in sparse dynamic tomography

T.A. Bubba<sup>a</sup>, M. März<sup>b</sup>, Z. Purisha<sup>a</sup>, M. Lassas<sup>a</sup>, and S. Siltanen<sup>a</sup>

<sup>a</sup>Department of Mathematics and Statistics, University of Helsinki, Gustaf Hällströmin katu 2, 00014 Helsinki, Finland

<sup>b</sup>Institut für Mathematik, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

## ABSTRACT

Classical tomographic imaging is soundly understood and widely employed in medicine, nondestructive testing and security applications. However, it still offers many challenges when it comes to dynamic tomography. Indeed, in classical tomography, the target is usually assumed to be stationary during the data acquisition, but this is not a realistic model. Moreover, to ensure a lower X-ray radiation dose, only a sparse collection of measurements per time step is assumed to be available. With such a set up, we deal with a sparse data, dynamic tomography problem, which clearly calls for regularization, due to the loss of information in the data and the ongoing motion. In this paper, we propose a 3D variational formulation based on 3D shearlets, where the third dimension accounts for the motion in time, to reconstruct a moving 2D object. Results are presented for real measured data and compared against a 2D static model, in the case of fan-beam geometry. Results are preliminary but show that better reconstructions can be achieved when motion is taken into account.

**Keywords:** tomography, dynamic tomography, sparse tomography, shearlets, regularization

## 1. INTRODUCTION

Computed tomography (CT) allows a non-invasive recovery of the internal structure of an object by recording the attenuation of the X-rays which are propagated through the object from multiple views. The inverse problem of reconstructing the X-rays attenuation (and thus the object structure) from the recorded measures, or *projections*, is well understood when the object is static and a full collections of projections (which corresponds to a dense sampling of the projection views or angles) is available. This “classical” setup is nowadays considered a mature field and, since the beginning of its development during the 1960s, gave rise to a deluge of applications in medicine, nondestructive testing and security application.

However, in many biomedical applications, the target is non-stationary and, in general, projection measurements are time-dependent: namely, the target changes in time between the recording of two consecutive projection images. This setting is usually referred to as *dynamic tomography*. Motion can be periodic, like in the case of the beating heart, or non-periodic, as the flow of a contrast agent in the blood vessels. Regular CT devices are already used dynamically, but the motion is usually not taken into account during the reconstruction task, especially when the changes are non-periodic. Indeed, *ad hoc* solutions exist for very slow or periodic movements: for instance, a heart can be satisfactorily imaged using gating,<sup>1</sup> and there exists techniques based on motion compensation.<sup>2–4</sup> In both cases, the result is the reconstruction of a static CT image from dynamic data. Moreover, many existing techniques for dynamic tomography are based on filtered back-projection algorithms, which requires a dense angular sampling. However, X-rays imaging comes with health hazards and, in the last years, lowering the X-rays exposure for the patients has been a major concern, if not a constraint due to realistic measurement protocols.

---

Further author information: (Send correspondence to T.A.B.)

T.A.B.: E-mail: [tatiana.bubba@helsinki.fi](mailto:tatiana.bubba@helsinki.fi), Telephone: +358 294151492

M.M.: E-mail: [maerz@math.tu-berlin.de](mailto:maerz@math.tu-berlin.de), Telephone: +49 3031427321

Z.P.: E-mail: [zenith.purisha@helsinki.fi](mailto:zenith.purisha@helsinki.fi), Telephone: +358 294151492

M.L.: E-mail: [matti.lassas@helsinki.fi](mailto:matti.lassas@helsinki.fi), Telephone: +358 505674417

S.S.: E-mail: [samuli.siltanen@helsinki.fi](mailto:samuli.siltanen@helsinki.fi), Telephone: +358 294151420

In this paper we propose a variational model for the reconstruction of a moving object that takes into account the space-time dynamics, without any assumption on the periodicity of the motion and in the case of sparsely sampled angular steps. The key of the proposed approach is to use 3D shearlets,<sup>5</sup> a recently proposed directional-aware representation systems for multidimensional data, to regularize the the ill-posedness deriving from the loss of information in the data and the ongoing motion. The resulting model is the following rather classical constrained analysis formulation:

$$\operatorname{argmin}_{\mathbf{f} \in \mathbb{R}_+^n} \frac{1}{2} \|\mathbf{A}\mathbf{f} - \mathbf{y}\|_2 + \lambda \|\Phi\mathbf{f}\|_1 \quad (1)$$

where  $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$ ,  $\mathbf{f}$  is the object to be reconstructed,  $\mathbf{A}$  is the measurement matrix,  $\mathbf{y}$  the measured data (or *sinogram*),  $\Phi$  the shearlet system, and  $\lambda$  is the regularization parameter. However, instead of considering a 2D shearlet system for each time frame of the acquisition, we explicitly take into account the motion using a 3D shearlet system where the third dimension accounts for the evolution in time, yielding a (2+1)-dimensional model. The rationale behind this model is that a continuous evolution of a cartoon-like object is well suited for the use of 3D shearlets. The solution of (1) is addressed by means of the variable metric inexact line-search algorithm (VMILA),<sup>6</sup> a very recently proposed proximal-gradient method, which enables the inexact computation of the proximal point defining the descent direction. Computational experiments are performed on both simulated and real data, the latter measured with a custom-built  $\mu$ CT system at the University of Helsinki<sup>7</sup> (Finland).

Sparse dynamic CT is lately receiving increasing attention (see 8,9 and the reference therein), but the use of shearlets to face it has never been investigated before. For static tomography,  $\ell^1$ -priors have been widely investigated, and were used for the first time on X-ray measurements with real data in 10,11. The results, compared against a 2D static model, are preliminary but show that better reconstructions can be achieved when motion is taken into account by including time as third dimension of the problem.

The remainder of this paper is organized as follows. In Section 2 we introduce a nonsmooth convex constrained minimization formulation for the 3D sparse dynamic CT problem in the discrete setting. The solution of such minimization problem is addressed in Section 3. The results of the proposed method are then presented in Section 4. Lastly, we draw some conclusions and we give an outlook to future work in Section 5.

## 2. A 3D SHEARLET-BASED MODEL FOR SPARSE DYNAMIC CT

In tomographic imaging, X-rays are emitted from a source (rotating on a circular trajectory around the target) from multiple directions of view (or angles). The rays are propagated through the object, being partially absorbed, and then recorded by a detector. The resulting data can be interpreted as a collection of line integrals of an unknown attenuation function  $f$ . Mathematically, this is modeled with the notion of the Radon transform:<sup>12</sup>

$$(\mathcal{R}f)(\omega, \xi) = \int_{\ell(\omega, \xi)} f(x) dx,$$

where  $\ell(\omega, \xi)$  denotes a ray and  $(\omega, \xi) \in [0, 2\pi) \times \mathbb{R}$ . In *dynamic* X-ray tomography the attenuation function  $f(x, t)$  is also a function of the time  $t$ :

$$(\mathcal{R}f)(\omega, \xi, t) = \int_{\ell(\omega, \xi)} f(x, t) dx. \quad (2)$$

Similarly to the static case, equation (2) can be discretized leading to the following system of linear equations:

$$\mathbf{A}\mathbf{f} = \mathbf{y} \quad (3)$$

where  $\mathbf{f} \in \mathbb{R}^n$  is a discrete representation of the object to be imaged,  $\mathbf{y} \in \mathbb{R}^m$  is a vector of the measured data (or *sinogram*) taken at given time instances  $\tau \in \mathbb{N}$  in a fixed time period  $[0, T]$ , with  $T \in \mathbb{N}$  being the number of time frames, and  $\mathbf{A} \in \mathbb{R}^{m \times n}$  describes the measurement process (dependent on time). The discretization of  $\mathbf{f} = \{\mathbf{f}_\tau\}_{\tau=0}^T$  is obtained by stacking column-wise each target frame  $\mathbf{f}_\tau \in \mathbb{R}^{N \times N}$  at time instance  $\tau$ , namely

$n = N^2T$  where  $N$  is the resolutions in pixels at each time frame. Similarly,  $\mathbf{y} = \{\mathbf{y}_\tau\}_{\tau=0}^T$  is obtained by column-wise stacking each sinogram frame  $\mathbf{y}_\tau \in \mathbb{R}^{N_\theta \times N_{\text{dte}}}$  at time instance  $\tau$ , *i.e.*,  $m = N_\theta N_{\text{dte}}T$ , where  $N_\theta$  is the number of projection views and  $N_{\text{dte}}$  the number of detector cells. The matrix  $\mathbf{A}$  is understood as a block diagonal matrix, where each block  $\mathbf{A}_\tau \in \mathbb{R}^{N_\theta N_{\text{dte}} \times N^2}$  describes the measurement process at time  $\tau$ :

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_0 & & & \\ & \mathbf{A}_1 & & \\ & & \ddots & \\ & & & \mathbf{A}_T \end{pmatrix}. \quad (4)$$

It is well known that CT is an ill-posed problem, leading to ill-condition in the discrete setting. In the following we are also assuming a sparse angular subsampling (namely,  $N_\theta$  is of the order of a few dozen) leading to a highly ill-posed discrete inverse problem. Under this condition, the linear system in (3) has infinitely many solutions, which can be also be dominated by noise, due to the ill-conditioning. Hence, regularization strategies are necessary. A classical strategy is to consider a functional of the form

$$G(\mathbf{f}) = G_0(\mathbf{f}) + G_1(\mathbf{f}) \quad (5)$$

where  $G_0$  is the data mismatch term, and  $G_1$  includes *a priori* information on the solution. The expression for  $G_0$  depends on the nature of the noise on the data. In CT, noise is usually the result of the combination of a Poisson and a Gaussian process. However, usually the mean for the Poisson process is large enough to approximate it with a Gaussian process. By using a maximum likelihood approach, this leads to the least squares functional:

$$G_0(\mathbf{f}) = \frac{1}{2} \|\mathbf{A}\mathbf{f} - \mathbf{y}\|_2^2. \quad (6)$$

On the other hand,  $G_1$  works as a regularization term, namely it should account for the ill-posedness of the problem by incorporating *a priori* information. A well-known approach to denoising is to consider the  $\ell_1$ -norm of a sparsifying transform. In our case, we are not only interested in denoising, but also in accounting for the ongoing motion. To this end, we consider shearlets, due to their localization and directional properties, which allow for a better understanding of multidimensional phenomena. In  $G_1$ , we consider also the prior given by the physics of the problem, that is, the knowledge of the nonnegativity of the solution  $\mathbf{f}$  ( $\mathbf{f}$  is a measure of attenuation and, hence, measures the incoming “number of photons”). Thus, our proposed regularization term reads as:

$$G_1(\mathbf{f}) = \lambda \|\Phi\mathbf{f}\|_1 + \iota_{\mathbb{R}_+}(\mathbf{f}),$$

where  $\lambda$  is the regularization parameter,  $\Phi$  is the matrix of the underlying 3D shearlet transform and  $\iota_{\mathbb{R}_+}$  is the indicator function of the nonnegative orthant. All together, this leads to the constrained minimization problem (1):

$$\operatorname{argmin}_{\mathbf{f} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{A}\mathbf{f} - \mathbf{y}\|_2^2 + \lambda \|\Phi\mathbf{f}\|_1 + \iota_{\mathbb{R}_+}(\mathbf{f}).$$

In the following, we will consider (and compare) two different scenarios:

- **Static – 2D:** this case corresponds to reconstruct each time frame as a 2D static problem, and motion is not taken into account at all. Hence, the measurements matrix  $\mathbf{A}$  is constructed by stacking *the same* 2D fan-beam measurements matrix  $\mathbf{A}_\tau$  always obtained by sampling the same  $N_\theta$  angles starting from angle 1. Correspondingly, we consider a 2D shearlet system, namely we consider the same 2D shearlet system for each time frame  $\tau$ .
- **Dynamic – 3D:** in this case, we explicitly account for the motion. Hence, the measurements matrix  $\mathbf{A}$  is obtained by stacking  $T$  2D fan-beam measurement matrices  $\mathbf{A}_\tau$ , each obtained with a shift of 1 degree (per frame  $\tau$ ) for the angles sampling. Namely, if for the first frame we consider the  $N_\theta$  angles starting from angle 1, for the second frame we consider  $N_\theta$  angles starting from angle 2, and so on, until the last block matrix  $\mathbf{A}_T$  where the angle sampling starts from angle  $T$ . Correspondingly, we consider a 3D shearlet system, where shearlets are understood as a tool for regularizing the motion (and the noise), given that they “act” also in the third dimension.

---

**Algorithm 1** Variable Metric Inexact Line-search Algorithm

---

Choose  $0 < \alpha_{\min} \leq \alpha_{\max}$ , set  $\beta, \gamma \in (0, 1)$ ,  $\eta \in (0, 1]$  and  $\mathbf{f}^{(0)} \geq 0$ .  
**for**  $k = 0, 1, 2, \dots$  **do**  
  Step 1. Choose the steplength  $\alpha_k \in [\alpha_{\min}, \alpha_{\max}]$  and the scaling matrix  $D_k$ ;  
  Step 2. Proximal approximation: compute the dual variable  $\mathbf{v}^{(k)}$  such that  
    
$$h(\mathbf{v}^{(k)} - \alpha_k D_k \mathbf{B}^T \mathbf{v}^{(k)}, \mathbf{f}^{(k)}) \leq \eta H(\mathbf{v}^{(k)}, \mathbf{f}^{(k)})$$
  
    and set  $\tilde{\mathbf{f}}^{(k)} = \mathbf{v}^{(k)} - \alpha_k D_k \mathbf{B}^T \mathbf{v}^{(k)}$ ;  
  Step 3. Compute the descent direction  $\mathbf{d}^{(k)} = \tilde{\mathbf{f}}^{(k)} - \mathbf{f}^{(k)}$ ;  
  Step 4. Modified Armijo linesearch:  
    **if**  $\Gamma(\mathbf{f}^{(k)} + \lambda_k \mathbf{d}^{(k)}) \leq \Gamma(\mathbf{f}^{(k)}) + \beta \lambda_k h(\mathbf{v}^{(k)}, \mathbf{f}^{(k)})$  **then**  
      go to Step 5.  
    **else**  
       $\lambda_k = \gamma \lambda_k$  and go to Step 4.  
  Step 5. Set  $\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \lambda_k \mathbf{d}^{(k)}$ .  
**end for**

---

Alongside these two different scenarios, we investigated also two “mixed” approaches, namely:

- **semi-static**: in this case, we combine the measurements matrix  $\mathbf{A}$  from the Dynamic – 3D approach with the 2D shearlet system from the Static – 2D scenario.
- **semi-dynamic**: in this case, we combine the measurements matrix  $\mathbf{A}$  from the Static – 2D approach with the 3D shearlet system from the Dynamic – 3D scenario.

For the sake of brevity, we do not report the results concerning these mixed approaches, but preliminary results show that using a 3D shearlet system is usually better than a 2D one, and also the 3D geometry helps improving the recovery.

### 3. A PROXIMAL-GRADIENT METHOD FOR DYNAMIC CT

For the solution of the optimization problem (1) we use the recently proposed variable metric inexact line-search algorithm (VMILA),<sup>6</sup> whose step-by-step description is outlined in Algorithm 1.

It is a proximal-gradient method suitable for minimizing the sum of a differentiable (possibly nonconvex) function plus a convex (possibly nondifferentiable) function. Note that these assumptions are satisfied in our model, where the differentiable part consists of the data mismatch term  $G_0$  (equation (5)) and the nondifferentiable contribute is given by sum  $G_1$  (equation (6)). The  $(k + 1)$ -th iteration of VMILA is given by:

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \lambda_k \left\{ \text{prox}_{\alpha_k G_1}^{D_k^{-1}} \left( \mathbf{f}^{(k)} - \alpha_k D_k \nabla G_0(\mathbf{f}^{(k)}) \right) - \mathbf{f}^{(k)} \right\}$$

where  $\lambda_k$  is a linesearch parameter, chosen accordingly to a generalized Armijo-like backtracking procedure, which ensures global convergence;  $\alpha_k \in [\alpha_{\min}, \alpha_{\max}] \subset \mathbb{R}^+$  is a positive steplength;  $D_k$  is a symmetric and positive definite scaling matrix such that all its eigenvalues are contained in an interval  $\left[ \frac{1}{L_k}, L_k \right]$  where  $L_k > 1$  and  $\sum_{k=0}^{\infty} (L_k^2 - 1) < \infty$ , and  $\text{prox}_{\alpha_k G_1}^{D_k^{-1}}$  denotes the proximal operator associated to  $G_1$  in the metric induced by the scaling matrix  $D_k$ . In particular, we select  $\alpha_k$  through an adaptive strategy based on the Barzilai-Borwein updating rules<sup>13,14</sup> and  $D_k$  according to a split-gradient idea<sup>15</sup> based on the decomposition of the gradient of  $G_0$  into a positive and a negative part and we adopt for  $L_k$  the following updating rule<sup>16</sup>  $L_k = \sqrt{1 + L/k^2}$ , where  $L$  is a suitable initial positive threshold.



(a) First frame.

(b) Middle frame.

(c) Last frame.

Figure 1: Pictures of three frames of the Emoji data set acquired at the University of Helsinki (Finland).

In many common situations, as in the case of (1), a closed-form of the proximal operator of  $G_1$  is not known. In this case, if the nondifferentiable term  $G_1$  can be expressed as

$$G_1(\mathbf{f}) = g(\mathbf{B}\mathbf{f}) \quad \text{or} \quad G_1(\mathbf{f}) = \sum_{i=1}^r g_i(\mathbf{B}_i\mathbf{f})$$

being  $r \in \mathbb{N}, r \geq 1$ ,  $\mathbf{B}$ ,  $\mathbf{B}_i$  linear operators, and  $g$ ,  $g_i$  proper, convex, lower semicontinuous functions with an easy-to-compute proximal operator, VMILA consists in a double loop method, where the inner loop is designed to inexactly solve the minimization problem associated to the definition of the proximal operator. The baseline idea is to write the minimization problem associated to the proximity operator in an equivalent dual form. The inner loop is devised with a suitable stopping criterion (Algorithm 1, Step 2), controlled by a parameter  $\eta \in (0, 1]$ . We refer to the original paper<sup>6</sup> for all the technical details. Here we only mention that in our case  $\mathbf{B} = (\Phi^T, \mathbb{1})^T$  being  $G_1(\mathbf{f}) = g_1(\mathbf{B}_1\mathbf{f}) + g_2(\mathbf{B}_2\mathbf{f})$  where  $g_1 = \|\cdot\|_1$ ,  $g_2 = \iota_{\mathbb{R}_+}$  and  $\mathbf{B}_1 = \Phi$ ,  $\mathbf{B}_2 = \mathbb{1}$ .

## 4. NUMERICAL EXPERIMENTS

In this Section, we present preliminary numerical results for the dynamic CT reconstruction problem. All the algorithms were implemented in Matlab 9.1 (R2016b), and performed on a MacBook Pro (Intel Core i7) at 3.3 GHz and CPU 16GB 2133 MHz LPDDR3 memory. To generate the shearlet operator  $\Phi$  and perform the tests, we used some functions of the ShearLab package,<sup>17,18</sup> available at [www.shearlab.org](http://www.shearlab.org).

### 4.1 Test problems

We test our approach on both simulated data (see Section 4.1.1) and real data (see Section 4.1.2). In both cases, we consider  $T = 33$  time instances and  $N_\theta = 30$  uniformly spaced angles over  $2\pi$  for each time frame.

#### 4.1.1 3D Shepp-Logan phantom

We use the 3D Shepp-Logan phantom sized  $N \times N \times T$ , with  $N = 128$ . After generating the  $128 \times 128 \times 128$  Shepp-Logan phantom with the routine `phantom3d.m` (available, for instance, on GitHub and implemented by Matthias C. Schabel), we downsample the third dimension choosing  $T = 33$  slices of interest, which will be understood as time frames. The corresponding projection data (*i.e.*, the sinogram) of the simulated phantom is corrupted by a white Gaussian process with zero mean and 0.5% variance.

#### 4.1.2 Emoji data set

We use tomographic X-ray real data measured with a custom-built CT device located in the radiation-shielded facilities of the Laboratory of Electronic Structure of the Department of Physics at the University of Helsinki (Finland). The target consist of 15 “emoji faces” (or *frames*) ideally evolving in time from an expressionless face (Figure 1(a)), namely, a face with narrow, closed eyes, and a straight mouth, to a smiling face with smiling

Table 1: Figures of merit for three frames of interest of the Shepp-Logan phantom (each slice is sized  $128 \times 128$ ) and for the 3D Shepp-Logan phantom (sized  $128 \times 128 \times 33$ ), obtained with  $\lambda = 10^{-2}$ . Corresponding reconstructed images are reported in Figure 2.

		PSNR	Rel. Err.	HaarPSI	SSIM	iter
frame 11	2D – static	24.66	0.12	0.61	0.88	43
	3D – dynamic	25.48	0.10	0.63	0.88	47
frame 17	2D – static	23.63	0.13	0.53	0.81	43
	3D – dynamic	24.57	0.09	0.57	0.83	47
frame 21	2D – static	24.09	0.13	0.57	0.84	43
	3D – dynamic	24.93	0.09	0.60	0.84	47
3D phantom	2D – static	25.90	0.002	0.67	0.90	43
	3D – dynamic	26.64	0.001	0.69	0.90	47

eyes (Figure 1(c)), that is, a face with wedge eyes and a thin convex mouth. In the intermediary frames, the mouth evolves from a straight line to a convex parabola, and the eyes (one by one) turn from a straight line to a wedge (Figure 1 (b)), while the outer circular boundary is fixed (*i.e.*, not in motion) for all the 15 “emoji faces”. The circular boundary, the mouth and the eyes of each frame are made of small squared ceramic stones (calcium) of about  $25 \text{ mm}^2$ . Each of the 15 frame consist of a 2D cross-section of the real 3D acquisition, each sized  $N \times N$ , with  $N = 64$ . To get the  $T = 33$  time instances, we doubled each of the 15 emoji face, and the last emoji face has been added 3 more times. Further details on the geometry of measurement setup can be found in the documentation of the acquisition of a stationary target in 2D, acquired with the same device.<sup>19</sup>

## 4.2 Parameters and stopping criterion

In all the experiments, the steplength  $\alpha_k$  is constrained in the interval  $[\alpha_{\min}, \alpha_{\max}]$  with  $\alpha_{\min} = 10^{-5}$  and  $\alpha_{\max} = 10^5$ , the initial threshold  $L$  for the scaling matrix  $D_k$  has been set equal to  $10^{10}$ . For the line search parameters we used  $\gamma = 0.4$ ,  $\beta = 10^{-4}$ . The parameter  $\eta \in (0, 1]$ , that controls the stopping criterion for the inner loop, is chosen equal to  $10^{-5}$ . As a suitable criterion for stopping the (outer) iterations of VMILA, we used the relative difference between two consecutive iterations. Namely, we stop the iterations when

$$\frac{\|\mathbf{f}^{(k)} - \mathbf{f}^{(k-1)}\|_2}{\|\mathbf{f}^{(k)}\|_2} \leq \tau_{\text{it}} \quad (7)$$

where we set  $\tau_{\text{it}} = 10^{-3}$ . Also, we set 200 as a safeguard maximum number of outer iterations, and 50 for the inner loop. Concerning the regularization parameter  $\lambda$ , we sampled the values  $10^\ell$ , with  $\ell = -4, -3, -2, -1, 0$ . The number of scales for both the 2D and 3D shearlet transform is set equal to 2, with 8 directions across each scale. As first estimate for  $\mathbf{f}^{(0)}$ , we choose to use the vector of all zeros, while for the initial guess of the inner loop of VMILA, at the first outer iterate it is chosen to be the vector of all zeros, and at all successive iterates a “warm start” is exploited, namely the inner solver is initialized with the dual solution computed at the previous iterate.

Lastly, for simulated data (see Section 4.1.1), all numerical tests were compared against different state-of-the-art figures of merit, namely the peak-signal-to-noise ratio (PSNR), the relative error (Rel. Err.), the structural similarity index (SSIM) and the Haar wavelet-based perceptual similarity index<sup>20</sup> (HaarPSI). Concerning real data (see Section 4.1.2), figures of merit estimates are not available, since a comparable groundtruth is not available. The goodness of the reconstruction is still retrievable on a qualitative basis, by looking at the picture taken at the original target (see Figure 1).

## 4.3 Numerical results

In Table 1 we present the results obtained with the simulated Shepp-Logan phantom for both scenarios considered in Section 2. In the columns, from left to right, we report the PSNR, the relative error, the HaarPSI, the SSIM, the number of performed iterations. All the results reported in Table 1 are obtained with  $\lambda = 10^{-2}$ , which gave



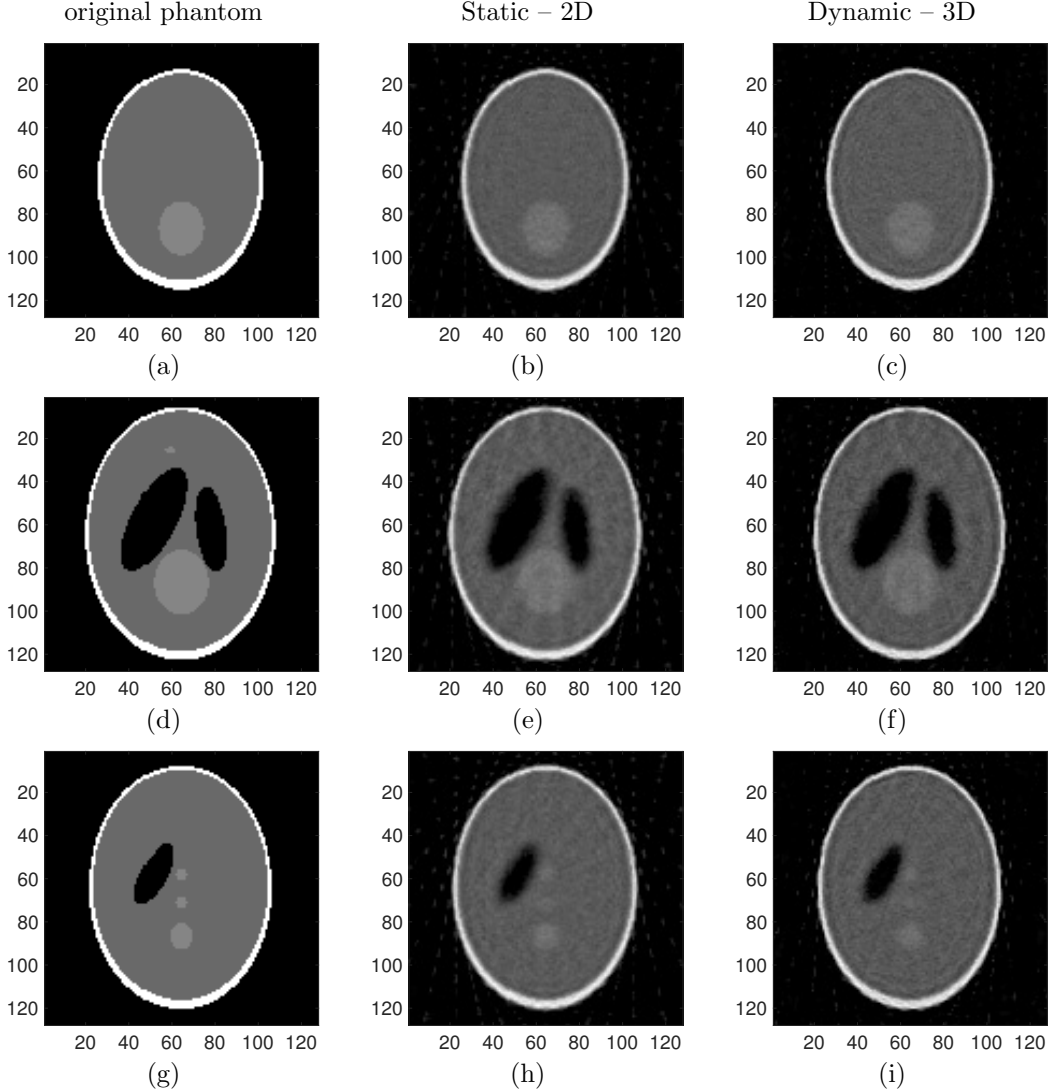


Figure 2: Shepp-Logan phantom (left) and corresponding reconstructions with static (middle) and dynamic (right) approaches ( $\lambda = 10^{-2}$ ). First row: frame 11. Second row: frame 17. Third row: frame 21.

the largest PSNR, HaarPSI and SSIM (and the smallest relative error) on the 3D phantom, for both approaches. The corresponding image reconstructions are collected in Figure 2. Here, in the first column we report the frame from the original phantom (panels (a), (d) and (g)), in the second the reconstructions obtained with the static (2D) setup (panels (b), (e) and (h)) and in the third the reconstructions obtained with the dynamic (3D) setup (panels (c), (f) and (i)). Each row in Figure 2 corresponds to a different frame in the original phantom.

The results in Table 1 show that the dynamic (3D) approach seems to outperform the static (2D) approach. Even if the difference is small, this is consistent for all three slices reported (and similar results are obtained on the other slices considered). This is confirmed by the corresponding images reported in Figure 1. At a first look, the two approaches seems to yield almost the same reconstruction, but with a closer look it can be noticed that the line artifacts (due mostly to the angular sub-sampling) are less evident in the dynamic (3D) reconstructions, and the edges are in general better recovered with the dynamic (3D) approach. In particular, by focusing on frame 21 (panels (h) and (i)) notice that the middle disk is almost lost with the static (2D) approach, while with the dynamic (3D) approach the dot has the right size with visible and well outlined edges.

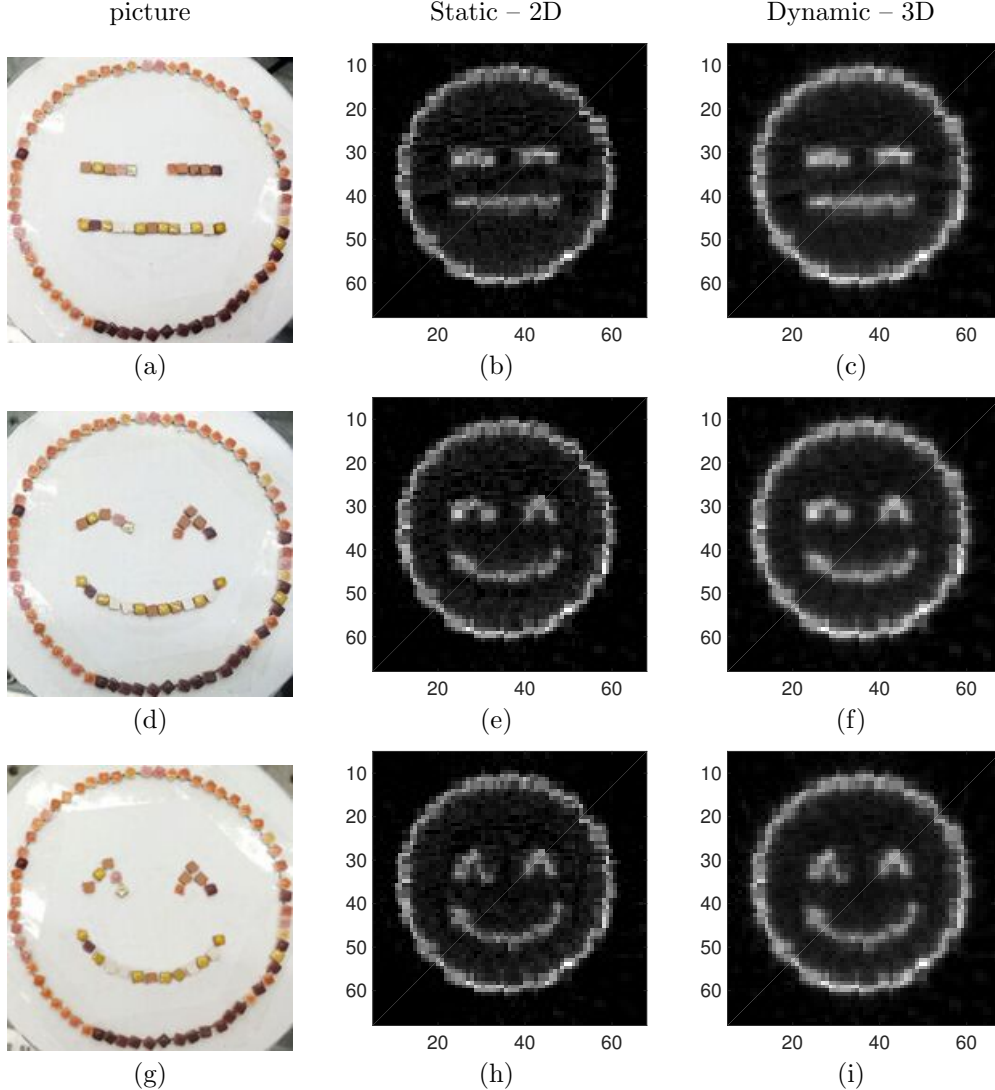


Figure 3: Emoji frames (left) and corresponding reconstructions with static (middle) and dynamic (right) approaches ( $\lambda = 10^{-2}$ ). First row: first frame. Second row: middle frame. Third row: last frame.

In Figure 3 we collect the reconstructions obtained with the Emoji data set. We recall that figures of merit estimates are not available, since a comparable groundtruth is not at disposal. Similarly to the simulated case, in the first column we report the original picture of the frame (panels (a), (d) and (g)), in the second the reconstructions obtained with the static (2D) setup (panels (b), (e) and (h)) and in the third the reconstructions obtained with the dynamic (3D) setup (panels (c), (f) and (i)). Also in this case, we found  $\lambda = 10^{-2}$  to yield the best reconstructions. Similarly to the Shepp-Logan phantom, both approaches seems to result in similar reconstructions. However, a closer look shows that the static (2D) reconstructions are more dominated by noise than the dynamic (3D) ones. This is evident by looking at the inner part of the phantom (inside the outer boundary of the emoji face). In both cases the boundary of each ceramic stone is not well detected, but this could be more an issue related to the small resolution ( $N = 64$ ) of each frame.

Overall, the reported results are preliminary but encouraging, indicating that a dynamic 3D shearlet approach could be beneficial. Indeed, by increasing the temporal resolutions (*i.e.*, considering more time frames) and by increasing the spatial resolution (*i.e.*, considering more detector cells) we could generate larger 3D shearlet system



with more scales, and this should result in even more accurate reconstructions, able to suppress the artifacts provoked by the motion.

## 5. CONCLUSIONS

In this paper we propose a 3D variational formulation based on shearlets for sparse dynamic tomography. The problem is formulated as a constrained minimum problem, where the functional to minimize is given by the sum of a data mismatch term with a nonsmooth sparsity-promoting regularizer. Our approach is tested by using both simulated and real data, and the numerical assessment is carried out via a recently proposed iterative proximal-gradient minimization method called VMILA. The results show that the dynamic (3D) approach outperforms the static (2D) one, on both simulated and real data. Even if the difference for figures of merit are limited, the results still indicate that the dynamic (3D) approach is able to suppress more noise and artifacts coming from the motion. Also, we expect to further improve these very preliminary results by considering more time frames and higher resolution for the targets.

## ACKNOWLEDGMENTS

This work has been supported by the Academy of Finland through the Finnish Centre of Excellence in Inverse Problems Research 2012–2017, decision number 250215. M. März acknowledges partial support by the DFG Priority Programme 1798 (CoSIP).

## REFERENCES

- [1] Ritman, E. L., “Cardiac computed tomography imaging: a history and some future possibilities,” *Cardiol. Clin.* **21**(4), 491–513 (2003).
- [2] Li, T., Schreibmann, E., Yang, Y., , and Xing, L., “Motion correction for improved target localization with on-board cone-beam computed tomography,” *Physics in medicine and biology* **51**, 253 (2005).
- [3] Hahn, B., “Reconstruction of dynamic objects with affine deformations in computerized tomography,” *Journal of Inverse and Ill-posed Problems* **22**, 323–339 (2014).
- [4] Hahn, B., “Null space and resolution in dynamic computerized tomography,” *Inverse Problems* **32**, 025006 (2016).
- [5] Kutyniok, G. and Labate, D., [*Shearlets. Multiscale Analysis for Multivariate Data*], Birkhäuser, Boston, MA (USA) (2012).
- [6] Bonettini, S., Loris, I., Porta, F., and Prato, M., “Variable metric inexact line-search based methods for nonsmooth optimization,” *SIAM J. Optim.* **26**, 891–921 (2016).
- [7] Meaney, A., *Design and Construction of an X-ray Computed Tomography Imaging System*, Master’s thesis, University of Helsinki, Department of Physics (2015).
- [8] Niemi, E., Lassas, M., Kallonen, A., Harhanen, L., Hämäläinen, K., and Siltanen, S., “Dynamic multi-source X-ray tomography using a spacetime level set method,” *Journal of Computational Physics* **291**, 218–237 (2015).
- [9] Burger, M., Dirks, H., Frerking, L., Hauptmann, A., Helin, T., and Siltanen, S., “A variational reconstruction method for undersampled dynamic X-ray tomography based on physical motion models,” *submitted*, arXiv:1705.06079 (2017).
- [10] Siltanen, S., Kolehmainen, V., Jarvenpaa, S., Kaipio, J., Koistinen, P., Lassas, M., Pirttila, J., and Somersalo, E., “Statistical inversion for X-ray tomography with few radiographs I: General theory,” *Physics in Medicine and Biology* **48**, 1437–1463 (2003).
- [11] Siltanen, S., Kolehmainen, V., Jarvenpaa, S., Kaipio, J., Koistinen, P., Lassas, M., Pirttila, J., and Somersalo, E., “Statistical inversion for X-ray tomography with few radiographs II: Applications to dental radiology,” *Physics in Medicine and Biology* **48**, 1465–1490 (2003).
- [12] Natterer, F. and Wübbeling, F., [*Mathematical methods in image reconstruction*], Siam: Society for Industrial and Applied Mathematics, Philadelphia, PA (USA) (2001).
- [13] Barzilai, J. and Borwein, J. M., “Two point step size gradient methods,” *IMA J. Numer. Anal.* **8**, 141–8 (1988).

- [14] Frassoldati, G., Zanghirati, G., and Zanni, L., “New adaptive stepsize selections in gradient methods,” *J. Industrial and Management Optim.* **4**(2), 299–312 (2008).
- [15] Lantéri, H., Roche, M., Cuevas, O., and Aime, C., “A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints,” *Signal Process.* **81**, 945–74 (2001).
- [16] Bonettini, S. and Prato, M., “New convergence results for the scaled gradient projection method,” *Inv. Prob.* **31**(9), 095008 (2015).
- [17] Kutyniok, G., Shahram, M., and Zhuang, X., “ShearLab: a rational design of a digital parabolic scaling algorithm,” *SIAM J. Imaging Sci.* **5**(4), 1291–1332 (2012).
- [18] Kutyniok, G., Lim, W.-Q., and Reisenhofer, R., “ShearLab 3D: Faithful digital shearlet transforms based on compactly supported shearlets,” *ACM Trans. Math. Software* **42**, 5 (2016).
- [19] Bubba, T. A., Hauptmann, A., Huotari, S., Rimpeläinen, J., and Siltanen, S., “Tomographic X-ray data of a lotus root filled with attenuating objects,” *arXiv:1609.07299* (2016).
- [20] Reisenhofer, R., Bosse, S., Kutyniok, G., and Wiegand, T., “A Haar wavelet-based perceptual similarity index,” *arXiv:1607.06140* (2017).