Warp-and-Project Tomography for Rapidly Deforming Objects

GUANGMING ZANG, RAMZI IDOUGHI, RAN TAO, GILLES LUBINEAU, PETER WONKA, and WOLFGANG HEIDRICH, King Abdullah University of Science And Technology, Saudi Arabia

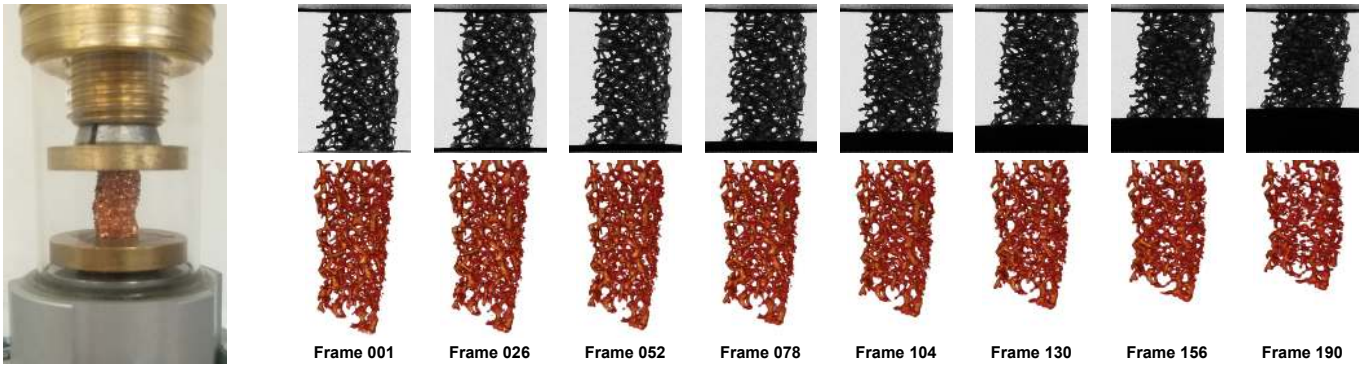

Fig. 1. We introduce a CT reconstruction method for objects that undergo rapid deformation during the scan. Shown here is a copper foam crumpling under a compressive force during the scan. The whole complex animation is reconstructed using only 192 projection images that all correspond to different deformation states of the foam.

Computed tomography has emerged as the method of choice for scanning complex shapes as well as interior structures of stationary objects. Recent progress has also allowed the use of CT for analyzing deforming objects and dynamic phenomena, although the deformations have been constrained to be either slow or periodic motions.

In this work we improve the tomographic reconstruction of time-varying geometries undergoing faster, non-periodic deformations. Our method uses a warp-and-project approach that allows us to introduce an essentially continuous time axis where consistency of the reconstructed shape with the projection images is enforced for the specific time and deformation state at which the image was captured. The method uses an efficient, time-adaptive solver that yields both the moving geometry as well as the deformation field.

We validate our method with extensive experiments using both synthetic and real data from a range of different application scenarios.

CCS Concepts: • **Computing methodologies** → *Computer graphics*; *3D imaging*; *Motion capture*.

Additional Key Words and Phrases: X-ray computed tomography, 4D reconstruction, deformation capture

ACM Reference Format:
Guangming Zang, Ramzi Idoughi, Ran Tao, Gilles Lubineau, Peter Wonka, and, Wolfgang Heidrich. 2019. Warp-and-Project Tomography for Rapidly Deforming Objects. *ACM Trans. Graph.* 38, 4, Article 86 (July 2019), 13 pages. https://doi.org/10.1145/3306346.3322965

Authors' address: Guangming Zang, guangming.zang@kaust.edu.sa; Ramzi Idoughi, ramzi.idoughi@kaust.edu.sa; Ran Tao, ran.tao@kaust.edu.sa; Gilles Lubineau, gilles.lubineau@kaust.edu.sa; Peter Wonka, peter.wonka@kaust.edu.sa;
Wolfgang Heidrich, wolfgang.heidrich@kaust.edu.sa, King Abdullah University of Science And Technology, Thuwal, 23955-6900, Saudi Arabia.

0730-0301/2019/7-ART86
https://doi.org/10.1145/3306346.3322965

1 INTRODUCTION

X-ray computed tomography (CT) is widely used in computer graphics and computer vision and even more frequently in medicine, biology and material science as a non-destructive imaging technique, able to reveal inner structures of the studied object. Until recently, X-ray CT was only used to scan static objects from different viewing angles. The need for a dynamic tomography reconstruction arises in applications where the scanned object undergoes deformation, or if the target of study is the motion itself. Even in the static case, tomography is often an ill-posed problem that requires hundreds of projections to reconstruct high-quality volumes. When the scanned object undergoes deformation, the number of projections for each deformation state is often insufficient for reconstructing each state with a traditional reconstruction methods. This makes the dynamic tomography reconstruction a highly challenging task.

Recently, Zang et al. [2018b] proposed a non-parametric Space-Time tomographic method (ST-tomography) to scan and analyze deforming objects and dynamic phenomena. While this method resulted in marked improvement of the state of the art, it does suffer from several shortcomings that we address in this work: First, ST-tomography was conceived for the case of relatively slow and smooth motion fields, where the deformation is negligible for short sequences of $\approx 10-60$ successive frames. Second, the method relies on an explicit tradeoff between spatial and temporal reconstruction quality. Finally, the temporal sampling is uniform, resulting in wasted computational effort for slow moving periods, as well as poor reconstruction quality for fast moving periods.

In this work, we propose a new warp-and-project approach for dynamic tomographic reconstruction. This new method, inspired by ST-tomography, relaxes the assumption of slow deformations in order to reconstruct objects with larger motion even between successive projections, where the ST-Tomography fails. Our proposed

method not only estimates the volume densities over time, but also the motion field. To this end, we move from a coarsely discretized time axis in ST-tomography to an essentially continuous time axis, where each projection image has its own time stamp, and warping is used to align the keyframes with the individual projections. We would like to highlight that the complexity of a deformation has no impact in our comparison to ST-tomography, only the speed of the motion matters. We also decouple the frame rate of the reconstructed volume sequence from the acquisition times for the captured projections. Finally, the temporal sampling is also adaptive, since additional volumes are reconstructed during periods of rapid motion. These improvements translate into significant improvements in the reconstruction results, as demonstrated by quantitative comparisons on simulated data, as well as qualitative comparisons on real data from a number of different application domains.

In summary, the main contributions of this work include:

- a new image formation model for dynamic tomography reconstruction, that takes into account the deformation occurring between successive captured projections.
- a temporal decoupling between the reconstructed key frames and the captured projections.
- a non-uniform temporal up-sampling, which will improve the quality of each reconstruction.
- a matrix-free solver for the proposed optimization algorithm.
- a strong evaluation of our approach both on simulated data, controlled experimental data (where a ground truth can be estimated) and several real data sets highlighting different application scenarios.

2 RELATED WORK

Dynamic 3D surface reconstruction. from color and depth sensors has been an active research topic in computer graphics. The state-of-the-art techniques allow real-time 4D reconstruction of non-rigidly deforming scenes using one [Innmann et al. 2016; Wang et al. 2018] or more [Dou et al. 2016] or RGB-D cameras. These approaches combine surface reconstruction over time with a parameterized nonrigid motion tracking in an optimization framework. Priors on motion are also integrated. In [Li et al. 2013; Zheng et al. 2017] a 3D scanner is used to capture 4D point cloud data, which is used respectively to reconstruct the plant growth and the flower blooming process. Furthermore, some approaches were developed for the capture of time-varying fluid surfaces, where specific priors to the fluid field are usually incorporated [Wang et al. 2009]. However, these surface-based methods cannot reconstruct internal structures of objects or occluded geometry.

Computed tomography. is a family of computational imaging techniques that reconstruct a density volume from a set of 2D images representing the projections of this volume along a set of directions. The classical reconstruction approaches are based on the use of the Radon transform and its inverse (e.g. the filtered back-projection algorithm [Feldkamp et al. 1984]). On the other hand, the Algebraic Reconstruction Technique (ART) [Gordon et al. 1970] and its numerous variants like the Simultaneous Algebraic Reconstruction Technique (SART) [Andersen and Kak 1984] use the Kaczmarz projection scheme, in order to iteratively update the reconstructed volume in a matrix-free fashion.

During the last two decades, CT found many applications in computer graphics and vision. Reche-Martinez et al. [2004] proposed a volumetric method to capture and render trees from multi-views photographs. Trifonov et al. [2006] developed a tomographic reconstruction method for transparent objects immersed inside a fluid having the same refractive index. Furthermore, visible-light tomography reconstruction has been widely used in fluid imaging, for several purposes like: estimating the 3D velocity field using the Tomo-PIV (Tomographic Particle Imaging Velocimetry) technique [Elsinga et al. 2006], reconstructing 3D flames [Hasinoff and Kutulakos 2007; Ihrke and Magnor 2004], imaging gas flows by reconstructing the refractive index variation of that gas due to the inner temperature changes [Atcheson et al. 2008], and capturing turbulent fluid mixtures [Gregson et al. 2014, 2012].

The main interest of using X-ray CT in computer graphics is to capture the internal structures of opaque objects [Anirudh et al. 2018; Zhao et al. 2011], as well as to retrieve complicated surfaces with occlusions like flowers [Ijiri et al. 2014; Stuppy et al. 2003; Zang et al. 2018a].

Dynamic tomographic reconstruction. is a challenging task, which becomes unavoidable in situations where the scan target deforms or degrades during the scanning process in a way that cannot be controlled or eliminated. Dynamic reconstruction also opens the door to the use of X-ray CT imaging for studies of dynamic phenomena, where the motion itself is the primary interest. Thus, several strategies have been proposed deal with dynamic objects.

In the medical field, the motion of the scanned organs is often periodic, e.g. for the heart or the lungs. Several methods have been based on this observation to provide a dynamic reconstruction of the heart or the lungs [Chen et al. 2012; Sonke et al. 2005]. Such methods reconstruct each phase of the motion cycle independently, by using only projections belonging to the same phase. This requires a large number of projections to cover all the phases of the cycle, which results in a higher radiation dose for the patient. In addition, the obtained reconstruction presents clinically prohibitive artifacts [Schmidt et al. 2014]. To overcome these issues, 4D iterative reconstructions have been proposed [Mory et al. 2014; Schmidt et al. 2014]. These methods employed either a spatial and temporal total variation regularization or an optical flow based registration between successive phases. However, the quality of the obtained reconstructions is still impacted by the low number of projections for each phase. Furthermore, a considerable effort has been made in this field to estimate the motion of the scanned object during the acquisition process [Li et al. 2007; Taubmann et al. 2015; Zeng et al. 2007]. Then the retrieved deformation field is employed to correct the reconstructed volumes. Often, this approach requires the knowledge of the initial state of the reconstructed volume, which will be used as a reference to estimate the motion.

On the hardware side, very fast CT scanning with hundreds of projections per second is enabled by bright X-ray sources such as synchrotrons [De Schryver et al. 2018]. This paves the way for

mechanical engineering and material science to obtain a better understanding of some dynamic processes. Indeed, CT and micro-CT devices are commonly used in these fields to study dynamic experiments like the compression of a composite materials [De Schryver et al. 2018; Weißenborn et al. 2016], fatigue cracks [Lachambre et al. 2015; Morgeneyer et al. 2013] or fluid flow in porous media [Shah et al. 2016, 2013; Shastry et al. 2018]. However, even with this fast hardware (which is far from commonly available), the acquisition time is still three orders of magnitude too slow for true video rate volume reconstruction using traditional algorithms, and so better space-time reconstruction algorithms are still highly desirable.

As a result, many studies use as stop-motion style acquisition of static states representing a supposedly continuous deformation [Hild et al. 2014; Lachambre et al. 2015; Morgeneyer et al. 2013]. The motion field can then be found by digital volume correlation (DVC, essentially 3D optical flow). Some variants of the DVC algorithm have been proposed to deal with continuous deformations, that occur during the scans. One improvement is made by projecting the volumes onto a set of basis functions (e.g. finite element basis) [Neggers et al. 2015]. Then the DVC is applied only on these basis functions, in order to speed up the estimation of the motion and reduce the degrees of freedom. Another variant is to use a projection-based DVC (P-DVC) to estimate the deformation of a known reference shape (template) from only few projections [Leclerc et al. 2015; Taillandier-Thomas et al. 2016]. Finally, Jailin et al. [2018] proposed a combination between a variant of the ART reconstruction and the basis functions P-DVC approach to reconstruct a deforming volume and retrieve its motion, in a multi-scale scheme. The main limitation of these approaches is their specificity to certain types of deforming objects, where it is easy to define the basis function and track the motion through them. For some phenomena such as rising dough, which undergoes significant topological change, these methods will fail to reconstruct the volumes.

Our work is most similar to the recent approach proposed by Zang et al. [2018b], which jointly reconstructs the 4D density volumes and the deformation fields between successive time frames. This method yields interesting results for many time-varying phenomena. However, as mentioned in the introduction, it still suffers from a number of shortcomings, most notably the inability to handle faster motions. Our work aims to address these shortcomings.

3 CONTINUOUS TOMOGRAPHY RECONSTRUCTION

3.1 Image Formation Model

In dynamic x-ray tomography the aim is to reconstruct a 4D volume $\mathbf{h}$ that represents the scanned deforming object, from N_p acquired projections. While previous work [Zang et al. 2018b] assumed that the object motion between a small number of successive frames is negligible, we target situations where the motion can be significant even between two images taken in immediate succession.

Assume that the CT scan consists of N_p projection images $\mathbf{p} = \{\mathbf{p}_1, \ldots, \mathbf{p}_j, \ldots, \mathbf{p}_{N_p}\}$ taken at times $\{t_1, \ldots, t_j, \ldots, t_{N_p}\}$. For the description in the following we loosely assume that the projection images are taken at regular intervals, i.e. $t_j = j \cdot \Delta t$, although the framework also works for irregular temporal sampling patterns. Due to the continuous motion of the scan target, the volume is

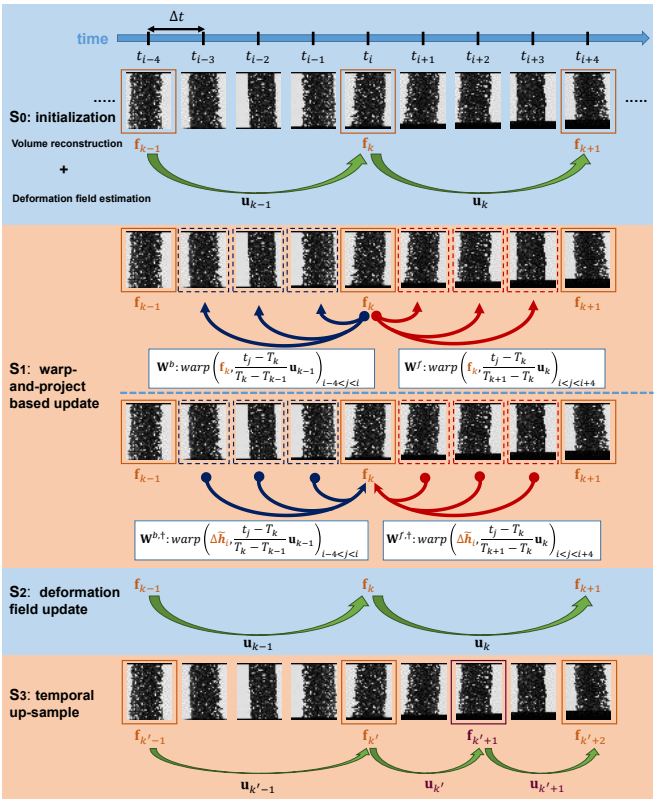

Fig. 2. Diagram of our method. S0: the key frames and the motion field between the key frames are initialized to 0. S1: volume reconstruction. The current volume estimate for a key frame is warped to the capture time of each X-ray image, and a residual image is computed by comparing the X-ray image with the projection of this warped volume. The residual is then back-projected into 3D, and warped back to the time of the key frame to update the volume estimate. S2: flow (velocity) reconstruction between subsequent key frames. S3: temporal up-sampling. New key frames are inserted where the motion is fast.

different at each capture time t_j, and is denoted as $\mathbf{h} = \{\mathbf{h}_j\}$. The relationship between the captured projections and the time varying volumes is then described as

$$\underbrace{\begin{pmatrix} \mathbf{A}_1 & & \\ & \ddots & \\ & \mathbf{A}_j & \\ & & \ddots \\ & & \mathbf{A}_{N_p} \end{pmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{pmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_j \\ \vdots \\ \mathbf{h}_{N_p} \end{pmatrix}}_{\mathbf{h}} = \underbrace{\begin{pmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_j \\ \vdots \\ \mathbf{p}_{N_p} \end{pmatrix}}_{\mathbf{p}}, \tag{1}$$

where $\mathbf{A}_j \in \mathbb{R}^{M \times N_v}$ is the projection matrix for *a single* projection, mapping the volume $\mathbf{h}_j$ with N_v voxels to a projection image $\mathbf{p}_j$ with M pixels. Also see Figure 2.

We note that Equation 1 is a heavily ill-posed problem that cannot be solved without additional priors such as motion models, since a single 2D projection image $\mathbf{p}_j$ is insufficient for reconstructing

a whole 3D volume $\mathbf{h}_j$. At the same time, it is often not necessary to reconstruct all N_p volumes to get a good representation of the shape and its deformation, since the changes can still be gradual and smooth. We therefore subsample the volumes $\mathbf{h}$ into a set of N_k key frames $\mathbf{f} = \{\mathbf{f}_1, \ldots, \mathbf{f}_k, \ldots, \mathbf{f}_{N_k}\}$ corresponding to times T_k. The output of our reconstruction method is this set of key frames, together with a set of motion fields $\mathbf{u} = \{\mathbf{u}_1, \ldots, \mathbf{u}_{N_k-1}\}$, where $\mathbf{u}_k$ describes the deformation between $\mathbf{f}_k$ and $\mathbf{f}_{k+1}$. We choose N_k adaptively, starting with a small number and increasing it until the deformation between key frames is small enough.

Given estimated key frames $\mathbf{f}_k$ and motion fields $\mathbf{u}_k$, we can approximate the volume at a time t_j with $T_k \leq t_j < T_{k+1}$ by warping (advecting) the key frame $\mathbf{f}_k$ *forward* in time:

$$\mathbf{h}_j \approx \tilde{\mathbf{h}}_j = \text{warp}(\mathbf{f}_k, \frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_k). \tag{2}$$

The warping operator $\text{warp}(\mathbf{f}_k, \frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_k)$ corresponds to the advection of the 3D volume $\mathbf{f}_k$ by the 4D deformation field $\frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_k$. In Equation 3 we illustrate how each voxel of the output field is computed. A 3$^{\text{rd}}$ order spatial interpolation is used in our implementation of the warping.

$$\tilde{\mathbf{h}}_j \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{f}_k \begin{pmatrix} x - \frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_{k,1}(x,y,z) \\ y - \frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_{k,2}(x,y,z) \\ z - \frac{t_j - T_k}{T_{k+1} - T_k}\mathbf{u}_{k,3}(x,y,z) \end{pmatrix} \tag{3}$$

Similarly, we can obtain another estimate by *backward* warping the next key frame:

$$\mathbf{h}_j \approx \tilde{\mathbf{h}}_j = \text{warp}(\mathbf{f}_{k+1}, \frac{t_j - T_{k+1}}{T_{k+1} - T_k}\mathbf{u}_k). \tag{4}$$

For ease of notation, we introduce warping operators $\mathbf{W}_j^f$ and $\mathbf{W}_j^b$ that respectively perform forward and backward warping to create two estimates of the intermediate frame $\tilde{\mathbf{h}}_j$ from the previous (resp. next) key frame, i.e. $\tilde{\mathbf{h}}_j = \mathbf{W}_j^f(\mathbf{f}_k)$ and $\tilde{\mathbf{h}}_j = \mathbf{W}_j^b(\mathbf{f}_{k+1})$. The image formation model from Equation 1 then corresponds to two separate data terms that can be utilized in an optimization-based reconstruction:

$$\mathbf{A} \cdot \mathbf{W}^f(\mathbf{f}) = \mathbf{p} \text{ and} \tag{5}$$

$$\mathbf{A} \cdot \mathbf{W}^b(\mathbf{f}) = \mathbf{p} \tag{6}$$

For implementing these data terms in an optimization approach, we also require the adjoint operators $\mathbf{W}^{f,\dagger}$ and $\mathbf{W}^{b,\dagger}$. However, these are easily implemented as the corresponding warps in the opposite direction, followed by an averaging of all warped volumes that contribute to a single keyframe.

Figure 2 shows a diagram of this warping-based interpolation of the intermediate volumes from the neighboring key frames, which is the key distinguishing characteristic of our method compared to the existing state of the art. Without this warping-based approach, all projections used in the reconstruction of a key frame are implicitly assumed to have been taken at the same time (i.e. representing the same shape). This is the approach taken by Zang et al. [Zang et al. 2018b], and it results in blurred reconstructions for faster motions.

3.2 Full Optimization Problem

Given the two data terms from above, we can now formulate an objective function for reconstructing the deforming geometry jointly with the motion field. Due to the ill-posed nature of this problem, this requires additional regularizers for both the key frames and the deformation field, which we adopt from the work of Zang et al. [2018b]:

$$\begin{aligned} \min_{\mathbf{f},\mathbf{u}} & \sum_{j=1}^{N_p} \left\| \mathbf{A}_j \mathbf{W}_j^f(\mathbf{f}_{j^-}) - \mathbf{p}_j \right\|_2^2 + \sum_{j=1}^{N_p} \left\| \mathbf{A}_j \mathbf{W}_j^b(\mathbf{f}_{j^+}) - \mathbf{p}_j \right\|_2^2 \\ & + \kappa_1 \sum_{k=1}^{N_k-1} \left\| \nabla_T \mathbf{f}_k + \nabla_S \mathbf{f}_k \cdot \mathbf{u}_k \right\|_1 \\ & + \sum_{k=1}^{N_k} \left[\kappa_2 \left\| \nabla_S \mathbf{f}_k \right\|_{\mathrm{H}_\epsilon} + \kappa_3 \left\| \nabla_T \mathbf{f}_k \right\|_2^2 \right] \\ & + \sum_{k=1}^{N_k-1} \sum_{i=x,y,z} \left[\kappa_4 \left\| \nabla_S \mathbf{u}_{k,i} \right\|_{\mathrm{H}_\tau} + \kappa_5 \left\| \nabla_T \mathbf{u}_{k,i} \right\|_2^2 \right], \end{aligned} \tag{7}$$

where $\mathbf{f}_{j^-}$ and $\mathbf{f}_{j^+}$ refer to the key frames immediately before and immediately after projection $\mathbf{p}_j$. Here, κ_1, κ_2, κ_3, κ_4 and κ_5 are weights of the different terms of the objective function. The operators ∇_T and ∇_S correspond to the discrete temporal and spatial gradients, implemented as one-sided divided differences. The first two terms correspond to the two warping-based data terms derived in Section 3.1. The second line in this objective function corresponds to a 3D version of the brightness constancy term in the Horn-Schunck optical flow [Horn and Schunck 1981]. In order to deal with large deformations we opted for a multi-scale implementation of the optical flow [Meinhardt-Llopis et al. 2013]. The next two terms in the third line correspond respectively to the spatial and temporal regularizations of the density volumes. A Huber penalty [Huber 2011] is used on the spatial gradient with a positive parameter ϵ, while we favor smooth behavior in the time domain with an L2-norm. Similar regularizations are also used for the deformation field in the two terms of the fourth line. τ is the positive parameter of the Huber penalty on the spatial gradient of the deformation field.

The framework that we propose in this paper is presented in Figure 2. A detailed description of this framework is also given as pseudo-code in Algorithm 1. First, an initialization of the density volumes and the deformation fields is performed (Step S0). Then, a Warp-and-Project update scheme (step S1) is applied to improve the quality of the reconstruction of the density volumes. Given the updated volumes, we re-estimate the deformation fields (Step S2). Finally, new intermediate key frames may be introduced (Step S3) in order to improve the temporal resolution and the spatial accuracy. The estimation of the deformation fields is done according to a multi-scale coarse-to-fine scheme [Meinhardt-Llopis et al. 2013]. The deformation fields are first estimated for the coarsest level ($s = N_{scales}$). Then, these fields are up-sampled scale-by-scale using the operator $\uparrow$ and re-estimated at each given scale. The output of

Algorithm 1 Warp-and-Project Tomography

procedure WP-TOMOGRAPHY($F_{\mathbf{f}}, G_{\mathbf{f}}, F_{\mathbf{u}}, G_{\mathbf{u}}, F_{\mathbf{u}}^{\mathbf{W}}, G_{\mathbf{u}}^{\mathbf{W}}, \rho$)
  // Step S0: initialize the key frames and deformation fields.
  $\mathbf{f} \leftarrow 0; \mathbf{u} \leftarrow 0;$
  while not converged **do**
    // Step S1: update volumes by comparing the projections
    // of warped volumes with captured projections
    $\mathbf{f} \leftarrow$ WARPANDPROJECT($F_{\mathbf{u}}^{\mathbf{W}}, G_{\mathbf{u}}^{\mathbf{W}}$)

    // Step S2: update deformation field between key frames
    // from the coarsest scale to the finest
    // generate multi-scale data
    $\mathbf{f}^1 \leftarrow \mathbf{f}, \mathbf{u}^1 \leftarrow \mathbf{u}$
    for s from 1 to $N_{scales} - 1$ **do**
      $\mathbf{f}^{s+1} \leftarrow \downarrow \mathbf{f}^s$
      $\mathbf{u}^{s+1} \leftarrow \rho \downarrow \mathbf{u}^s$
    end for
    for s from $N_{scales} - 1$ to 1 **do**
      $\mathbf{u}^s \leftarrow$ ESTIMATEDEFORMATIONS($F_{\mathbf{f}^s}(\mathbf{u}^s), G_{\mathbf{f}^s}(\mathbf{u}^s)$)
      $\mathbf{u}^{s-1} \leftarrow \frac{1}{\rho} \uparrow \mathbf{u}^s$
    end for
    $\mathbf{u} \leftarrow \mathbf{u}^1$
  end while

  // Step S3: temporal up-sample
  if not converged and motion too fast **then**
    $\mathbf{f}, \mathbf{u} \leftarrow$ TEMPORALUPSAMPLE($\mathbf{f}, \mathbf{u}$)
  end if

  return $\mathbf{f}$ and $\mathbf{u}$
end procedure

this step are the estimated deformation fields at the finest scale ($s = 1$). The up-sampling $\uparrow$and the down-sampling $\downarrow$are done by a factor of ρ, using a cubic interpolation. The steps (S1, S2 and S3) are repeated in a loop until convergence. For the last iteration, only the steps S1 and S2 are performed in order to get the best reconstruction at the last temporal sampling. The operators $F_{\mathbf{f}}, G_{\mathbf{f}}, F_{\mathbf{u}}, G_{\mathbf{u}}, F_{\mathbf{u}}^{\mathbf{W}}$ and $G_{\mathbf{u}}^{\mathbf{W}}$ will are explained below.

3.3 Solver

3.3.1 Step S0: Initialization. In the initialization of the algorithm, we set the deformation fields $\mathbf{u}_k = 0$. As a result, the warping operators simplify to identity operators, which means that in the first iteration the volume reconstruction (Step S1 below) is static volume reconstruction without the warp-and-project approach.

We also initialize the number of key frames N_k to be small, giving a coarse temporal resolution, and causing the initial volume estimates to be reconstructed from a large number of projections.

3.3.2 Step S1: Warp-and-Project based volume update. To solve the joint optimization problem in Equation 7, we split it into two sub-problems that we solve alternately. The sub-problem for reconstruction of the density volume is solved during this first step of our framework. It is described by the following minimization problem:

$$\begin{aligned}\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} \; & \sum_{j=1}^{N_p} \left\| \mathbf{A}_j \mathbf{W}_j^f(\mathbf{f}_{j^-}) - \mathbf{p}_j \right\|_2^2 + \sum_{j=1}^{N_p} \left\| \mathbf{A}_j \mathbf{W}_j^b(\mathbf{f}_{j^+}) - \mathbf{p}_j \right\|_2^2 \\ & + \kappa_1 \sum_{k=1}^{N_k - 1} \left\| \nabla_T \mathbf{f}_k + \nabla_S \mathbf{f}_k \cdot \mathbf{u}_k \right\|_1 \\ & + \sum_{k=1}^{N_k} \left[\kappa_2 \left\| \nabla_S \mathbf{f}_k \right\|_{\mathrm{H}_\epsilon} + \kappa_3 \left\| \nabla_T \mathbf{f}_k \right\|_2^2 \right]\end{aligned} \tag{8}$$

With the novel data terms used in this objective function, the comparison of the captured $\mathbf{p}_k$ and the simulated projection of the forward-warped key frame, $\mathbf{A}_k \mathbf{W}_k^f \mathbf{f}_k$ is performed between pairs of projections corresponding to the same time t_i, and hence represent the same geometric configuration of the scan target. The same holds for the backward-warped key frame in the second data term.

In the following, we will drop the f and b superscript, and simply refer to the warping operator as $\mathbf{W}$. To simplify the notation, we will further use the non-linear $\mathbf{W}$ like a matrix, and refer to its adjoint operator as $\mathbf{W}^T$.

To solve the optimization problem in Equation 8, we use a first-order primal-dual framework [Chambolle and Pock 2011] as shown in Algorithm 2. This optimization problem is then split into two sub-problems, that we solve alternately using proximal operators. The first sub-problem contains only the least squares data fitting terms of this equation, that we denote as $F_{\mathbf{u}}^{\mathbf{W}}$. The remaining terms, denoted as $G_{\mathbf{u}}^{\mathbf{W}}$, are grouped together into the second sub-problem.

Algorithm 2 CP-based method for tomographic reconstruction

procedure WARPANDPROJECT($F_{\mathbf{u}}^{\mathbf{W}}, G_{\mathbf{u}}^{\mathbf{W}}$)
  while not converged **do**
    // update slack variable
    $\mathbf{g}^{j+1} \leftarrow \mathbf{prox}_{\lambda_1 G_{\mathbf{u}}^{\mathbf{W}^*}}(\mathbf{g}^j + \lambda_1 \mathbf{K}_{\mathbf{u}} \bar{\mathbf{f}}^j)$
    // update volume variable
    $\mathbf{f}^{j+1} \leftarrow \mathbf{prox}_{\mu_1 F_{\mathbf{u}}^{\mathbf{W}}}(\mathbf{f}^j - \mu_1 \mathbf{K}_{\mathbf{u}}^{\mathrm{T}} \mathbf{g}^{j+1})$
    // update dual variable
    $\bar{\mathbf{f}}^{j+1} \leftarrow 2 \cdot \mathbf{f}^{j+1} - \mathbf{f}^j$
  end while
  return $\mathbf{f}$
end procedure

The Algorithm 2 shows the pseudo-code used to solve the Warp-and-Project strategy, which is based on the CP-algorithm. The notations $\mathbf{g}$, $\bar{\mathbf{f}}$, $\mathbf{prox}_{\lambda_1 G_{\mathbf{u}}^{\mathbf{W}^*}}$ and $\mathbf{prox}_{\mu_1 F_{\mathbf{u}}^{\mathbf{W}}}$ are used respectively for the slack and dual variables, and the proximal operators of the functions $G_{\mathbf{u}}^{\mathbf{W}^*}$ and $F_{\mathbf{u}}^{\mathbf{W}}$. The operator $\mathbf{K}_{\mathbf{u}}$ is defined by:

$$\mathbf{K}_{\mathbf{u}} = \left(\nabla_S, \; \nabla_T, \; \mathbf{W} \right)^T \tag{9}$$

Since we are using the same priors for the density volumes as Zang et al. [2018b], the proximal operator $\mathbf{prox}_{\lambda_1 G_{\mathbf{u}}^{\mathbf{W}^*}}$ is also the same, and we refer to their paper for the details.

On the other hand, we proposed a new derivation to solve the proximal operator $\mathbf{prox}_{\mu_1 F_{\mathbf{u}}^{\mathbf{W}}}$, given by:

$$\mathbf{prox}_{\mu_1 F_{\mathbf{u}}^{\mathbf{W}}}(\mathbf{f}_v) = \underset{\mathbf{f}}{\operatorname{argmin}} \ \|\mathbf{AWf} - \mathbf{p}\|_2^2 + \frac{1}{2\mu_1}\|\mathbf{f} - \mathbf{f}_v\|_2^2. \quad (10)$$

This proximal operator problem is equivalent to solving the following minimization problem:

$$\begin{aligned} &\underset{\widehat{\mathbf{X}}}{\operatorname{argmin}} \ \left\|\widehat{\mathbf{X}}\right\|_2^2 \\ &\text{subject to:} \quad \widehat{\mathbf{A}} \cdot \widehat{\mathbf{X}} = \widehat{\mathbf{p}}. \end{aligned} \quad (11)$$

where:

$$\begin{aligned} \widehat{\mathbf{p}} &= \sqrt{2\mu_1}\,(\mathbf{p} - \mathbf{AW} \cdot \mathbf{f}_v) \\ \widehat{\mathbf{A}} &= \begin{bmatrix} \mathbb{I} & \sqrt{2\mu_1}\mathbf{AW} \end{bmatrix} \\ \widehat{\mathbf{X}} &= \begin{bmatrix} \sqrt{2\mu_1}\,(\mathbf{p} - \mathbf{AWf}) & \mathbf{f} - \mathbf{f}_v \end{bmatrix}^T \end{aligned} \quad (12)$$

In the pseudo-code detailed in the Algorithm 3, we present the solver that we used for the proximal operator in Equation 10. For each key frame k, we warp the density volume $\mathbf{f}_k$ to the times t_j of its neighboring projections. Then, we project the obtained volume $\tilde{\mathbf{h}}_j$ using the corresponding viewing angle θ_j, and compute the residual image with respect to the captured projection $\mathbf{p}_j$. A correction volume $\Delta\tilde{\mathbf{h}}_j$ is then computed by back-projecting the residual image Δp. Finally, this correction volume is warped from the time t_j of the j^{th} projection to the time T_k, and is then used to update the key frame $\mathbf{f}_k$.

Algorithm 3 Solver for the proximal operator in Equation 10

procedure WARPING-CORRECTION($\mathbf{u}, \mathbf{f}_v$)
  initialize: $\alpha \in \mathbb{R}$, $\mathbf{f} = \mathbf{f}_v$ and $\mathbf{q} = \mathbf{f} - \mathbf{f}_v = 0$
  for k from 1 to N_k **do**
    while not converged **do**
      for all j with $T_{k-1} \le t_j \le T_{k+1}$ **do**
        // warp the key frame to a given time (Eqs. 2, 4)
        $\tilde{\mathbf{h}}_j = \mathbf{W}\mathbf{f}_k$
        // residual in image space
        $\widehat{\mathbf{p}}_j = \sqrt{2\mu_1}\left(\mathbf{p}_j - \mathbf{A}\tilde{\mathbf{h}}_j\right)$
        $\widehat{\mathbf{X}}_j = \left[\sqrt{2\mu_1}\left(\mathbf{p}_j - \mathbf{A}\tilde{\mathbf{h}}_j\right) \quad \mathbf{q}_k\right]^T$
        $\Delta p_j = \widehat{\mathbf{p}}_j - \widehat{\mathbf{A}} \cdot \widehat{\mathbf{X}}_j$
        // compute the correction volume
        $\Delta\tilde{\mathbf{h}}_j = \widehat{\mathbf{A}}^T \cdot \Delta p_j$
        // warp the residual to key frame and update
        $\mathbf{f}_k = \mathbf{f}_k + \alpha\mathbf{W}^T\Delta\tilde{\mathbf{h}}_j$
        $\mathbf{q}_k = \mathbf{q}_k + \alpha\Delta p_j$
      end for
    end while
  end for
  return f
end procedure

One can notice that the proximal operator in Equation 10 is very similar to the case of a linear least squares problem [Parikh et al. 2014] where the solution is given by:

$$\mathbf{prox}_{\mu_1 F_{\mathbf{u}}^{\mathbf{W}}}(\mathbf{f}_v) = \left(2\mu_1\,(\mathbf{AW})^T\,\mathbf{AW} + \mathbb{I}\right)^{-1}\left(2\mu_1\,(\mathbf{AW})^T\,\mathbf{p} + \mathbf{f}_v\right) \quad (13)$$

Actually, our proposed solver is also quite similar to this solution, except that the warping operator $\mathbf{W}$ and its adjoint are non-linear operators that are implemented procedurally.

3.3.3 Deformation field estimation (S2). After reconstructing volume estimates with the warp-and-project strategy, the deformation fields between successive key frames have to be updated. For this purpose, we solve the deformation field estimation sub-problem derived from the Equation 7. This optimization sub-problem is described as follows:

$$\begin{aligned} \mathbf{u}^* \quad = \quad & \underset{\mathbf{u}}{\operatorname{argmin}} \ \kappa_1 \sum_{k=1}^{N_k-1} \|\nabla_T \mathbf{f}_k + \nabla_S \mathbf{f}_k \cdot \mathbf{u}_k\|_1 \\ & + \sum_{k=1}^{N_k-1} \sum_{i=x,y,z} \left[\kappa_4 \left\|\nabla_S \mathbf{u}_{k,i}\right\|_{\mathrm{H}_\tau} + \kappa_5 \left\|\nabla_T \mathbf{u}_{k,i}\right\|_2^2\right] \end{aligned} \quad (14)$$

The adopted approach to solve this optimization is the same as in the ST-tomography framework [Zang et al. 2018b].

3.3.4 Temporal up-sampling (S3). In the state-of-the-art 4D CT reconstruction methods, the temporal sampling is dependent on the number of captured projections N_θ used to reconstruct the density volume for one key frame. However, one of the advantages of the proposed warp-and-project strategy is to decouple the key frames from their association to these sets of N_θ captured projections. This allows us to adjust the temporal sampling by choosing the number of key frames, without taking into account the number of captured projections N_θ needed to initialize the reconstruction of each key frame. Then, we can on the one hand increase the temporal resolution of our reconstruction. On the other hand, we also increase the accuracy of our reconstruction as we will show in Section 4. In addition, the combination between the warp-and-project strategy and the temporal up-sampling opens the way to the reconstruction of fast phenomena with a good accuracy.

The temporal up-sampling is performed only if a deformation field $\mathbf{u}_k$ exceeds a given threshold velocity. In this case, we introduce an intermediate key frame between the key frames $\mathbf{f}_k$ and $\mathbf{f}_{k+1}$. The density volume for this new key frame $\mathbf{f}_{k'+1}$ is computed by warping $\mathbf{f}_k$ to the time $T_{k'+1}$ using the deformation field $\mathbf{u}_k$. Moreover, the deformation field $\mathbf{u}_k$, is split linearly into two parts, each covering half the motion. For the last outer iteration of our framework (see Algorithm 1), the temporal up-sampling is skipped. This is to ensure the final density volumes reconstruction result has been updated by the Warp-and-Project strategy. The obtained temporal sampling for the key frames is mostly non-uniform. With this sampling choice, we reduce the memory cost for saving the density volumes as well as the deformation fields.

4 RESULTS AND DISCUSSION

In the following, we first validate our proposed approach by quantitative comparisons on both simulated data (fluid flow) as well

as real data where ground truth is available from high resolution scanning using a stop-motion approach (copper foam). Then we show the results of our reconstruction on six different data sets, where the deformation is relatively fast. All experiments use the low-discrepancy view sequence proposed by Zang et al [2018b].

In the experiments shown in this section, many parameters are common. For the multi-scale optical flow calculation, we use $N_{scales} = 3$, $\sigma = 0.65$, and $\rho = 0.5$. The weights for the Huber penalty priors are set to $\kappa_2 = 0.08$ and $\kappa_4 = 1.2$. We tested a range of parameters and observed that the ranges (0.01, 0.1) for κ_2, (0.8, 1.5) for κ_4, and (0.01, 0.05) for κ_5 provide good results. For volume density reconstruction, $\kappa_1 = 0.25$, the relaxation parameter α for SART is set to 0.3, and two inner SART iterations are applied for all methods. For the proximal framework, $\lambda_1 = \mu_1 = 0.3$. Our algorithm was implemented in C++. The run times for the density volume reconstruction and the flow field estimation, the dynamic range of density levels, and other specific acquisition parameters are given in Table 2. Comparisons are made with two reference methods. The first is **SART-ROF**, a state-of-the-art static tomographic reconstruction method that incorporates the Rudin-Osher-Fatemi Total Variation prior [Getreuer 2012]. The resulting optimization problem per frame is solved using a primal-dual scheme, and the SART algorithm is used as the solver for the data term. The second reference method is the Space-Time Tomography (**ST-Tomography**) proposed by Zang et al. [2018b].

4.1 Quantitative validation

4.1.1 Synthetic plume data. Our first validation experiment is done using synthetic fluid flow data, generated by mantaflow [Thuerey and Pfaff 2018]. The purpose of this data set is to quantify the quality of the estimated deformation field, since simulation is the only way to generate dense ground truth motion. The quality of the volume reconstructions is assessed with real scans later in this section.

The simulation domain (resolution $100 \times 150 \times 100$) contains a cylindrical source emitting a (spatially and temporally) non-uniform density. The aim of this non-uniformity is to introduce some texture inside the plume for tracking. The emitted density is transported through the domain by a 3D incompressible flow over 300 time steps. For this experiment, the velocity of the density transport is controlled by the buoyancy parameter in the domain. We used 5 different values for this parameter ($v = 0.1, 0.41, 0.73, 1.28$ and 3.0 voxels/Δt). Frames from the fastest animation are shown in Figure 3. The leftmost image shows the starting configuration, which is the same for all simulations.

Only the volumes having an odd index have been employed for the reconstruction of the plumes, but all 300 simulated volumes are used later for the comparison between the different obtained reconstructions and the ground truth. From each volume with an odd index a projection is computed with view points according to the low discrepancy view sampling strategy [Zang et al. 2018b] The reconstruction methods (ST-tomography and our method) are then performed using these 150 projections with the following strategies to select the number of key frames. For ST-tomography, we combine 10 projections together to reconstruct each key frame. For our method, 5 levels of uniform temporal up-sampling are performed to obtain 15, 30, 60, 120 or 150 key frames. Here, the level 1 is the level with the smallest number of reconstructed key frames.

Table 1. Numerical comparisons with ground truth data for different algorithms: average end-point error (EE) in voxels and average angular error (AE) in degrees.

	Speed (voxels/Δt)	ST-tomo.	Ours L1	Ours L2	Ours L3	Ours L4	Ours L5
EE	3.0	0.58	0.49	0.42	0.36	0.33	0.32
	1.28	0.44	0.37	0.32	0.27	0.25	0.21
	0.73	0.37	0.32	0.28	0.23	0.21	0.19
	0.41	0.31	0.27	0.24	0.20	0.18	0.17
	0.1	0.27	0.24	0.21	0.18	0.16	0.15
AE	3.0	17.5	17.1	16.2	14.6	13.9	13.8
	1.28	15.8	15.4	14.5	13.3	11.9	11.5
	0.73	14.3	14.1	13.3	12.2	11.3	10.9
	0.41	13.3	12.9	12.0	11.2	10.7	10.5
	0.1	12.7	12.0	11.2	10.7	10.3	10.2

Table 1 shows comparisons of the Average Endpoint Error (EE) and the Mean Angular Error (AE) for the competing methods. EE and AE are standard error metrics for assessing optical flow methods. Shown are ST-tomography and several levels of our method. We can see that our method dominates ST-tomography even after just one iteration, and then continues to improve over the next few iterations as the adaptive temporal sampling and the warp-and-project method improve the motion estimates. Furthermore, the method degrades gracefully with increased speed.

4.1.2 In-situ transformations of a copper foam. To perform quantitative evaluation of the volume reconstructions, we turn to real X-ray images of copper foam. This dataset is inspired by composite material analysis in mechanical engineering, where the deformation characteristics of such metal foams and similar kinds of materials are being investigated. Two controlled transformations (translation and compression) were conducted using a CT5000 5kN compression load stage (Deben UK Ltd., Suffolk, UK). This stage contains two parallel surfaces: a fixed one (top) and a moving one (bottom). The latter is controlled with high precision using a software interface. This setup allows us to capture high quality scans of many poses of the deformation using a stop-motion approach.

Vertical translation of the copper foam. For this experiment, a thin slice of copper foam (dimensions: $8.69 \times 8.42 \times 1.73$ mm^3) was set on the moving part of the compression stage. The fixed part of the stage has been removed for this experiment. The sample was scanned for 42 different positions, using 90 viewing angles each. Between successive times, the sample was translated by 0.1 *mm*.

The Figure 4 illustrates 4 positions of the sample during the translation. While the first row shows projections of the sample at the same viewing angle, the second row shows the obtained reconstruction of the copper foam using the SART-ROF algorithm. This reconstruction is very accurate, since the object is static for each position, and we use a large number of projections. For the following comparison of the different reconstruction methods these

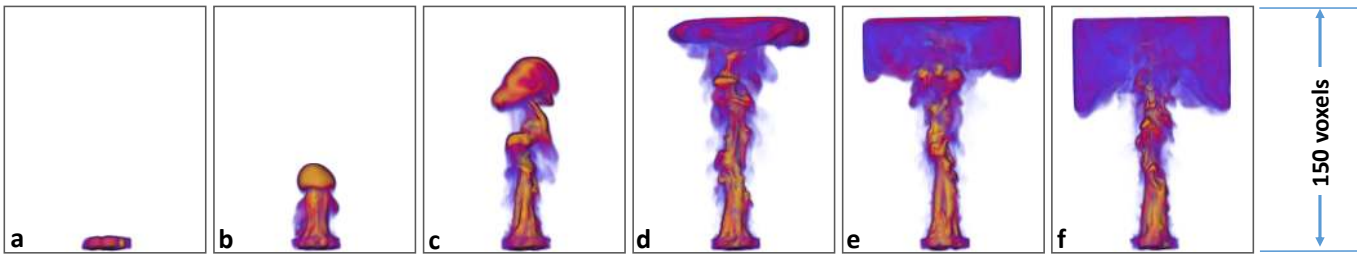

Fig. 3. Several frames from the fastest fluid animation.

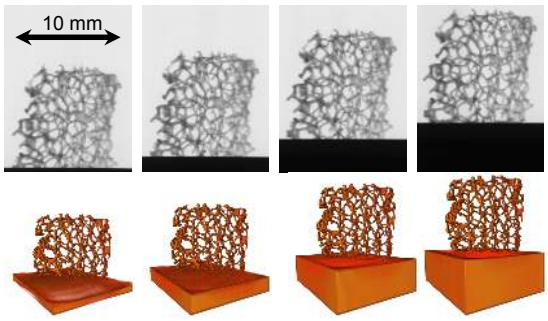

Fig. 4. The translation of the copper foam. First row: captured projections, second row: reconstructed volumes at time frame 2, 12, 30, 40.

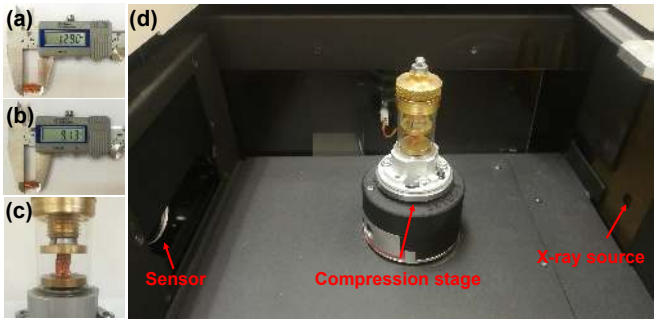

Fig. 6. (a-b): The height (in mm) of the copper foam before and after the compression process, respectively. The total displacement during the compression is 3.77 mm, uniformly distributed over 192 scans. 60 projections are obtained for each scan. (c): the states of the foam after 192 scans. (d) The Micro-CT setup for the in-situ compressions of the foam.

reconstructed volumes are considered as the ground truth volumes for the copper foam at each position.

For the dynamic reconstructions, from each position only 6 projections are used. Moreover, the projections from each 5 successive positions are combined in one time frame for the SART-ROF and the ST-tomography as well as for the initialization of our method. After convergence, the SART-ROF and the ST-tomography approaches reconstruct only 8 volumes, while with our method a volume is reconstructed for each position of the copper foam.

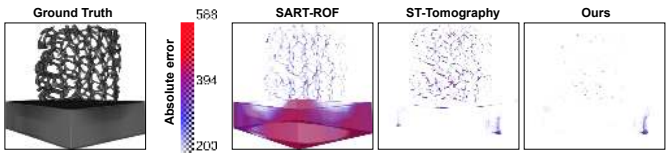

Fig. 5. Absolute error for SART-ROF, ST-Tomography and ours for time frame 25. The dynamic range of density values is (0, 1800).

For one intermediate volume the absolute error is given in Figure 5 between the ground truth and the different reconstruction methods. This figure shows the improved accuracy of our reconstruction compared to the state-of-the-art methods.

Compression of the copper foam. In this experiment, we scanned the foam crumpling under a compressive force using the setup shown in Figure 6. This is a real-world scenario that is of interest in mechanical engineering applications. In order to obtain ground-truth data, we again employed a stop-motion scanning strategy, this time acquiring 192 individual scans with 60 projections each. To test our method against the competing approaches, we then selected 1 projection for each scan according to the low-discrepancy sequence. Results for a single frame are shown in Figure 7. The top row shows the reconstructions, while the bottom row shows the error of the different methods. We can see that our new method significantly reduces the reconstruction error compared to both SART-ROF and ST-tomography. Please also refer to the video.

To quantify this effect further, we show numerical results in Figure 8 and Table 3. Figure 8 plots the PSNR of each key frame for SART-ROF, ST-tomography, and several iterations of our method. As expected, the stationary reconstruction by SART-ROF has the worst performance, with ST-tomography improving the result by a 2−dB on average. Our warp-and-project method is slightly superior to ST-tomography even in the first iteration, and then continues to improve by 2 – 3dB over the next 4 iterations. This is because the adaptive insertion of key frames allows for more accurate estimations of the motion field, which in term allow for better reconstructions of the volume densities with the warp-and-project approach. Table 3 shows numerical results of the same experiment, aggregated over all frames, but with separate statistics for the top, middle, and bottom parts of the volume. These exhibit different motion speeds since the foam is only compressed from the bottom, while the top edge is stationary. To test multiple virtual scan speeds, we also performed temporally sub-sampled reconstructions in which only every other or every third time step is used (i.e. reconstructions from 96 and 64

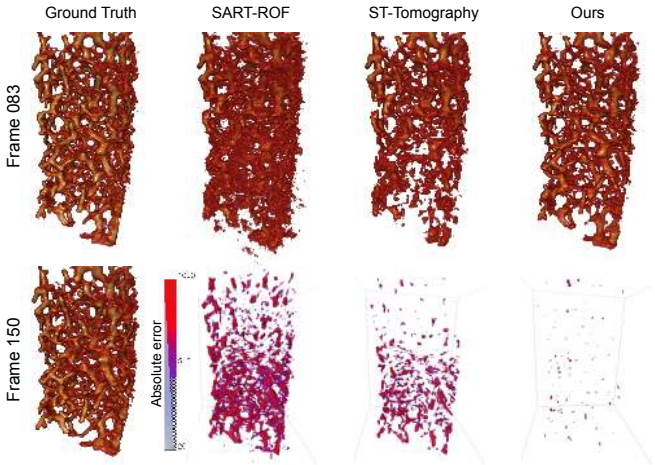

Fig. 7. Algorithm comparison for the compression of copper foam. First row: results from different reconstruction methods compared to the ground truth; Second row: the absolute error for time frame 150. The dynamic range of density values is (0, 4600).

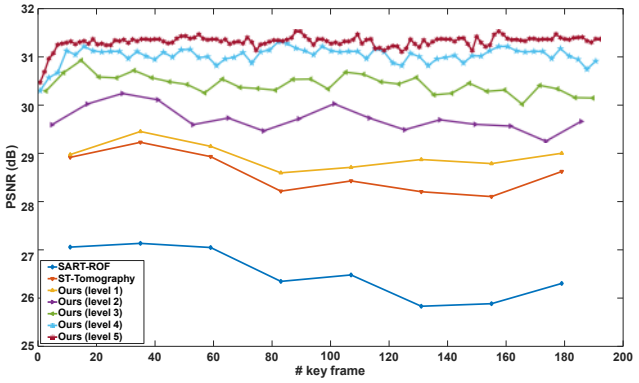

Fig. 8. PSNR values of the volume reconstructions for each key frame. We compare SART-ROF, ST-tomography, and several iterations of our method. Please see text for details.

projections, respectively). From the results presented in Table 3, we notice that with the third experiment (Sampling 3), we have a PSNR lower than 20 for the bottom region of the foam corresponding to the fastest motions. We can consider this experiement as the failure case of our method. In this experiment, a 30% compression of the foam is performed approximately over 64 "scans", which corresponds to a 0.47% compression per scan. We illustrate in Figure 9 the degradation of the reconstruction quality when the number of the used projections in the reconstruction is reduced.

4.2 Application to additional real scans

We also provide qualitative results for a number of fast-moving data sets. The algorithm parameters for the individual dataset are shown in Table 2 along with the reconstruction times. In addition

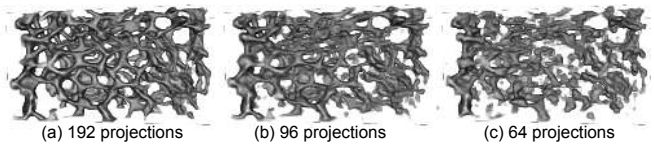

Fig. 9. The degradation of the reconstruction quality with reduced number of used projections, for the bottom part (i.e. fastest region) of the compressed copper foam. (a-c): 192, 96 and 64 projections are respectively involved for the reconstruction.

to the visual results presented here, we also encourage the reader to consult the video for the full time-varying reconstruction results.

Rising dough. The first data set is a scan of fast rising dough, shown in Figures 13, 11, and 10. As the dough rises due to the yeast secreting CO_2 gas, we see bubbles forming at both microscopic and mesoscopic scales. The microscopic bubbles manifest themselves as a change in the density of the dough, while the meso-scale bubbles result in drastic topology changes of the dough that are resolved well with our method, but are blurred out in the comparison methods (see slices in Figure 11). A quantitative analysis is also conducted by comparing the PSNR and SSIM between the acquired projection image from real scan, and the projected images from the reconstructed volumes using different methods, as shown in Table 4. Note that the projection used for this comparison is not involved in the different reconstructions.

Fig. 10. Rising dough reconstructed by our method. First row: direct volume rendering. Second row: bubble surfaces.

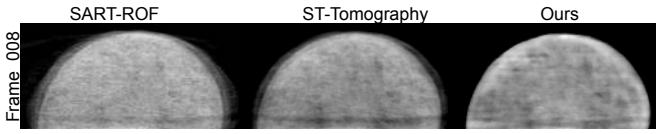

Fig. 11. Comparison between SART-ROF, ST-tomography and ours for a slice visualization of the dough dataset.

Capillarity effect in porous rock. The next experiment is a standard test deployed in geology, and specifically oil and gas exploration [Shah et al. 2013]. The goal of such studies is to measure the porosity or the permeability of rock samples in order to quantify

Table 2. Parameters used in the acquisition and for the tomographic reconstruction.

Dataset	# proj.	Cap. time [h:mm]	N_θ (ST-tomo.)	Keyframes (Ours)	Volume size	S1 step [h:mm]	S2 step [h:mm]	# Outer iterations	κ_3	κ_5	Dynamic range
rock	800	1:22	20	80	400×500×400	4:42	6:30	5	0.3	0.05	(0.0, 86.3)
fungus	600	0:38	30	128	250×300×250	2:20	4:36	4	0.2	0.03	(0.0, 19.5)
hydro-gel	640	0:43	16	70	423×320×423	3:15	5:20	5	0.3	0.05	(0.0, 24.8)
liquids	600	0:46	30	120	180×200×180	3:44	5:50	4	0.1	0.03	(0.0, 110.2)
pills	30	0:02	3	30	400×200×400	1:20	2:25	5	0.1	0.03	(0.0, 38.5)

Table 3. Calculated PSNR [dB], and SSIM for different reconstruction methods: SART-ROF, ST-Tomography, and our method.

Region	Top (slowest)		Center (medium)		Bottom (fastest)	
Sampling 1: 192 projections						
Metric	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
SART-ROF	29.14	0.68	26.27	0.56	24.11	0.28
ST-Tomo.	32.11	0.72	27.85	0.64	24.68	0.41
Ours (level 1)	32.28	0.72	28.48	0.65	24.90	0.46
Ours (level 2)	32.72	0.75	29.10	0.67	26.25	0.53
Ours (level 3)	33.32	0.77	29.93	0.68	28.08	0.56
Ours (level 4)	33.74	0.78	30.57	0.70	29.18	0.61
Ours (level 5)	33.78	0.78	30.87	0.72	29.35	0.62
Sampling 2: 96 projections						
SART-ROF	28.55	0.65	23.61	0.24	21.31	0.19
ST-Tomo.	31.41	0.69	24.76	0.47	23.14	0.36
Ours (level 1)	31.51	0.69	25.14	0.50	23.39	0.37
Ours (level 2)	31.86	0.71	26.29	0.56	23.95	0.39
Ours (level 3)	32.43	0.72	27.84	0.60	24.45	0.43
Ours (level 4)	32.59	0.72	28.88	0.63	24.72	0.44
Ours (level 5)	32.88	0.73	29.08	0.64	24.87	0.46
Sampling 3: 64 projections						
SART-ROF	26.83	0.58	21.48	0.19	16.85	0.13
ST-Tomo.	28.73	0.63	22.92	0.33	17.92	0.19
Ours (level 1)	28.79	0.63	22.92	0.34	17.96	0.19
Ours (level 2)	28.94	0.64	23.28	0.37	18.25	0.20
Ours (level 3)	29.11	0.65	23.68	0.40	18.72	0.23
Ours (level 4)	29.28	0.67	23.98	0.42	18.94	0.24
Ours (level 5)	29.35	0.67	24.17	0.44	19.08	0.24

Table 4. Calculated PSNR [dB] and SSIM for projection-based comparison over the dough data set. We compare here the acquired projection at time frame 8 with the projections obtained from the different reconstruction methods: SART-ROF, ST-Tomography, and our proposed method.

Metric	SART-ROF	ST-Tomography	Ours
PSNR	27.61	29.89	30.25
SSIM	0.82	0.86	0.90

the ability of the rock to store oil. Here, a cylindrical sample of the rock is placed in a dish with the bottom dipped into a liquid contrast agent. Due to capillary action, the liquid is absorbed into the sample over the duration of the scan (one hour and 22 minutes). A 3D visualization of this experiment showing the absorbed liquid can be found in Figure 12.

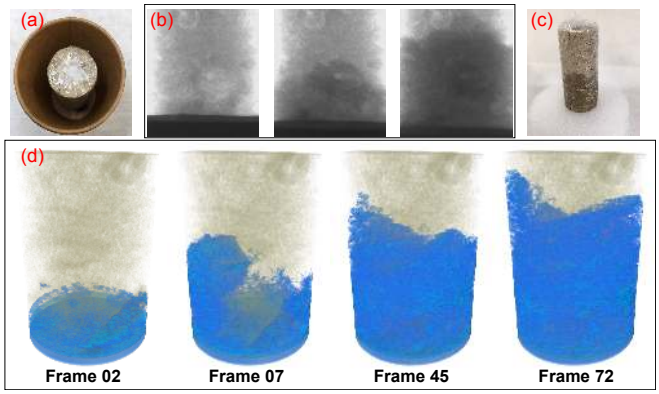

Fig. 12. Reconstruction results for the porous rock dataset. Images (a) and (c) represent respectively the rock before and after the scan. Some representative CT projections are given in (b). The rendering results in (d) show the absorption of liquid over time.

Dried snow fungus. Next, we show the re-hydration of dried mushrooms (*tremella fuciformis*, Figure 14). A direct comparison with the reference methods shows significantly improved detail and reconstruction quality compared to both SART-ROF and ST-tomography. Note that the first iteration of our algorithm produces approximately similar results to ST-tomography, but later iterations improve the result due to both the added key frames and the warp-and-project method using the estimated motion fields.

Objects in high viscosity liquid. In Figure 15 we show an experiment where solid objects were dropped into a high viscosity fluid and sink to the bottom under gravity. In this case the motion is fast enough to violate the assumption of negligible deformation between successive frames, which ST-tomography is based on. As a result, the shapes are blurred out and fine detail is lost, while our method manages sharp reconstructions of the solid.

Pills dissolved in water. In many medical pills, the drug is coated with a protective hull that dissolves in water, releasing the drug at a designed rate. Figure 16 shows the results of capturing this process in a CT scan. This is a challenging data set, since the motion is initially quite fast, and the drug particles are quite small. The whole sequence is only 3 minutes long and consists of 30 projections. As

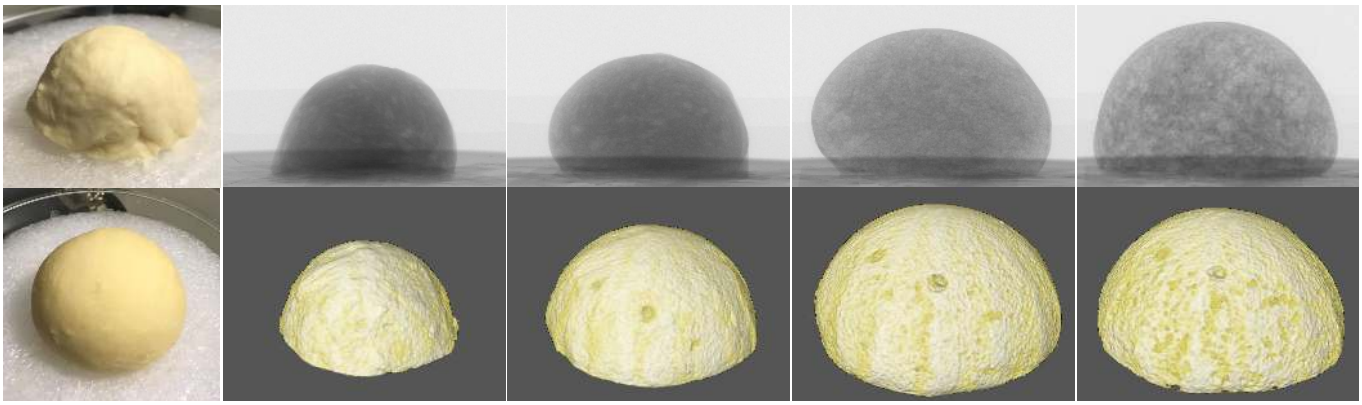

Fig. 13. 3D reconstructions of the dough dataset, with corresponding 2D X-ray images, as well as before-and-after photographs.

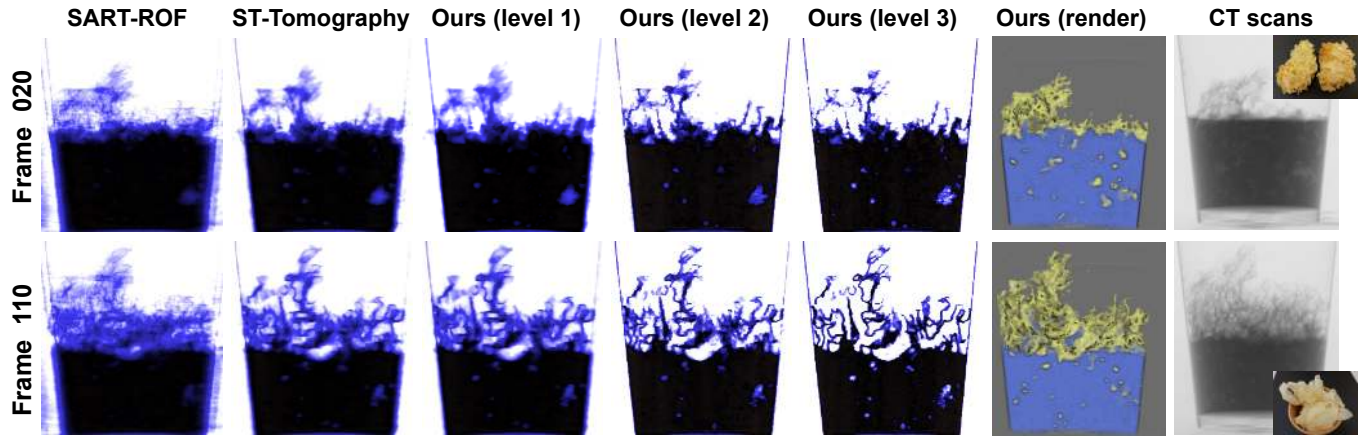

Fig. 14. Re-hydration of dried *tremella fuciformis*. This process results in a rapid volume expansion that tends to blur out the thin structures of the fungus in the comparison methods, while our method manages to reconstruct fine detail. From left to right: slice visualizations, 3D renderings, acquired projections during scanning and photographs of the sample.

with the previous datasets, we can see that our method produces crisper, more detailed results than the comparison methods.

Hydro-gel balls. Finally, in Figures 18 and 17 we show 2D slice visualizations and 3D renderings of Hydro-gel balls (Orbeez) absorbing water. Since the density of the water and the Hydro-gel balls is very similar, there is very little contrast between the two in the X-ray images or in the reconstruction. However, we can again see that the rapid absorption of water and the associated volume change create fast motions that cannot reliably be reconstructed by the competitor methods, whereas our approach produces sharp, clearly defined shapes for the balls.

5 CONCLUSIONS AND FUTURE WORK

In summary, we have presented a new method, warp-and-project tomography, for tomographic reconstruction of deforming objects. We perform quantitative comparisons on simulated data, as well as qualitative comparisons on real data from a number of different application domains. These experiments clearly show a significant improvement in the reconstruction quality compared to the state-of-the-art in both static and dynamic tomographic reconstruction.

We observed two major limitations of our method. First, the tracking algorithm requires features in order to work. If we are scanning a nearly feature-less volumetric object, the reconstruction of the motion field will only work on the boundary, but not inside the feature-less region. Second, our method will degrade with increasing speed of motion and an increasing geometric complexity of the volume. For example, objects that have a lot of high frequency details generally require more scans in a static reconstruction. For these objects our method can only work with slower motions. Also as a minor limitation, we do not consider motion blur during the acquisition of a single frame.

We believe that the developed method has significant applications in many domains. In fact the metal foam and rock data sets are already starting points of applications in mechanical engineering and geology / oil and gas exploration. In the future we plan on deeper investigations of these applications in collaboration with domain experts. We also intend to combine the method described here with

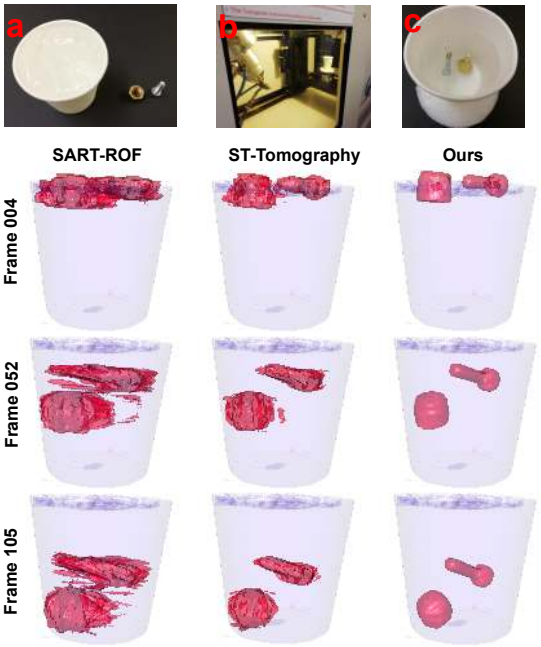

Fig. 15. Metal objects being dropped into a high viscosity liquid. First row: (a) and (b) show the metal objects respectively before and after the scan. Image (b) illustrates the scan acquisition. On the other rows, we represent the 3D results, where our method produces the sharpest reconstructions.

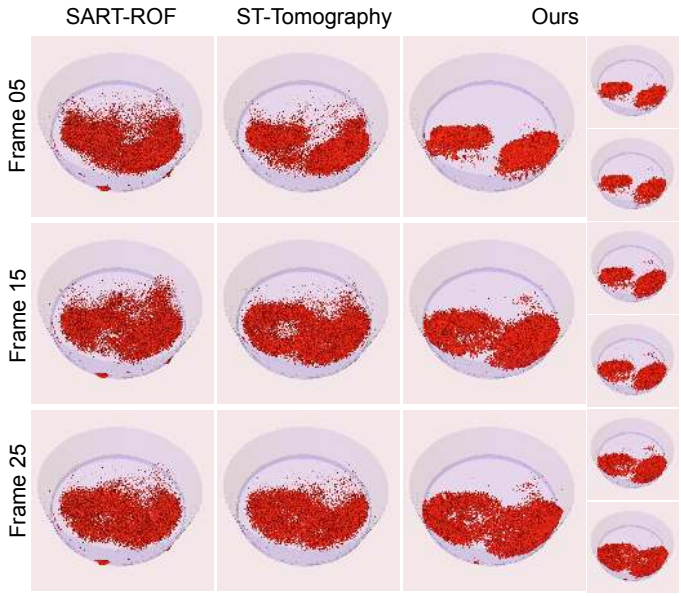

Fig. 16. Pills dissolving in water. This is our fastest dataset with only 30 projections captured in 3 minutes.

other X-ray tomography techniques such as phase contrast imaging, in order to boost the contrast in certain problematic data sets, such as the hydro-gel balls.

ACKNOWLEDGMENTS

This work was supported by King Abdullah University of Science and Technology as part of VCC Center Competitive Funding. The authors would like to thank the anonymous reviewers for their valuable comments. We thank Thomas Theussl for helping with the volume renderings, and Samuel Kortas for providing his assistance and efficient tools for high performance computing. We also thank Viswasanthi Chandra, Vinicius Lube, Qiang Fu, El-Hocine Bergou, Jing Ren, and Yuansi Tian for insightful discussions on the datasets.

REFERENCES

Anders H Andersen and Avinash C Kak. 1984. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic Imaging* 6, 1 (1984), 81–94.

Rushil Anirudh, Hyojin Kim, Jayaraman J Thiagarajan, K Aditya Mohan, Kyle Champley, and Timo Bremer. 2018. Lose the views: Limited angle CT reconstruction via implicit sinogram completion. In *Proc. CVPR*, Vol. 2.

Bradley Atcheson, Ivo Ihrke, Wolfgang Heidrich, Art Tevs, Derek Bradley, Marcus Magnor, and Hans-Peter Seidel. 2008. Time-resolved 3D capture of non-stationary gas flows. *ACM Trans. Graph.* 27, 5 (2008), 132.

Antonin Chambolle and Thomas Pock. 2011. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging and Vision* 40, 1 (2011), 120–145.

Guang-Hong Chen, Pascal Thériault-Lauzier, Jie Tang, Brian Nett, Shuai Leng, Joseph Zambelli, Zhihua Qi, Nicholas Bevins, Amish Raval, Scott Reeder, et al. 2012. Time-resolved interventional cardiac C-arm cone-beam CT: An application of the PICCS algorithm. *IEEE Trans. Med. Img.* 31, 4 (2012), 907–923.

Thomas De Schryver, Manuel Dierick, Marjolein Heyndrickx, Jeroen Van Stappen, Marijn A Boone, Luc Van Hoorebeke, and Matthieu N Boone. 2018. Motion compensated micro-CT reconstruction for in-situ analysis of dynamic processes. *Scientific reports* 8, 1 (2018), 7655.

Mingsong Dou, Sameh Khamis, Yury Degtyarev, Philip Davidson, Sean Ryan Fanello, Adarsh Kowdle, Sergio Orts Escolano, Christoph Rhemann, David Kim, Jonathan Taylor, et al. 2016. Fusion4D: Real-time performance capture of challenging scenes. *ACM Trans. Graph.* 35, 4 (2016), 114.

Gerrit E Elsinga, Fulvio Scarano, Bernhard Wieneke, and Bas W van Oudheusden. 2006. Tomographic particle image velocimetry. *Exp. Fluids* 41, 6 (2006), 933–947.

LA Feldkamp, LC Davis, and JW Kress. 1984. Practical cone-beam algorithm. *JOSA A* 1, 6 (1984), 612–619.

Pascal Getreuer. 2012. Rudin-Osher-Fatemi total variation denoising using split Bregman. *Image Processing On Line* 2 (2012), 74–95.

Richard Gordon, Robert Bender, and Gabor T Herman. 1970. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology* 29, 3 (1970), 471–481.

James Gregson, Ivo Ihrke, Nils Thuerey, and Wolfgang Heidrich. 2014. From capture to simulation: connecting forward and inverse problems in fluids. *ACM Trans. Graph.* 33, 4 (2014), 139.

James Gregson, Michael Krimerman, Matthias B Hullin, and Wolfgang Heidrich. 2012. Stochastic tomography and its applications in 3D imaging of mixing fluids. *ACM Trans. Graph.* 31, 4 (2012), 52–1.

Samuel W Hasinoff and Kiriakos N Kutulakos. 2007. Photo-consistent reconstruction of semitransparent scenes by density-sheet decomposition. *IEEE Trans. PAMI* 29, 5 (2007), 870–885.

François Hild, Hugo Leclerc, and Stéphane Roux. 2014. Performing DVC at the Voxel Scale. In *Advancement of Optical Methods in Experimental Mechanics, Volume 3*. 209–215.

Berthold KP Horn and Brian G Schunck. 1981. Determining optical flow. *Artificial Intelligence* 17, 1-3 (1981), 185–203.

Peter J Huber. 2011. Robust statistics. In *Int. Encyclopedia of Statistical Science*. 1248–1251.

Ivo Ihrke and Marcus Magnor. 2004. Image-based tomographic reconstruction of flames. In *Proc. SCA*. 365–373.

Takashi Ijiri, Shin Yoshizawa, Hideo Yokota, and Takeo Igarashi. 2014. Flower modeling via X-ray computed tomography. *ACM Trans. Graph.* 33, 4 (2014), 48.

Matthias Innmann, Michael Zollhöfer, Matthias Nießner, Christian Theobalt, and Marc Stamminger. 2016. VolumeDeform: Real-time volumetric non-rigid reconstruction. In *Proc. ECCV*. Springer, 362–379.

Clément Jailin and Stéphane Roux. 2018. Dynamic Tomographic Reconstruction of Deforming Volumes. *Materials* 11, 8 (2018), 1395.

Joël Lachambre, Julien Réthoré, Arnaud Weck, and Jean-Yves Buffiere. 2015. Extraction of stress intensity factors for 3D small fatigue cracks using digital volume correlation and X-ray tomography. *Int. J. Fatigue* 71 (2015), 3–10.

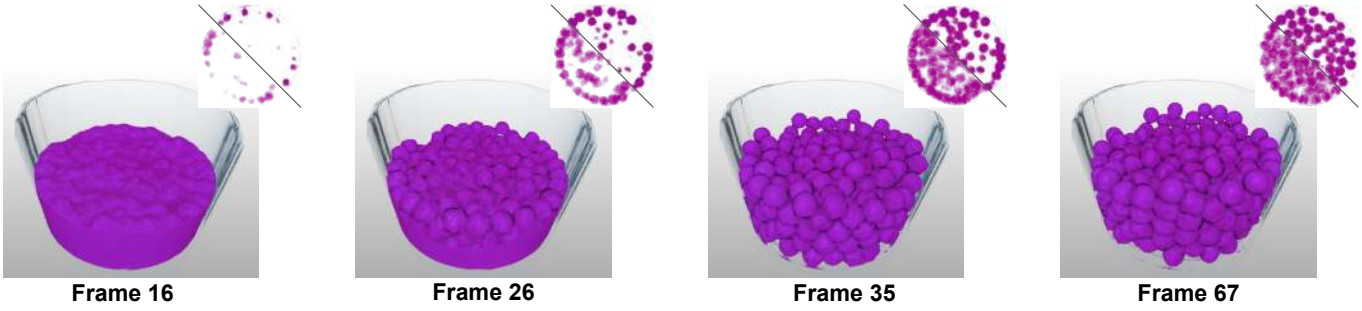

Fig. 17. 3D rendering of water absorption by Hydro-gel balls, with a slice comparison of the ST-tomography approach (lower left) and our method (upper right).

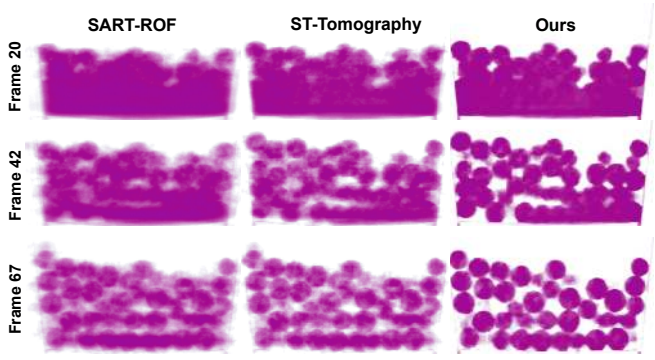

Fig. 18. Slice rendering of water absorption by Hydro-gel balls.

Hugo Leclerc, Stéphane Roux, and François Hild. 2015. Projection savings in CT-based digital volume correlation. *Exp. Mech.* 55, 1 (2015), 275–287.

Tianfang Li, Albert Koong, and Lei Xing. 2007. Enhanced 4D cone-beam CT with inter-phase motion model. *Medical physics* 34, 9 (2007), 3688–3695.

Yangyan Li, Xiaochen Fan, Niloy J Mitra, Daniel Chamovitz, Daniel Cohen-Or, and Baoquan Chen. 2013. Analyzing growing plants from 4D point cloud data. *ACM Trans. Graph.* 32, 6 (2013), 157.

Enric Meinhardt-Llopis, Javier Sánchez Pérez, and Daniel Kondermann. 2013. Horn-Schunck Optical Flow with a Multi-Scale Strategy. Image Processing On Line, 2013: 151–172, 2013. (2013).

Thilo F Morgeneyer, Lukas Helfen, Hazem Mubarak, and François Hild. 2013. 3D digital volume correlation of synchrotron radiation laminography images of ductile crack initiation: an initial feasibility study. *Exp. Mech.* 53, 4 (2013), 543–556.

Cyril Mory, Vincent Auvray, Bo Zhang, Michael Grass, Dirk Schäfer, S James Chen, John D Carroll, Simon Rit, Françoise Peyrin, Philippe Douek, et al. 2014. Cardiac C-arm computed tomography using a 3D + time ROI reconstruction method with spatial and temporal regularization. *Med. Phys.* 41, 2 (2014).

J Neggers, JPM Hoefnagels, MGD Geers, F Hild, and S Roux. 2015. Time-resolved integrated digital image correlation. *Internat. J. Numer. Methods Engrg.* 103, 3 (2015), 157–182.

Neal Parikh, Stephen Boyd, et al. 2014. Proximal algorithms. *Foundations and Trends® in Optimization* 1, 3 (2014), 127–239.

Alex Reche-Martinez, Ignacio Martin, and George Drettakis. 2004. Volumetric reconstruction and interactive rendering of trees from photographs. *ACM Trans. Graph.* 23, 3 (2004), 720–727.

Mai L Schmidt, Per R Poulsen, Jakob Toftegaard, Lone Hoffmann, David Hansen, and Thomas S Sørensen. 2014. Clinical use of iterative 4D-cone beam computed tomography reconstructions to investigate respiratory tumor motion in lung cancer patients. *Acta Oncologica* 53, 8 (2014), 1107–1113.

SM Shah, F Gray, JP Crawshaw, and ES Boek. 2016. Micro-computed tomography pore-scale study of flow in porous media: Effect of voxel resolution. *Advances in water resources* 95 (2016), 276–287.

SM Shah, J Yang, John P Crawshaw, O Gharbi, Edo S Boek, et al. 2013. Predicting porosity and permeability of carbonate rocks from core-scale to pore-scale using medical CT, confocal laser scanning microscopy and micro CT. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Abhishek Shastry, Paolo Palacio-Mancheno, Karl Braeckman, Sander Vanheule, Ivan Josipovic, Frederic Van Assche, Eric Robles, Veerle Cnudde, Luc Van Hoorebeke, and Matthieu Boone. 2018. In-Situ High Resolution Dynamic X-ray Microtomographic Imaging of Olive Oil Removal in Kitchen Sponges by Squeezing and Rinsing. *Materials* 11, 8 (2018), 1482.

Jan-Jakob Sonke, Lambert Zijp, Peter Remeijer, and Marcel van Herk. 2005. Respiratory correlated cone beam CT. *Med. Phys.* 32, 4 (2005), 1176–1186.

Wolfgang H Stuppy, Jessica A Maisano, Matthew W Colbert, Paula J Rudall, and Timothy B Rowe. 2003. Three-dimensional analysis of plant structure using high-resolution X-ray computed tomography. *Trends in Plant Science* 8, 1 (2003), 2–6.

Thibault Taillandier-Thomas, Stéphane Roux, and François Hild. 2016. Soft route to 4D tomography. *Phys. Rev. Letters* 117, 2 (2016), 025501.

Oliver Taubmann, Günter Lauritsch, Andreas Maier, Rebecca Fahrig, and Joachim Hornegger. 2015. Estimate, compensate, iterate: joint motion estimation and compensation in 4-D cardiac C-arm computed tomography. In *Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*. 579–586.

Nils Thuerey and Tobias Pfaff. 2018. MantaFlow. (2018). *http://mantaflow.com.*

Borislav Trifonov, Derek Bradley, and Wolfgang Heidrich. 2006. Tomographic reconstruction of transparent objects. In *Proc. EGSR.*

Huamin Wang, Miao Liao, Qing Zhang, Ruigang Yang, and Greg Turk. 2009. Physically guided liquid surface modeling from videos. *ACM Trans. Graph.* 28, 3 (2009), 90.

Sen Wang, Xinxin Zuo, Chao Du, Runxiao Wang, Jiangbin Zheng, and Ruigang Yang. 2018. Dynamic Non-Rigid Objects Reconstruction with a Single RGB-D Sensor. *Sensors* 18, 3 (2018), 886.

O. Weißenborn, S. Geller, M. Gude, F. Post, S. Praetorius, A. Voigt, and S. Aland. 2016. Deformation Analysis of Polymer Foams under Compression Load using in situ computed Tomography and Finite Element Simulation Methods. In *17th European conference on Composite Materials (ECCM-17).*

Guangming Zang, Mohamed Aly, Ramzi Idoughi, Peter Wonka, and Wolfgang Heidrich. 2018a. Super-Resolution and Sparse View CT Reconstruction. (2018).

Guangming Zang, Ramzi Idoughi, Ran Tao, Gilles Lubineau, Peter Wonka, and Wolfgang Heidrich. 2018b. Space-time Tomography for Continuously Deforming Objects. *ACM Trans. Graph.* 37, 4 (2018), 36.

Rongping Zeng, Jeffrey A Fessler, and James M Balter. 2007. Estimating 3-D respiratory motion from orbiting views by tomographic image registration. *IEEE Trans. Med. Img.* 26, 2 (2007), 153–163.

Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala. 2011. Building volumetric appearance models of fabric using micro CT imaging. *ACM Trans. Graph.* 30, 4 (2011), 44.

Qian Zheng, Xiaochen Fan, Minglun Gong, Andrei Sharf, Oliver Deussen, and Hui Huang. 2017. 4D Reconstruction of Blooming Flowers. *CGF* 36, 6 (2017), 405–417.