TomoFluid: Reconstructing Dynamic Fluid From Sparse View Videos

TomoFluid: Reconstructing Dynamic Fluid from Sparse View Videos

Guangming Zang[1] Ramzi Idoughi[1] Congli Wang[1] Anthony Bennett[1] Jianguo Du[1]
Scott Skeen[2] William L. Roberts[1] Peter Wonka[1] Wolfgang Heidrich[1]
[1]KAUST [2]Sandia National Laboratories

Abstract

Visible light tomography is a promising and increasingly popular technique for fluid imaging. However, the use of a sparse number of viewpoints in the capturing setups makes the reconstruction of fluid flows very challenging. In this paper, we present a state-of-the-art 4D tomographic reconstruction framework that integrates several regularizers into a multi-scale matrix free optimization algorithm. In addition to existing regularizers, we propose two new regularizers for improved results: a regularizer based on view interpolation of projected images and a regularizer to encourage reprojection consistency. We demonstrate our method with extensive experiments on both simulated and real data.

1. Introduction

Capturing fluid flows is a challenging reconstruction problem that is of great interest in computer vision and many scientific fields, *e.g.* for retrieving the fluid properties (*e.g.* temperature, species concentration, density and velocity) [43], validating simulation results [14, 40] or allowing flow editing and re-simulation [20].

A number of fluid imaging methods have been developed to tackle the different usage scenarios (we provide a brief literature review in the next section). Among those techniques, visible light tomography is the most used in computer vision and also finds frequent use in scientific imaging applications. In visible light tomography, simultaneous video sequences of the fluid are acquired from different angles, using either monochromatic or color cameras. From those videos tomographic reconstruction algorithms can be applied to reconstruct a sequence of 3D volumes, representing the densities of the fluid at different times. Some proposed approaches reconstruct the volumes at each time frame independently of each other, while others jointly optimize all frames in a single time-dependent reconstruction [20, 13, 14]. This second approach allows the introduction of temporal priors that improve the reconstruction quality. However, reconstructing complex three-dimensional fluid flows using visible-light tomography remains a difficult task for several reasons: First, many phenomena like absorption, scattering, refraction, occlusion impact the captured data. Second, the involved setups are often sophisticated and require a difficult calibration step. Finally, the data acquired is typically very sparse with few cameras capturing the fluid flow. Indeed, usually the cameras are placed relatively close to the fluid, which limits the number of cameras that can be used. Also, many real-world fluid experiments in scientific imaging take place inside special containers, and it is difficult or impossible to add many optical access windows to these containers without affecting the fluid flow. Finally, many fluid phenomena of interest are quite fast, which necessitates the use of high-speed or ultra high-speed cameras that add a significant expense to the setup (e.g. $80k or more per camera). For these reasons, it is often only feasible to use 2 or 3 cameras for an experiment. This leads to an under-constrained reconstruction problem that heavily relies on the use of good regularizers to produce satisfactory results.

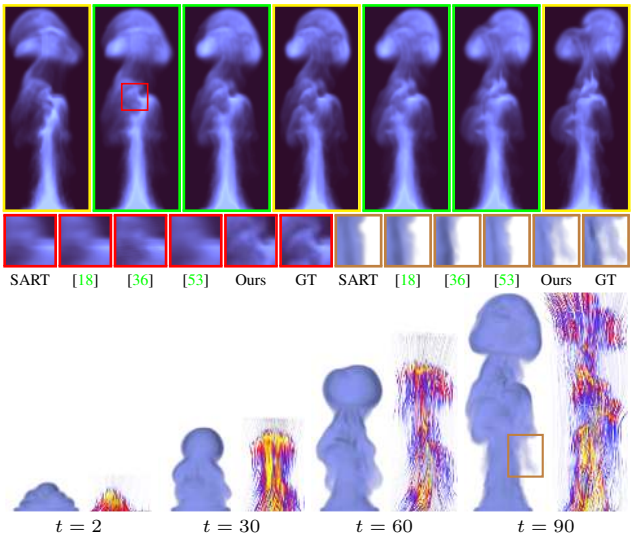

Figure 1: With sparse view videos as input (yellow box), TomoFluid generates three types of results: Novel view images for each time frame (green box) in the first row, physical-based temporally coherent density volumes and velocity fields (illustrated with streamlines) in the last row. Comparison between ground truth (GT) and different approaches is shown in the middle.

As shown in Figure 1, our work aims to reconstruct a 4D sequence of fluid flow from a sparse set of captured videos of this fluid. Our proposed approach is based on three primary observations. (1) In visible-light tomography, we often use very narrow fields of view, so that the camera geometry is almost orthographic, and thus the projection at 180° is significantly similar to the mirrored projection at 0°. (2) Although different points in the volume move differently in the image plane according to the camera parallax, the apparent motion of each point in the images is smooth with changes in viewing angle. (3) Furthermore, each pixel in one of the projection images corresponds to a line integral of the volume densities along a ray, and since the volume densities are the same in all views, this suggests that views can be interpolated with optimal transport-type methods. In practice, however, we found that direct interpolation of intensity images produces superior results (see supplemental material for more details).

Based on these observations, we propose an approach that can reconstruct high quality fluid flows from a limited number of input views (two or three). In particular, the main contributions of our work are:

- We built an effective framework for the reconstruction of fluid flows captured by a sparse set of high speed cameras. This framework provides both the 4D density field of the fluid and its deformation field.
- We propose a new regularizer based on view-interpolation and a reprojection consistency constraint that we incorporate in our framework.
- We extensively validate our method using both simulated and real experimental data from a range of different application scenarios.

2. Related Work

Dynamic Fluid Capture To characterize fluid flows, several modalities have been used to measure either 3D scalar-fields (*e.g.* temperature, species concentration and density) or 3D vector fields (*e.g.* velocity and vorticity) [43]. Planar Laser Induced Fluorescence (PLIF) [11], for example, measures the concentration of a fluorescent dye in fluid flows. A 3D extension of this technique has also been proposed [46]. To capture the 3D density field of the fluid several scanning approaches have been developed like laser line scanning [25, 17] or compressive structured light scanning [22]. However, the characterization of fluid flows requires the retrieval of the 3D velocity vector field. In the literature, two families of techniques have been used to measure the velocity field of a fluid: the tracer-based approaches and the tracer-free methods. The first family of methods consists of introducing tracers (particles, dye, *etc.*) in the fluid, then the velocity of the fluid is retrieved by tracking these tracers. Particle Image Velocimetry (PIV) and its different variants, such as tomographic PIV [15], synthetic aperture PIV [6], structured-light PIV [48, 47, 1] and plenoptic PIV [16, 41] are widely used in different fields to characterize fluid flows. For tracer-free approaches like Background Oriented Schlieren tomography (BOS) [19, 4], the phase change due refractive index differences in the fluid to track the flow. Most of these approaches retrieve only either the density field of the fluid or the velocity field. Moreover, they are specific to a given family of fluid. By using visible light tomography, we propose to jointly retrieve the density and the velocity fields of a large family of fluids.

Visible-light tomography In computer vision and graphics, most methods for fluid imaging are based on a tomographic reconstruction. This approach consists of retrieving a 3D density field representing the state of the fluid at a given time, from a set of 2D captured images (projections). In contrast to the medical field where computed tomography is based on X-ray scanning, visible-light is used to scan the fluids. Therefore, visible-light tomography methods were proposed to reconstruct 3D flames [24, 29], to image gas flows using Schlieren tomography [5] or to capture turbulent fluid mixtures using emission tomography [21]. The main shortcoming of visible light tomography for imaging fluids is the small number of vie points / projection images due to constraints in the hardware setup (*e.g.* cost of the cameras and space limitations). Usually less than 16 projections are used in the reconstruction, while in X-ray tomography hundreds or even thousands of projections can be used. This problem is commonly called the sparse-view tomography reconstruction problem, and it is heavily ill-posed – often the number of unknowns (voxels) exceeds the number of knowns (pixels) by one or two orders of magnitude.

Sparse-view reconstruction In medical X-ray CT applications, sparse-view scanning has been introduced to reduce radiation dose and shorten the acquisition time. Several techniques were proposed to improve the quality of the reconstruction. These methods are based on compressed sensing [38, 8], total variation based regularization [32, 31, 2], dictionary learning [9, 10, 33] or deep learning [23, 30, 37]. In the medical field, the deep learning approaches outperform the other methods, thanks to the large amount of available training data and the limited space of reconstructed shapes. However, it is very challenging to apply learning approaches to 4D fluid applications, because of the lack of training data and the diversity of flows. In addition, the memory consumption and time for training are also potentially serious issues. Finally, it is very hard to train one network to reconstruct a wide range of datasets in different application scenarios as presented in our work.

The sparse-view problem occurs also for dynamic X-ray

reconstructions, since it is impossible to take a large number of projections representing the same state of the scanned object. Zang *et al*. [52, 53, 50] proposed to jointly reconstruct the density and the deformation fields, by incorporating some spatial and temporal priors on these two fields. These approaches provide accurate reconstructions for the scanned objects. Nevertheless, the proposed scanning strategy to improve the reconstruction cannot be applied easily to high-speed fluid imaging setups, where the cameras have a fixed position and orientation.

Some attempts to improve the results of sparse tomographic reconstruction have been also proposed in the fluid imaging field. Gregson *et al*. [20] introduced a physically-based prior to reconstruct incompressible (i.e. divergence-free) flows. Okabe *et al*. [36] reconstructed a smoke volume from a sparse set of projections, by transferring the appearance information from the captured projections to novel viewing angles. This statistics-based approach (*i.e*. histogram matching and normalization) provides a plausible visualization of the smoke. However, there is no guarantee that the retrieved results correspond to the captured smoke. Eckert *et al*. [13] pushed the limit of the sparse tomography problem by introducing a single-view reconstruction approach for plumes. They compensate the lack of information by using physics-based and geometric priors. To improve the reconstruction of plumes, they added in [14] an inflow estimation module that only applies in very specific settings, and they used 5 viewing angles instead of one. While these methods result in *plausible* reconstruction for graphics, the overall accuracy and agreement with ground truth measurements is relatively poor and insufficient for scientific or engineering purposes.

3. Optimization Framework

System Overview We propose a novel dynamic reconstruction framework called TomoFluid. In this framework, several regularizers are introduced to constrain the solution space and achieve unprecedented reconstruction quality for fluid imaging. In the following section, we first present the camera model and pre-processing step of the input videos captured from optical cameras. Then we formulate the data-fitting term that involves the input videos. Next, the synthetic view regularizer, which is one of our key technical contributions will be explained in detail. This regularizer consists of using new estimated views to constrain the reconstructed volumes. The view estimation is done using an interpolation from existing videos. In addition, since only 2 or 3 views are available in our setting, we propose a visual hull regularizer to tackle this highly under-determined inverse problem. Finally, other regularizers, such as smoothness constraint, temporal coherence prior, density consistency prior, as well as a physically plausible regularizer used commonly for the incompressible flow estimation (*i.e*. divergence-free constraint) will be explained.

3.1. Framework Details

3.1.1 Camera Model

For a tomographic reconstruction, the fluid, represented by a sequence of discretized density fields $\mathbf{x} = (\mathbf{x}_t)_{1\leq t\leq T}$, is captured simultaneously by N cameras. Ideally, these cameras have to cover a $180°$ angular range around the fluid. At each time step t with $1 \leq t \leq T$, the i^{th} camera ($1 \leq i \leq N$) corresponding to the angle $\phi_i = (i-1)/N \cdot 180°$, captures the projection image $\mathbf{f}_{i,t}$ given by:

$$\mathbf{f}_{i,t} = \mathbf{K}_i \mathbf{x}_t + \mathbf{n_i} \tag{1}$$

where $\mathbf{K}_i$ and $\mathbf{n_i}$ are respectively the projection matrix (Radon transform operator) and the noise distribution corresponding to the i^{th} camera.

To retrieve the 3D volume sequence $\mathbf{x}$ from the captured videos, using the classical tomography reconstruction, we have to optimize the following loss function:

$$\mathcal{L}_{\text{Recon}} = \sum_{t=1}^{T} \sum_{i=1}^{N} \|\mathbf{K}_i \mathbf{x}_t - \mathbf{f}_{i,t}\|_2^2 \tag{2}$$

In our case, however, only few cameras are used to capture the fluid flow. Thus, at each time step t, only a sparse set of projection images $\mathbf{f}_{1,t} \dots \mathbf{f}_{M,t}$ is known, for $M < N$. This results in a sparse sampling in the frequency domain, according to the Fourier Slice Theorem [28]. Strong regularization is then needed to overcome this issue. For the following, we denote $\mathcal{M}$ as the set of viewing angles corresponding to the captured projections. Then, $\mathbf{f}_{i,t}$ ($\phi_i \in \mathcal{M}$) are the known projection images, and $\mathbf{f}_{j,t}$ ($\phi_j \in \mathcal{N} - \mathcal{M}$) are the unknown ones that we would like to estimate.

3.1.2 Data-fitting Term

We define our data-fitting term as the contribution of the M known views in the loss function given in Equation (2). This term can be written as follows:

$$\mathcal{L}_{\text{data}} = \sum_{t=1}^{T} \sum_{i|\phi_i \in \mathcal{M}} \|\mathbf{K}_i \mathbf{x}_t - \mathbf{f}_{i,t}\|_2^2 \tag{3}$$

3.1.3 Novel View Regularizer

The aim of introducing this regularizer is to take into account the contribution of the $N - M$ non-captured projections, in order to improve the reconstruction quality. To achieve this goal, we first estimate the novel views from the captured ones, as described below.

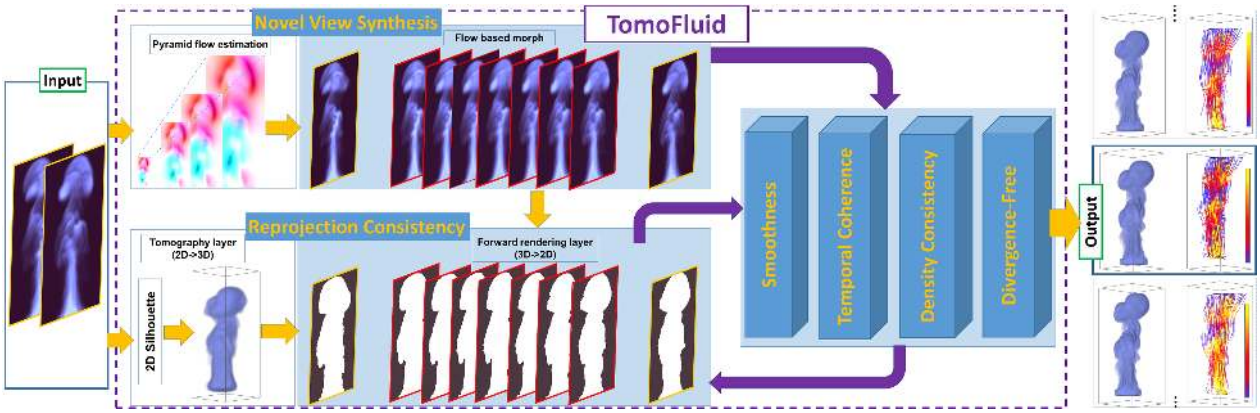

Figure 2: Overview of the architecture of our framework.

Flow-based warping: To estimate the novel views from the known orthographic view geometry, we interpolate the missing projections (*i.e.* novel views) $\mathbf{f}_{j,t}$ from its two neighboring captured projections $\mathbf{f}_{i,t}$ and $\mathbf{f}_{k,t}$, given $\phi_i < \phi_j < \phi_k$. Notice that: $\phi_j \notin \mathcal{M}$ and $\phi_i, \phi_k \in \mathcal{M}$.

We first compute the optical flow $\mathbf{v}_{i\to k}$ from the acquired image pair $\mathbf{f}_{i,t}$ and $\mathbf{f}_{k,t}$. A multi-scale scheme is utilized to tackle the large deformation [35]. Different weight settings are given to the horizontal and vertical directions, since the deformation occurs primly in the direction perpendicular to the rotation axis of our setting. Based on the angle indexing ratio $s = (j-i)/(k-i)$, we obtain two intermediate images $\widetilde{\mathbf{f}}_{i,t}$ and $\widetilde{\mathbf{f}}_{k,t}$ by warping $\mathbf{f}_{i,t}$ and $\mathbf{f}_{k,t}$, respectively:

$$\widetilde{\mathbf{f}}_{i,t} = \text{warp}(\mathbf{f}_{i,t}, s\mathbf{v}_{i\to k}) \tag{4}$$

$$\widetilde{\mathbf{f}}_{k,t} = \text{warp}(\mathbf{f}_{k,t}, -(1-s)\mathbf{v}_{i\to k}) \tag{5}$$

Morphing-based view interpolation: The missing projection $\mathbf{f}_{j,t}$ is interpolated using the warped images $\widetilde{\mathbf{f}}_{i,t}$ and $\widetilde{\mathbf{f}}_{k,t}$, as follows:

$$\mathbf{f}_{j,t} = (1-s)\widetilde{\mathbf{f}}_{i,t} + s\widetilde{\mathbf{f}}_{k,t} \tag{6}$$

which is very similar to a morphing process. At this step, all the $N - M$ missing projections are interpolated from the captured images.

Loss function: The loss function of the proposed synthetic view constraint is given in Equation (7). The accuracy of the projection estimation module depends highly on the angular distance to the captured projections (please refer to Section 4 for detail). We translate this point in the optimization framework, by introducing a confidence estimation weight ($0 < \omega_j < 1$) in the term $\mathcal{L}_{\text{novel}}$. The closer the angles are to the captured projections, the higher is the confidence in the estimation. For these closer angles the weight will be set close to 1. For the other angles, the weight will decrease gradually.

$$\mathcal{L}_{\text{novel}} = \sum_{t=1}^{T} \sum_{j|\phi_j \notin \mathcal{M}} \omega_j \left\| \mathbf{K}_j \mathbf{x}_t - \mathbf{f}_{j,t} \right\|_2^2 \tag{7}$$

Analysis: Our main goal in introducing the morphing-based interpolation is to regularize the missing projections in an energy consistent manner w.r.t. their neighboring captured projections. As mentioned in Section 1, the volume densities are the same in all views. Thus intuitively, an optimal-transport based interpolation is the approach to use for estimating the missing views from the captured ones. However, in practice, the proposed flow based warping operations outperform optimal-transport based methods (see supplementary for the experiments). In another perspective, considering the huge number of missing projections and interpolation operation involved, we anticipate our morphing method to be simple and efficient, especially in our 4D application scenario where hundreds of time frames are involved, making this extra time cost subtle for the full tomography reconstruction. Figure 3 shows a comparison between state-of-the-art optical flow with morphing [39], appearance transfer view interpolation [36], and ours, which is quantitatively the best and the fastest.

3.1.4 Reprojection Consistency Constraint

In our optimization framework, we introduce a reprojection consistency term ($\mathcal{L}_{\text{reprojection}}$), which is a variation of image-based visual hulls [34, 5, 13] and implemented in a simpler and more efficient way. However, in the referenced approaches the constraint is constructed only using the captured projections. In our work, we also generate silhouette masks corresponding to the missing views in the novel view interpolation step. These masks are then also involved in the computation of the visual hull from our reprojection consistency prior.

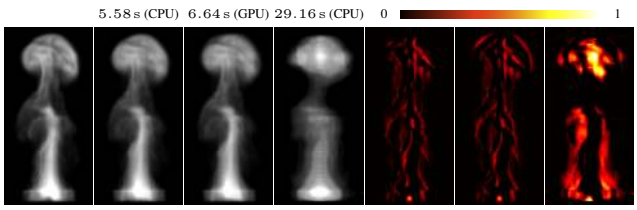

Figure 3: Flow-based warping comparison at 22° (morphed from 0° and 45°). Our simple yet efficient approach achieves satisfactory result compared to state-of-the-art optical flow [39] and appearance transfer [36] methods with the shortest time, where GT denotes ground truth.

As highlighted in Figure 2, for each time frame, we first create the silhouette images for the captured projections. Then, using these 2D projections and their silhouette images, the TV-regularized [18] method is applied to reconstruct a smooth volume due to lack of input information and artifact removal. A forward rendering operation is then applied on the initial volume to retrieve the object shape at different angles, corresponding to the estimated projections. These obtained silhouettes, combined with the masks of interpolated projections, correspond to the final results of reprojection consistency constraints for the estimated views. Though the mask from forward rendering is more conservative than that from the view interpolation step, in practice, by combining these two, a smooth visual hull in both spatial and temporal domain can be obtained.

In the forward rendering step, for each voxel at each angle ϕ_i, a ray will be generated between this voxel and the camera. The loss $\mathcal{L}_{\text{reprojection}}$ is defined as follows: For each intersection point of the generated ray and the projection images, if both the voxel value of the volume and the mask value of the interpolated projection is non-zero, then we set the pixel as 1, otherwise the value 0 will be given.

3.1.5 Other Regularizers

$$\mathcal{L}_{\text{smooth}}(\alpha,\beta,\gamma) = \sum_{t=1}^{T}\left[\alpha\left\|\nabla_S \mathbf{x}_t\right\|_1 + \beta\left\|\nabla_T \mathbf{x}_t\right\|_2^2\right] + \gamma\sum_{t=1}^{T-1}\sum_{d=x,y,z}\left\|\nabla_S \mathbf{u}_{t,d}\right\|_1 \quad (8)$$

Smoothness constraint: The fourth term that we introduce in our objective function is a smoothness prior that we apply both on the density volume of the fluid and the flow field: $\mathcal{L}_{\text{smooth}}$. In this prior we incorporate a spatial component with an L1-norm on both fields ($\mathbf{x}$ and $\mathbf{u}$) and a temporal component with an L2-norm only on the density volume. Indeed, the fluid flow is smoother according to the temporal dimension than the spatial dimension, where some discontinuities should be allowed.

where: ∇_S and ∇_T represent the spatial and temporal gradient operators. α, β and γ are respectively the weights of the spatial and temporal smoothness prior of the density volume and the spatial smoothness on the flow field.

Algorithm 1 TomoFluid algorithm

for t from 1 to T **do**
  $\mathbf{f}_{j,t} = \text{ViewSynthesis}(\mathbf{f}_{i,t})$ // interpolation
  $\mathbf{x}_t = \text{Tomography}(\mathbf{f}_{i,t})$
  $\mathbf{m}_t = \text{ReProjection}(\mathbf{x}_t)$
end for
repeat
  for t from 1 to T **do**
    $\mathbf{x}_t^* = \underset{\mathbf{x}_t}{\text{argmin}}\ \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{smooth}}(\alpha,\beta,0) + \delta\cdot\mathcal{L}_{\text{of}}$
      $+\mathcal{L}_{\text{novel}}$, w.r.t $\mathcal{L}_{\text{reprojection}}(\mathbf{m}_t) > 0$
    $\mathbf{u}_t^* = \underset{\mathbf{u}_t}{\text{argmin}}\ \mathcal{L}_{\text{smooth}}(0,0,\gamma) + \delta\cdot\mathcal{L}_{\text{of}}$
      $+\zeta\cdot\mathcal{L}_{\text{incompressible}}$
  end for
until Converge

Density consistency prior: The next prior that we add to our loss function is a density consistency over time. This term can be seen as a 3D version of the brightness constancy prior in the optical flow [27]. It ensures that the retrieved density volume in successive time steps should be consistent with a warping using the estimated flow field between these time steps. This assumption is generally valid for most fluid experiments, so long as diffusion happens at a time scale much slower than the camera frame rate. Note that this term involves both the density volume and the flow fields.

$$\mathcal{L}_{\text{of}} = \sum_{t=1}^{T-1}\left\|\nabla_T \mathbf{x}_t + \nabla_S \mathbf{x}_t\cdot\mathbf{u}_t\right\|_1 \quad (9)$$

Divergence-free prior: In the fluid simulation and imaging community, it is common to constrain the divergence of the flow to be equal to zero for incompressible fluid flows. This is simply the result of the mass-conservation assumption for the fluid. When the captured fluid flow can be assumed as incompressible, we incorporate the divergence-free prior ($\mathcal{L}_{\text{incompressible}}$) to our loss function.

$$\mathcal{L}_{\text{incompressible}} = \sum_{t=1}^{T-1}\text{DIV}(\mathbf{u}_t) \quad (10)$$

3.1.6 Optimization Details

By combining all the terms described previously, our loss function is given by:

$$(\mathbf{x}^*, \mathbf{u}^*) = \underset{\mathbf{x},\mathbf{u}}{\operatorname{argmin}} \; \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{novel}} + \mathcal{L}_{\text{smooth}} + \delta \cdot \mathcal{L}_{\text{of}} + \zeta \cdot \mathcal{L}_{\text{incompressible}} \tag{11}$$

where δ and ζ correspond respectively to the weights of spatial smoothness of the density volume, the density consistency and the divergence-free priors.

To solve the joint optimization problem in Equation (11), we split it into two sub-problems that we solve separately in an iterative and alternative fashion. Consequently, we get the scheme highlighted respectively in line 8 and line 10 of Algorithm 1. We apply the primal-dual Chambolle Pock algorithm [7] to efficiently tackle the involved discontinuities in the L1-terms in each sub-problem.

x-problem: The first sub-problem aims to reconstruct the density volume of the captured fluid. It involves the data-fitting term, the novel view regularizer, the reprojection consistency prior, the spatial and temporal regularizers of the density field (two first components of $\mathcal{L}_{\text{smooth}}$) and the density consistency prior. We follow the work [51] and use the PSART algorithm as solver to tackle the proximal operators of $\mathcal{L}_{\text{data}}$ and $\mathcal{L}_{\text{novel}}$ efficiently in a matrix-free manner.

u-problem: The second sub-problem is the flow field estimation. The corresponding objective function encompasses the spatial smoothness of the flow field (last term in $\mathcal{L}_{\text{smooth}}$), the density consistency prior and the divergence-free constraint. A multi-scale strategy [35] is applied to enable large deformations between the flow field volumes.

4. Results and Discussion

Baselines: In the following, we compare our approach to four baseline reconstruction techniques. The first baseline is an iterative tomographic reconstruction method named Simultaneous Algebraic Reconstruction Technique (**SART**) [3, 49], since it produces reasonably high-quality results while still being applicable to arbitrary camera models and application scenarios in practice [21, 26, 51]. The second baseline is the Bregman algorithm of Goldstein and Osher for TV-regularized denoising (**Getreuer**) [18]. The appearance transfer based method (**Okabe *et al.***) for fluid modeling proposed by [36] is compared as the third baseline. The last baseline (**Zang *et al.***) reconstructs fast deforming object with internal structure by X-ray CT scanner via a so called warp-and-project strategy proposed in [53]. Since the problem we are tackling is highly underdetermined, improved results are achieved for each method by constraining the reconstruction to the visual hull [29] of each data.

Parameters: The framework is implemented in C++ and it is parallelized using OpenMP. The experiments are conducted on a computer with 512 GB RAM and a dual-core 3.00GHZ Intel Xeon 2687W processor. In novel view synthesis stage, we estimate the flow with typical vertical and horizontal smoothness parameters of 1 and 10 respectively, performed at 8 different image pyramid scales. Additional details for parameters are provided in the supplement.

Table 1: PSNR/SSIM measurements for each approach. The value is averaged projection images over all 92 time frames at several degree (note as °). For each approach, average measurements for all generated projections (Avg. Proj.) and all reconstructed volumes (Avg. Vol.) are also presented. Refer to supplementary for more numerical analysis.

Method	05°	15°	25°	35°	Avg. Proj.	Avg. Vol.
SART	32.16/.872	27.55/.846	26.02/.852	28.25/.896	29.43/.868	25.54/.505
Getreuer	32.40/.879	27.58/.856	26.03/.857	28.33/.897	29.56/.876	25.58/.512
Okabe	29.80/.871	26.82/.832	25.96/.807	27.89/.829	28.23/.843	25.24/.467
Zang	32.69/.923	28.59/.895	27.15/.883	28.74/.903	29.81/.905	25.76/.538
Ours	**36.55/.978**	**30.09/.936**	**28.80/.923**	**31.17/.946**	**32.36/.950**	**27.72/.671**

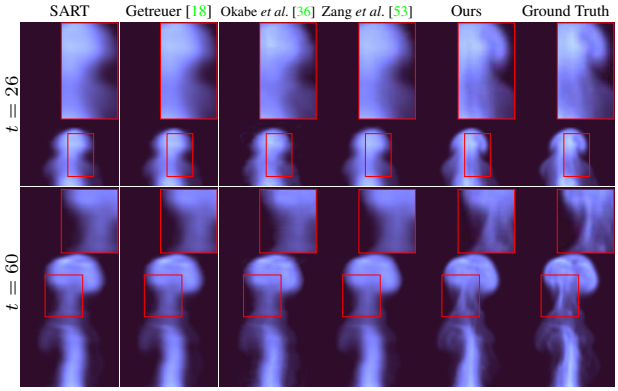

Figure 4: Projection images comparison for different methods.

4.1. Synthetic results

We first validate our algorithm on the synthetic fluid flow data, generated from Mantaflow [42]. In this experiment, as shown in Figure 4 and Table 1, a comprehensive evaluation is conducted between our method and the baselines. A time sequence of 100 fluid volumes is generated with a size of $100 \times 150 \times 100$ for each volume. For each time step, 180 orthogonal projection images, uniformly distributed over 180°, are then generated. The 100 volumes and 18000 projection images act as ground truth data. To simulate the real capture environment, three images at respectively 0°, 45°, and 90° for each time step are used as input for all methods.

The values of PSNR and SSIM, for both projections and volumes, are shown in Table 1 for all compared methods. We can observe that for all methods, the quality of projections is higher when they are closer to the real captured angles (*i.e.* 0°, 45°, or 90°). On the other hand, the comparison for the angles 25° and 65° yield the lowest values of

PSNR and SSIM (see the supplementary table). As shown in Figure 4, although a sharper boundary can be visually observed comparing to **SART**, **Getreuer** and **Zang *et al.***, the appearance transfer based method **Okabe *et al.*** fails to achieve a better numerical result than other methods. As mentioned in their paper, **Okabe *et al.*** mainly focus on how to apply this statistics based method to obtain a better visualization results instead of numerical analysis. In contrast, **Zang *et al.*** achieves the best results among all the baseline methods, mainly due to their effective spatial and temporal priors. They failed, however, to obtain good reconstructions at the novel estimated views in this sparse view setting, and significant blurring can be observed, as shown in Figure 4. Finally, our method (**Ours**) achieves best results for projections at any angle and reconstructed volumes, in terms of both PSNR/SSIM evaluation in Table 1 as well as for the qualitative comparison in Figure 4. In Table 1, the PSNR/SSIM results at several views are presented, and the averaged values of PSNR/SSIM for all projections and volumes are also evaluated.

4.2. Real world results

4D Soot Imaging of the turbulent flame: The setup for capturing the combustion process is done using a non-premixed turbulent jet flame with ethylene fuel at a Reynolds (Re) number of 2,600 and 10,400. As shown in Figure 5, the high speed videos were collected from three Photron SA-Z cameras placed at 0° (Camera 1), 45° (Camera 2), and 90° (Camera 3). For a qualitative evaluation for the reconstruction methods, we applied planar laser induced incandescence (LII) at 117° (ICCD) with a frequency doubled Quanta-Ray Pro-Series pulsed Nd:YAG laser at 10 Hz. A projection image from the 2,600 Re is shown in Figure 5(b). The normalized LII image, and reconstructed slices from our method (**Ours**) and **Zang *et al.*** [53] method at 117° are shown respectively in Figure 5(c-e) indicating sharper results with reasonably finer soot structures that can be observed from our method.

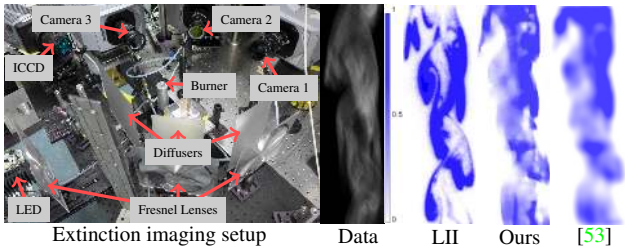

Figure 5: Extinction imaging and planar laser induced incandescence setup (LII), where one soot raw image from the video captured by camera 1 is shown. We compare normalized LII image taken at 117° (ground truth) with the same slice reconstructed by our method and **Zang *et al.*** [53].

Volume reconstructions using our method at 4 time steps can be seen in the top row of Figure 6. A comparison between each of the baseline methods and our method can be observed in the bottom of Figure 6 with each from split between baseline methods on the left and our method on the right. Figure 7 shows a further comprehensive comparison between our approach and the-state-of-the-art method [53], in terms of the projection images, the volume slices, and dynamic volumes at different angles and time frames. From all of these metrics, the unprecedented reconstruction quality from our method can be observed.

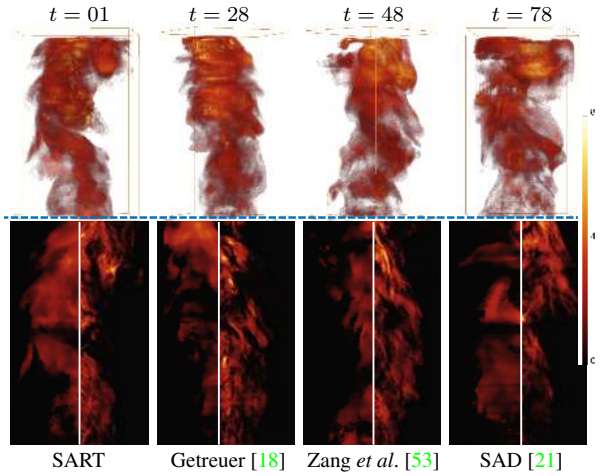

Figure 6: Reconstructed results for a highly turbulent flame. First row: Reconstructed results from our method for different time frames and viewing angles. Second row: A comparison of the baseline approach/Our approach (split left/right respectively) for volume slice. The baseline methods on the left side are: Plain **SART** reconstruction, **Getreuer** [18], **Zang *et al.*** [53], and sum-of-absolute-differences (**SAD**) [21] regularized reconstruction.

4D Imaging of a Mixing Fluid Process: The second application we performed is the imaging of a mixing fluid process with two high speed cameras. The experimental setup is shown in Figure 8. There are two cameras placed respectively at 0° and 90°. The light source is generated from the bottom of the glass tank with the help of reflector. The results from our method with different time frames and rotated angles are presented in Figure 8(c). Additional volumetric slice comparison for two different time steps between **Ours** and **Zang *et al.*** [53] is also performed in Figure 8(d). Although only two videos are available in this setting, a significant improvement in terms of reconstruction quality can be noticed from **Ours**.

4D Imaging of a Fuel Injection Process: In the final application, we tried to model the process of fuel injection for spray plume data. The setup for the experiments is illustrated in Figure 9(a). The details of this experiment are described in [12]. The light source is generated from a pulsed-driven

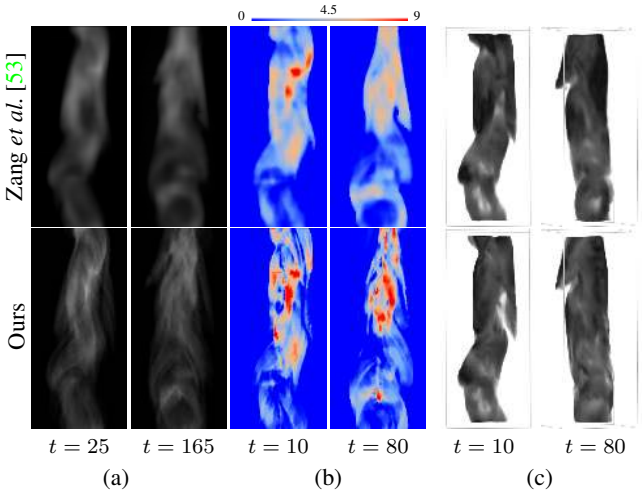

Figure 7: Reconstruction results for soot data, acquired using setup in Figure 5. (a) Single view comparison (at 25°). (b) Volume comparison (sliced). (c) Volume comparison (rotated).

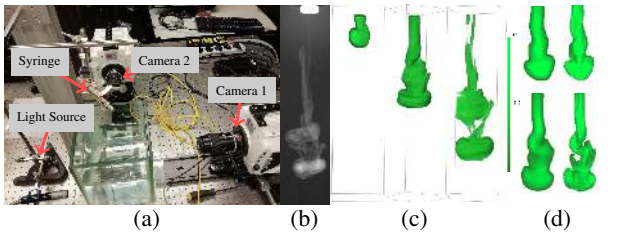

Figure 8: (a) The set up for capturing fluid images with two views. (b) One captured image from the setup (c) Our reconstruction at different time frames and viewing angles. (d) Slice comparison between **Zang *et al.*** [53] and our method for two different time frames.

white LED. A Photron SA-X2 high-speed camera was employed at 40 kfps to capture the extinction. The size of the captured image is 552×512. The injector was rotated at 0°, 40°, and 90°, respectively. At each view angle, the averaged image of three repeats was used for the 3D reconstruction. Note that this experiment is highly repeatable, if the same initial conditions are applied. Figure 9(b) shows two images captured at different time frames from 0 degree with the ambient gas pressure as 4kPa (left), and their flow field visualization estimated (right) from our TomoFluid framework (illuminated with streamlines). In Figure 9, from left to right, we compare respectively the reconstruction results with the sophisticated visible-light tomography method that regularized with the sum-of-absolute-differences (**SAD**) prior [21], state-of-the-art dynamic tomographic reconstruction method (**Zang *et al.***), and **Ours**. The first row illustrates the volume slice comparison, while the second row shows the volume reconstruction for one time frame. We can observe that there are obvious discontinuity between different time frames for **SAD** regularized method [21], since only a spatial prior is applied for this approach. In contrast, with both spatial and temporal priors, better results can be noticed when **Zang *et al.*** [53] is used. Finally, with a combination of powerful regularizers for view interpolation, reprojection consistency, spatial and temporal smoothness, and jointly optimization framework, our method provides more details of the spray.

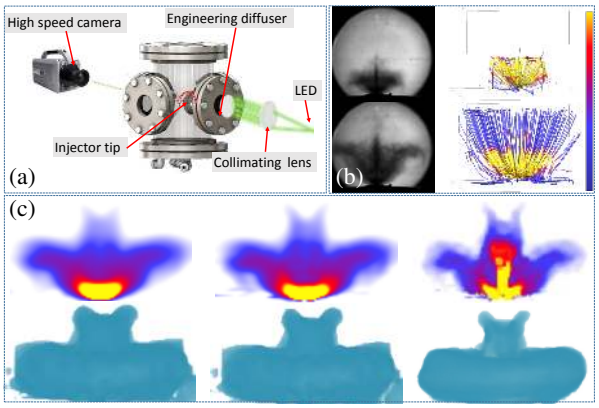

Figure 9: (a) The set up for capturing spray images. (b) Captured images at two time frames (left) and estimated flow field illustrated with streamlines. (c) Reconstruction results for **SAD** regularized method [21], **Zang *et al.*** [53], and **Ours**. First and second row of (c) represents respectively reconstructed slice and the volume.

5. Conclusion and Future work

We presented TomoFluid, a framework for the fluid reconstruction from sparse view videos. Although our work mainly focuses on dynamic fluid imaging, the presented methods could potentially be useful for various tasks such as X-ray tomography reconstruction with flow-based view interpolation, or fluid re-simulation with an estimated flow field between successive volumes. One limitation of our work is that the introduced novel view regularizer lacks a thorough physical interpretation, although we have demonstrated that it works very well in practice. Nonetheless we believe that for the future there lies much promise in exploring physically motivated view interpolation methods to achieve even better reconstruction results for sparse view tomography problems. Other future directions include combining phase contrast techniques [5, 44, 45] to handle transparent phenomena, such as gas flows.

6. Acknowledgements

This work was supported by KAUST as part of VCC and CCRC Center Competitive Funding and KAUST Competitive Research Grants. We thank the anonymous reviewers for their insightful comments, and Yuansi Tian for helping with the data collection.

References

[1] A. Aguirre-Pablo, A. B. Aljedaani, J. Xiong, R. Idoughi, W. Heidrich, and S. T. Thoroddsen. Single-camera 3D PTV using particle intensities and structured light. *Experiments in Fluids*, 60(2):25, 2019. 2

[2] M. Aly, G. Zang, W. Heidrich, and P. Wonka. TRex: A tomography reconstruction proximal framework for robust sparse view X-ray applications. *arXiv preprint arXiv:1606.03601*, 2016. 2

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

[4] B. Atcheson, W. Heidrich, and I. Ihrke. An evaluation of optical flow algorithms for background oriented Schlieren imaging. *Experiments in Fluids*, 46(3):467–476, 2009. 2

[5] B. Atcheson, I. Ihrke, W. Heidrich, A. Tevs, D. Bradley, M. Magnor, and H.-P. Seidel. Time-resolved 3D capture of non-stationary gas flows. *ACM Trans. Graph.*, 27(5):132, 2008. 2, 4, 8

[6] J. Belden, T. T. Truscott, M. C. Axiak, and A. H. Techet. Three-dimensional synthetic aperture particle image velocimetry. *Measurement Science and Technology*, 21(12):125403, 2010. 2

[7] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging and Vision*, 40(1):120–145, 2011. 6

[8] G.-H. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical physics*, 35(2):660–663, 2008. 2

[9] Y. Chen, L. Shi, Q. Feng, J. Yang, H. Shu, L. Luo, J.-L. Coatrieux, and W. Chen. Artifact suppressed dictionary learning for low-dose CT image processing. *IEEE Trans. Med. Img.*, 33(12):2271–2292, 2014. 2

[10] Y. Chen, Z. Yang, Y. Hu, G. Yang, Y. Zhu, Y. Li, W. Chen, C. Toumoulin, et al. Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means. *Physics in Medicine & Biology*, 57(9):2667, 2012. 2

[11] J. Crimaldi. Planar laser induced fluorescence in aqueous flows. *Experiments in Fluids*, 44(6):851–863, 2008. 2

[12] J. Du, G. Zang, B. Mohan, R. Idoughi, J. Sim, T. Fang, P. Wonka, W. Heidrich, and W. Roberts. Study of spray structure from non-flash to flash boiling conditions with space-time tomography. *International Symposium on Combustion*, 2020. 7

[13] M.-L. Eckert, W. Heidrich, and N. Thuerey. Coupled fluid density and motion from single views. In *CGF*, volume 37, pages 47–58. Wiley Online Library, 2018. 1, 3, 4

[14] M.-L. Eckert, K. Um, and N. Thuerey. Scalarflow: a large-scale volumetric data set of real-world scalar transport flows for computer animation and machine learning. *ACM Trans. Graph.*, 38(6):1–16, 2019. 1, 3

[15] G. E. Elsinga, F. Scarano, B. Wieneke, and B. W. van Oudheusden. Tomographic particle image velocimetry. *Experiments in Fluids*, 41(6):933–947, 2006. 2

[16] T. W. Fahringer, K. P. Lynch, and B. S. Thurow. Volumetric particle image velocimetry with a single plenoptic camera. *Measurement Science and Technology*, 26(11):115201, 2015. 2

[17] C. Fuchs, T. Chen, M. Goesele, H. Theisel, and H.-P. Seidel. Density estimation for dynamic volumes. *Computers & Graphics*, 31(2):205–211, 2007. 2

[18] P. Getreuer. Rudin-Osher-Fatemi total variation denoising using split bregman. *Image Processing On Line*, 2:74–95, 2012. 1, 5, 6, 7

[19] E. Goldhahn and J. Seume. The background oriented schlieren technique: sensitivity, accuracy, resolution and application to a three-dimensional density field. *Experiments in Fluids*, 43(2-3):241–249, 2007. 2

[20] J. Gregson, I. Ihrke, N. Thuerey, and W. Heidrich. From capture to simulation: connecting forward and inverse problems in fluids. *ACM Trans. Graph.*, 33(4):139, 2014. 1, 3

[21] J. Gregson, M. Krimerman, M. B. Hullin, and W. Heidrich. Stochastic tomography and its applications in 3D imaging of mixing fluids. *ACM Trans. Graph.*, 31(4):52–1, 2012. 2, 6, 7, 8

[22] J. Gu, S. K. Nayar, E. Grinspun, P. N. Belhumeur, and R. Ramamoorthi. Compressive structured light for recovering inhomogeneous participating media. *IEEE Trans. PAMI*, 35(3):1–1, 2012. 2

[23] Y. Han and J. C. Ye. Framing U-Net via deep convolutional framelets: Application to sparse-view CT. *IEEE Trans. Med. Img.*, 37(6):1418–1429, 2018. 2

[24] S. W. Hasinoff and K. N. Kutulakos. Photo-consistent reconstruction of semitransparent scenes by density-sheet decomposition. *IEEE Trans. PAMI*, 29(5):870–885, 2007. 2

[25] T. Hawkins, P. Einarsson, and P. Debevec. Acquisition of time-varying participating media. Technical report, UNIVERSITY OF SOUTHERN CALIFORNIA MARINA DEL REY CA INST FOR CREATIVE , 2005. 2

[26] F. Heide, J. Gregson, G. Wetzstein, R. Raskar, and W. Heidrich. Compressive multi-mode superresolution display. *Optics Express*, 22(12):14981–14992, 2014. 6

[27] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, 1981. 5

[28] J. Hsieh et al. Computed tomography: principles, design, artifacts, and recent advances. SPIE Bellingham, WA, 2009. 3

[29] I. Ihrke and M. Magnor. Image-based tomographic reconstruction of flames. In *Proc. SCA*, pages 365–373, 2004. 2, 6

[30] A. Kofler, M. Haltmeier, C. Kolbitsch, M. Kachelrieß, and M. Dewey. A U-Nets cascade for sparse view computed tomography. In *International Workshop on Machine Learning for Medical Image Reconstruction*, pages 91–99. Springer, 2018. 2

[31] Y. Liu, Z. Liang, J. Ma, H. Lu, K. Wang, H. Zhang, and W. Moore. Total variation-stokes strategy for sparse-view X-ray CT image reconstruction. *IEEE Trans. Med. Img.*, 33(3):749–763, 2013. 2

[32] Y. Liu, J. Ma, Y. Fan, and Z. Liang. Adaptive-weighted total variation minimization for sparse data toward low-dose x-

ray computed tomography image reconstruction. *Physics in Medicine & Biology*, 57(23):7923, 2012. 2

[33] Y. Lu, J. Zhao, and G. Wang. Few-view image reconstruction with dual dictionaries. *Physics in Medicine & Biology*, 57(1):173, 2011. 2

[34] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan. Image-based visual hulls. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 369–374. ACM Press/Addison-Wesley Publishing Co., 2000. 4

[35] E. Meinhardt-Llopis, J. S. Pérez, and D. Kondermann. Horn-schunck optical flow with a multi-scale strategy. *Image Processing On Line*, 2013:151–172, 2013. 4, 6

[36] M. Okabe, Y. Dobashi, K. Anjyo, and R. Onai. Fluid volume modeling from sparse multi-view images by appearance transfer. *ACM Trans. Graph.*, 34(4):93, 2015. 1, 3, 4, 5, 6

[37] T. Shen, X. Li, Z. Zhong, J. Wu, and Z. Lin. R^2-Net: Recurrent and recursive network for sparse-view CT artifacts removal. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 319–327. Springer, 2019. 2

[38] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine & Biology*, 53(17):4777, 2008. 2

[39] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8934–8943, 2018. 4, 5

[40] T. Takahashi and M. C. Lin. Video-guided real-to-virtual parameter transfer for viscous fluids. *ACM Trans. Graph.*, 38(6):237, 2019. 1

[41] Z. P. Tan and B. S. Thurow. Time-resolved 3D flow-measurement with a single plenoptic-camera. In *AIAA Scitech 2019 Forum*, page 0267, 2019. 2

[42] N. Thuerey and T. Pfaff. MantaFlow, 2018. *http://mantaflow.com*. 6

[43] C. Tropea and A. L. Yarin. *Springer handbook of experimental fluid mechanics*. Springer Science & Business Media, 2007. 1, 2

[44] C. Wang, X. Dun, Q. Fu, and W. Heidrich. Ultra-high resolution coded wavefront sensor. *Optics Express*, 25(12):13736–13746, 2017. 8

[45] C. Wang, Q. Fu, X. Dun, and W. Heidrich. Quantitative phase and intensity microscopy using snapshot white light wavefront sensing. *Scientific Reports*, 9:13795, 2019. 8

[46] R. Wellander, M. Richter, and M. Aldén. Time-resolved (khz) 3d imaging of oh plif in a flame. *Experiments in fluids*, 55(6):1764, 2014. 2

[47] J. Xiong, Q. Fu, R. Idoughi, and W. Heidrich. Reconfigurable rainbow PIV for 3D flow measurement. In *2018 IEEE International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2018. 2

[48] J. Xiong, R. Idoughi, A. A. Aguirre-Pablo, A. B. Aljedaani, X. Dun, Q. Fu, S. T. Thoroddsen, and W. Heidrich. Rainbow particle imaging velocimetry for dense 3D fluid velocity imaging. *ACM Trans. Graph.*, 36(4):36, 2017. 2

[49] M. Yan. Convergence analysis of SART: optimization and statistics. *International Journal of Computer Mathematics*, 90(1):30–47, 2013. 6

[50] G. Zang. *Space-Time Tomographic Reconstruction of Deforming Objects*. PhD thesis, 2020. 3

[51] G. Zang, M. Aly, R. Idoughi, P. Wonka, and W. Heidrich. Super-resolution and sparse view CT reconstruction. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 137–153, 2018. 6

[52] G. Zang, R. Idoughi, R. Tao, G. Lubineau, P. Wonka, and W. Heidrich. Space-time tomography for continuously deforming objects. *ACM Trans. Graph.*, 37(4):36, 2018. 3

[53] G. Zang, R. Idoughi, R. Tao, G. Lubineau, P. Wonka, and W. Heidrich. Warp-and-project tomography for rapidly deforming objects. *ACM Trans. Graph.*, 38(4), 2019. 1, 3, 6, 7, 8