Statistical Tomography of Microscopic Life

Aviad Levis,   Yoav Y. Schechner,   Ronen Talmon
Viterbi Faculty of Electrical Engineering
Technion - Israel Institute of Technology, Haifa, Israel
aviad.levis@gmail.com ; yoav@ee.technion.ac.il ; ronen@ee.technion.ac.il

Abstract

We achieve tomography of 3D volumetric natural objects, where each projected 2D image corresponds to a different specimen. Each specimen has unknown random 3D orientation, location, and scale. This imaging scenario is relevant to microscopic and mesoscopic organisms, aerosols and hydrosols viewed naturally by a microscope. In-class scale variation inhibits prior single-particle reconstruction methods. We thus generalize tomographic recovery to account for all degrees of freedom of a similarity transformation. This enables geometric self-calibration in imaging of transparent objects. We make the computational load manageable and reach good quality reconstruction in a short time. This enables extraction of statistics that are important for a scientific study of specimen populations, specifically size distribution parameters. We apply the method to study of plankton.

1. Introduction

The environment is teeming with microscopic and mesoscopic objects, whose study both structurally and as a population is central to science [39]. Examples include aerosols, hydrosols, and small organisms which have wide varieties of classes. It is possible to take samples to a lab and measure their 3D statistics, while they are few and sit still. However, such an approach has drawbacks. In a lab, organisms do not behave as in nature. To minimize biasing nature, in-situ microscopes are used, as illustrated in Fig. 1. There, and also in many lab studies, specimens do not sit still. Hence each image is of a different specimen at random orientation, scale, and location. Moreover, as there is a priority on gathering large data, time may not be allocated to scan each of the specimens in 3D using common microscopic methods. Due to the randomness of projection of volumetric objects, the data are not well amenable to 2D image analysis. Thus, manual labor is used, greatly limiting the statistical and structural understanding of nature. Consider plankton (Fig. 2), mesoscopic organisms which are a fundamental building block in the oceanic food chain. Their population state is indicative of climate changes and the state of the marine eco-system [17]. In-situ microscopes [39,40] give
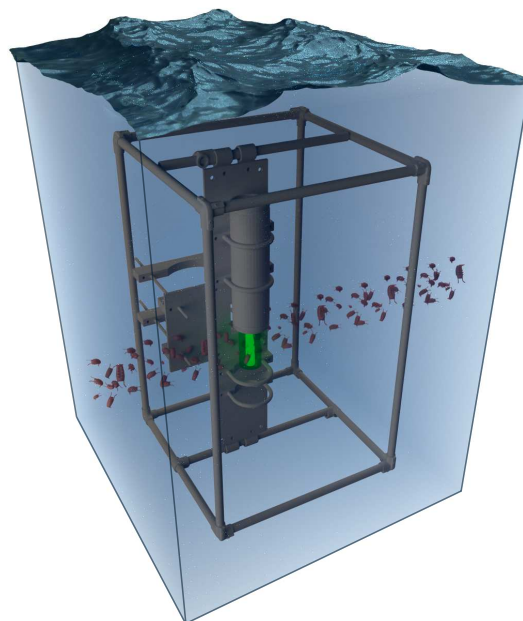

Figure 1. A random set of moving organisms at different orientations and scales is imaged using a submerged microscope. Statistical tomographic reconstruction yields self-calibrated estimation of the 3D density map, along with the unknown orientation and scale of each sample. Estimation of the population's size-distribution is thus enabled, jointly with 3D recovery (3D-POP). Microscope 3D model courtesy of Jaffe Lab.

high-temporal resolution in 2D [1,15], while 3D imaging of fixated plankton is done in the laboratory [22,27].

To address these challenges, we seek in-situ volumetric recovery and natural population statistics. In the case of biological specimens, this can provide high-quality estimates of bio-mass [38] and size-distributions [6]. The idea is to use an ensemble of 2D images, each of a different specimen, and recover all that is needed based on this random data. Our aim is to recover a statistical 3D tomography of microscopic life. Unlike common computational tomography [16,18,22] the images acquired are uncalibrated: the rotation, scale, and locations of the specimens are random and unknown.

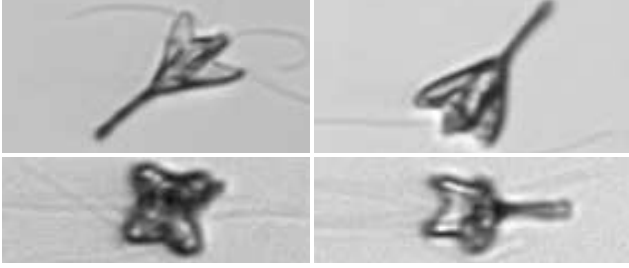We build on *Single Particle Reconstruction* (SPR) [12].

Figure 2. In-situ images of *Pyramimonas Longicauda*, from the Woods Hole Oceanographic Institute annotated plankton dataset [46]. Each organism has a different orientation and possibly scale, however, there is 3D structural resemblance.

In Cryo-Electron Microscopy (Cryo-EM) [25,34,52], multiple replicas of a molecule of a particular class are frozen. Each molecule has a random unknown orientation. Organisms, however, have greater variations within a class. The most dominant variation is scale. These variations cause typical SPR algorithms to fail when applied to biological ensembles. The task defined in this work is estimation of the scale of each sample, in conjunction with the orientation and locations, as part of the volume recovery process. As this task seeks recovery of 3D structure and the size-distribution of a population, we denote it *3D-POP*.

Meeting this challenge is computationally demanding. However, it has become feasible with increasing computing power. This feasibility is similar to enabling of other complex computer vision inverse problems, involving multiple reflections, non-line-of-sight imaging [5,7], scattering media [19,21,23,24,28,29,51] and blind estimation [32,47,48]. In a sense, we tackle a problem of self-calibration and 3D scene recovery. These topics were usually dealt in the context of opaque objects [2,35]. Under the microscope, however, many interesting organisms are transparent [20,33], hence we need to generalize self-calibration to multi-view imaging of transparent objects.

2. Theoretical Background

Computational tomography (CT) estimates a volumetric density distribution using multi-view projections. This section describes core concepts of tomographic recovery and orientation estimation using projections.

2.1. Tomography Using Known Orientations

Denote 3D spatial coordinates by $\boldsymbol{x}=(x_1, x_2, x_3)^{\mathrm{T}}$, where $(\cdot)^{\mathrm{T}}$ denotes transposition. The orientation of a sample with volumetric distribution $\beta(\boldsymbol{x})$ is defined by a 3D rotation matrix $\mathbf{R}$. This rotation is relative to the microscope reference frame. Define a projection operator of a rotated sample as

$$P_{\mathbf{R}}(\beta)(x_1, x_2) = \int \beta\left(\mathbf{R}^{-1}\boldsymbol{x}\right) \mathrm{d}x_3. \qquad (1)$$

Here x_3 is the integration axis and x_1, x_2 are the image axes. Denote a set of N orientation matrices as

$$\Omega = \{\mathbf{R}_1, ..., \mathbf{R}_N\}. \qquad (2)$$

The sets of measured and modeled projections corresponding to the orientations are $\mathbf{y}_\Omega = \{\boldsymbol{y}_1, ..., \boldsymbol{y}_N\}$ and $P_\Omega(\beta) = \{P_{\mathbf{R}_1}(\beta), ..., P_{\mathbf{R}_N}(\beta)\}$, respectively. Tomographic reconstruction seeks minimization of a re-projection cost

$$\hat{\beta}(\boldsymbol{x}) = \arg\min_\beta \left\{ \mathcal{D}\left[\mathbf{y}_\Omega, P_\Omega(\beta)\right] + \mathcal{R}(\beta) \right\}, \qquad (3)$$

where $\mathcal{R}$ is a regularization term that expresses prior knowledge about β, while $\mathcal{D}$ is a data (fidelity) term. The particular choice of the functionals $\mathcal{R}$ and $\mathcal{D}$ affects the solution and its computational speed.

2.2. Orientation Estimation

In most tomography applications, the rotation set Ω is known. This is not the case in SPR or our problem domain: specimen appear at arbitrary, unknown orientations in the projection imaging system. Here, 3D reconstruction of β becomes coupled to the estimation of Ω. SPR approaches are roughly divided into two classes of algorithms: reference-free and iterative refinement [11]. Iterative refinement methods use an initial $\{\beta(\boldsymbol{x}), \Omega\}$, then gradually adjust this set of unknowns to fit the data and priors similarly to Eq. (3). This process is computationally complex [8], thus slow.

To succeed in practice, a good initialization is required. Such initialization is provided using reference-free reconstruction, which is dramatically faster: Ω is derived without explicit recovery of $\beta(\boldsymbol{x})$, and afterwards $\beta(\boldsymbol{x})$ is estimated in a single step. Due to the computational advantage, our work focuses on a reference-free approach. The result can subsequently serve as initialization for improvement. The background for reference-free SPR is thus detailed.

Denote a 3D spatial frequency by $\boldsymbol{k}=(k_1, k_2, k_3)$. The 3D Fourier transform of the volumetric object is

$$\mathcal{F}\{\beta\}(\boldsymbol{k}) = \int \beta(\boldsymbol{x})e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{k}}\mathrm{d}\boldsymbol{x}, \qquad (4)$$

where $i=\sqrt{-1}$. A *central slice* is a frequency-plane in the 3D Fourier space, that passes through the origin, $\boldsymbol{k}=0$. For example, the domain of a horizontal frequency-slice is $[k_1, \ k_2, \ 0]^{\mathrm{T}}$, $\forall k_1, k_2 \in \mathbb{R}$. The domain of any central slice can be written in terms of a *rotated* horizontal slice,

$$\mathbf{R}^{-1}[k_1, \ k_2, \ 0]^{\mathrm{T}}. \qquad (5)$$

Figure 3 illustrates two central slices in 3D Fourier space (3D k-space). They intersect at a *common tilt axis*, also termed *common line*.

The 3D Fourier transform of a volumetric object and the 2D Fourier transform of any projected image of this object
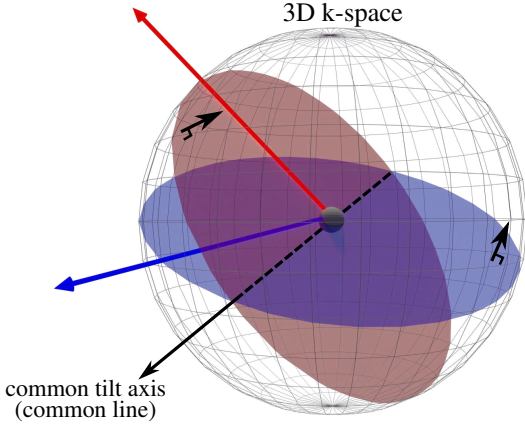
Figure 3. Two central slices in the 3D Fourier domain. Blue represents the coordinate system of $\tilde{\mathbf{y}}_n$. Red represents the coordinate system of $\tilde{\mathbf{y}}_m$. The slices intersect at a *common tilt axis* (common line). It is perpendicular to the direction-vector of either projection.
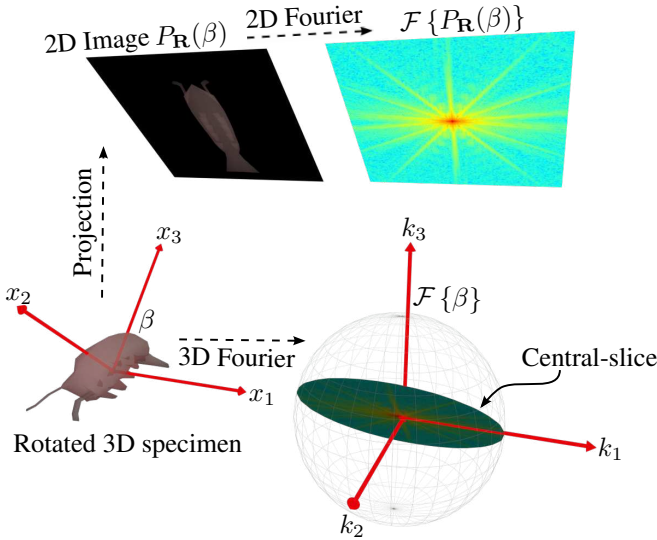

Figure 4. Illustration of two main concepts: projection and the central-slice theorem (Eq. 6). The 2D Fourier transform of any projected image of a rotated object, $\mathcal{F}\{P_{\mathbf{R}}(\beta)\}$, is equivalent to the 3D Fourier of the volumetric object, sampled on a central slice in the frequency domain.

are related by the *Central-Slice Theorem* [41] (Fig. 4).

$$\mathcal{F}\{\beta\}\left(\mathbf{R}^{-1}[k_1,\ k_2,\ 0]^{\mathrm{T}}\right) = \mathcal{F}\{P_{\mathbf{R}}(\beta)\}(k_1,k_2) \quad (6)$$

$$\begin{bmatrix} \text{3D Fourier sampled on a} \\ \text{rotated central slice} \end{bmatrix} = \begin{bmatrix} \text{2D Fourier of a} \\ \text{rotated projection} \end{bmatrix}$$

This theorem is at the core of Fourier-based tomography. It is also a key for orientation estimation, based on the following *common-line theory*.

2D k-space of $\tilde{\mathbf{y}}_n(\rho,\psi)$ 2D k-space of $\tilde{\mathbf{y}}_m(\rho,\psi')$
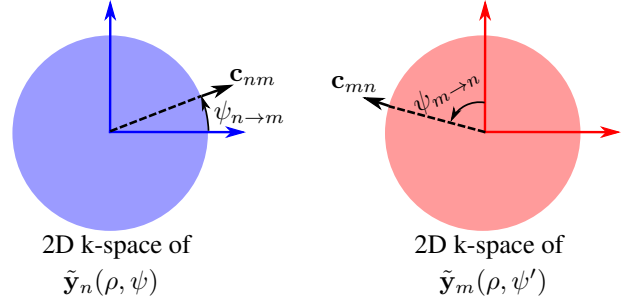
Figure 5. A pair of central slices in the 3D Fourier domain are equivalent to a pair of 2D Fourier images of the projected image data. The common line is realized in the Fourier images by a line at angles $\psi_{n\to m}$ and $\psi_{m\to n}$ according to the coordinate system.

Define polar coordinates

$$x_1 = r\cos\theta \qquad k_1 = \rho\cos\psi \qquad (7)$$
$$x_2 = r\sin\theta \qquad k_2 = \rho\sin\psi. \qquad (8)$$

The 2D continuous polar Fourier transform [13] is

$$\mathcal{F}\{f\}(\rho,\psi) = \int_\pi^\pi \int_0^\infty f(r,\theta)e^{-2\pi i r\rho\cos(\psi-\theta)}r\,\mathrm{d}r\,\mathrm{d}\theta. \quad (9)$$

Let $\mathbf{y}_n, \mathbf{y}_m$ be the n'th and m'th projected measured images. Respectively, denote the polar Fourier transform of each image by $\tilde{\mathbf{y}}_n(\rho,\psi)$ and $\tilde{\mathbf{y}}_m(\rho,\psi')$. According to (6), each empirical Fourier-image $\tilde{\mathbf{y}}_n$ is equivalent to a central slice of the 3D Fourier representation of an object. Recall that a slice-pair intersects in a *common line*. This common line thus appears, theoretically, both in $\tilde{\mathbf{y}}_n(\rho,\psi)$ and $\tilde{\mathbf{y}}_m(\rho,\psi')$, for particular angles $\psi=\psi_{n\to m}$ and $\psi'=\psi_{m\to n}$ (Fig. 5). The angles $\psi_{n\to m}, \psi_{m\to n}$ *partly* indicate the relative orientation of the two projections.

Define 2D unit row vectors (Fig. 3)

$$\mathbf{c}_{nm} = [\cos\psi_{n\to m},\ \sin\psi_{n\to m}], \quad (10)$$
$$\mathbf{c}_{mn} = [\cos\psi_{m\to n},\ \sin\psi_{m\to n}]. \quad (11)$$

Recall from Eq. (1), that the matrices $\mathbf{R}_n, \mathbf{R}_m$ express rotation relative to the microscope reference frame. These are the matrices we seek to recover. They satisfy [52]

$$[\mathbf{c}_{nm},\ 0]^{\mathrm{T}} = \mathbf{R}_n^{-1}\mathbf{R}_m[\mathbf{c}_{mn},\ 0]^{\mathrm{T}}. \quad (12)$$

If the pairwise angles $\psi_{n\to m}, \psi_{m\to n}$ are known (i.e. $\mathbf{c}_{nm}, \mathbf{c}_{mn}$ are set) then Eq. (12) is a constraint on the *relative 3D rotation* between the 3D specimen that had yielded the measured images $\mathbf{y}_n, \mathbf{y}_m$. Note that Eq. (12) is invariant to a global rotation of the microscope frame by matrix $\mathbf{O}$, i.e., $\mathbf{R}_m \to \mathbf{O}\mathbf{R}_m$ and $\mathbf{R}_n \to \mathbf{O}\mathbf{R}_n$. Thus $\mathbf{O}$ is unresolvable. Furthermore, Fourier measurements on the common line are also invariant to any rotation of either Fourier central-slice about the common-tilt axis (see Fig. 3). Hence (12) is not sensitive to a degree of freedom,

which is the relative rotation angle around the common-tilt axis [50]. However, if three images (projections) are used, this degree of freedom is generally resolved [50].

Let there be $N \geq 3$ measured images. Their transformed set of Fourier-images is $\tilde{\mathbf{y}}_n(\rho, \psi) \equiv \mathcal{F}\{\mathbf{y}_n\}$, $n=1,...,N$. Based on the pairwise linear constraints as (12), all rotations are estimated by

$$\hat{\mathbf{R}}_1, ..., \hat{\mathbf{R}}_N = \underset{\mathbf{R}_1,...,\mathbf{R}_N}{\operatorname{argmin}} \sum_{n \neq m} \|\mathbf{R}_n \mathbf{c}_{nm} - \mathbf{R}_m \mathbf{c}_{mn}\|_2^2 \quad (13)$$

$$\text{s.t. } \mathbf{R}_n \mathbf{R}_n^{\mathrm{T}} = \mathbf{I}, \ \det(\mathbf{R}_n) = 1, \ n = 1, ..., N.$$

Here $\mathbf{I}$ is a 3×3 identity matrix. Due to non-convexity of Eq. (13), a solution based on relaxation [52] is employed.

Pairwise Common-line Detection

Recall that Eq. (13) assumes $\psi_{n \rightarrow m}$ is known $\forall n, m$. This means that a common line first has to be identified in the 2D Fourier images taken from the observed data. We describe a common method for common-line detection. In both Fourier images $\tilde{\mathbf{y}}_n, \tilde{\mathbf{y}}_m$, consider pairs of radial lines parameterized by the angles ψ, ψ'. Define a cross-correlation function, based on the inner product between $\tilde{\mathbf{y}}_n(\rho, \psi), \tilde{\mathbf{y}}_m(\rho, \psi')$:

$$\Sigma_{nm}(\psi, \psi') = \int_{-\rho_{\max}}^{\rho_{\max}} \tilde{\mathbf{y}}_n^*(\rho, \psi) \tilde{\mathbf{y}}_m(\rho, \psi') \mathrm{d}\rho. \quad (14)$$

Here $\rho_{\max}$ is the radial frequency cap. Common-line detection can be achieved [52] by maximizing the normalized cross-correlation

$$\begin{bmatrix} \hat{\psi}_{n \rightarrow m} \\ \hat{\psi}_{m \rightarrow n} \end{bmatrix} = \underset{\psi, \psi'}{\operatorname{argmax}} \frac{\Sigma_{nm}(\psi, \psi')}{\sqrt{\Sigma_{nn}(\psi, \psi)} \sqrt{\Sigma_{mm}(\psi', \psi')}}. \quad (15)$$

Equations (13,14,15) derive the set Ω (Eq. 2), *without* explicit tomographic reconstruction of β.

3. Scale and Shift Estimation

Unlike molecules of a particular material, organisms of the same genus usually exhibit great variability in size. This leads to failure of typical object-free approaches, which fundamentally rely on Eqs. (14,15). Applying an SPR [52] approach to synthetic data, without scale estimation, results in a poor recovery due to an inability to detect common lines.

We demonstrate this in simulations using a biologically-realistic spatial distribution model. The model was attained by 3D imaging in the lab of a real organism: a Hydra polyp (Fig. 6). Its length is in the order of a hundred microns. It was acquired with a fluorescence light sheet microscope [31], which fast scans the 3D object. We took the 3D volumetric data and computationally projected it onto 100 orientations, simulating random uniform orientation sampled on the unit sphere. Each projection of size 200×200
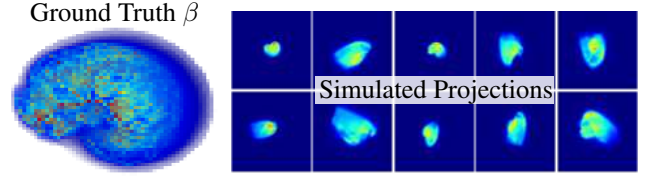

Figure 6. A Hydra polyp volume, observed in 3D, is randomly shifted, scaled and projected onto 100 random orientations, shifts and scales (Eqs. 16, 17). A subset of the projections is displayed.
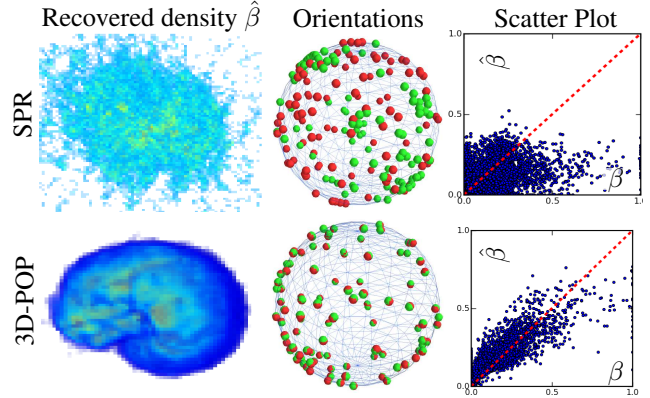

Figure 7. [Top] An SPR recovery [52] without scale estimation. From left to right: A 3D density map with a correlation coefficient of $\sim$0.56 compared to ground-truth volume. For visualization, the recovered density is cut-off at $\beta < 0.1$. Ground-truth (green) and estimated (red) orientations on the unit sphere. A diluted scatter plot (randomized 10% of data). [Bottom] 3D-POP recovery. The correlation coefficient is $\sim$0.94.

pixels is then scaled and shifted uniformly (Fig. 6)

$$\text{shift} \sim \mathcal{U}(-10, 10) \text{ [pixels]}, \quad (16)$$

$$\log_e(\text{scale}) \sim \mathcal{U}(-0.7, 0.7). \quad (17)$$

This sampling results in a scale difference of $\times 4$ between the largest and smallest Hydra length. After drawing samples, the computed numerical mean of $\log(\text{scale})$ was subtracted to give a zero mean distribution.

To further realistically emulate imaging, photon (Poisson) noise statistics are synthetically applied. Here, it is assumed that in the simulated imaging system, each pixel has full-well depth of 10000 photoelectrons. Hence the data has peak signal to noise ratio (SNR) of 100 (40db). A standard SPR recovery fails due to an inability to detect common lines (Fig. 6 [Top]).

For our objective, we must estimate the scale M_n of the object that appears in projection m. Had all observed organisms included a known spherical feature, then each image would have included a projection of the sphere, indicating M_n, $\forall n$. However, in general, an organism has a complex structure and an anisotropic aspect ratio. Hence, we must rely on the multi-view data to jointly derive the scales (and rotations) of all individual organisms, each of which is projected only once. A scale M_n results in an inverted scale in

the Fourier domain. We use this insight to generalize pairwise common-line detection.

Furthermore, organism samples are generally not centered consistently in images that are empirically acquired at random. Therefore, any projection of an object may be shifted by an unknown 3D offset $\boldsymbol{\tau}_n$. An orthographic projection $\mathbf{y}_n$ is invariant to the component of $\boldsymbol{\tau}_n$ along the projection axis. Let us focus on the complementing plane offset in 2D, $\boldsymbol{\tau}_n^{\mathrm{2D}}$. In the 2D Fourier domain an offset creates a linear phase $2\pi[k_1, k_2]^{\mathrm{T}}\boldsymbol{\tau}_n^{\mathrm{2D}}$. This phase, as described below, can be inferred from the data.

Equations (14,15) mean that for a Fourier pair, $\tilde{\mathbf{y}}_n(\rho, \psi_{n\to m}), \tilde{\mathbf{y}}_m(\rho, \psi_{m\to n})$ the common line has matching content on a 1D domain $-\rho_{\max} \leq \rho \leq \rho_{\max}$. Since relevant content is only on the common line (i.e. a function of ρ), then only the linear-phase component of $\boldsymbol{\tau}_n^{\mathrm{2D}}$ along this line affects content for matching. The effect is a linear phase difference, $2\pi\rho\tau$, where τ is a scalar-offset.

Denote a relative scale and offset between $\tilde{\mathbf{y}}_n$ and $\tilde{\mathbf{y}}_m$ as $M_{n\to m}$ and $\tau_{n\to m}$, respectively. Let us generalize Eq. (14) to

$$\Sigma_{nm}(\psi, \psi', M, \tau) =$$
$$\int_{-\rho_{\max}}^{\rho_{\max}} \tilde{\mathbf{y}}_n^*\left(\frac{\rho}{M}, \psi\right) \tilde{\mathbf{y}}_m(\rho, \psi')e^{2\pi i\rho\tau}\mathrm{d}\rho. \quad (18)$$

Eq. (15) is then generalized to

$$\{\hat{\psi}_{n\to m}, \hat{\psi}_{m\to n}, \hat{M}_{n\to m}, \hat{\tau}_{n\to m}\} =$$
$$\underset{\psi,\psi',M,\tau}{\mathrm{argmax}} \frac{\Sigma_{nm}(\psi, \psi', M, \tau)}{\sqrt{\Sigma_{nn}(\psi, \psi, M, \tau)}\sqrt{\Sigma_{mm}(\psi', \psi', M, \tau)}}. \quad (19)$$

The discussion of how we practically solve Eq. (19) per pair m, n is deferred to Sec. 4.

For the moment, assume that the set

$$\mathcal{P} = \{\hat{\psi}_{n\to m}, \hat{\psi}_{m\to n}, \hat{M}_{n\to m}, \hat{\tau}_{n\to m}\}_{\forall n,m} \quad (20)$$

is given. All these pairwise partial constraints of relative rotations, relative scales and relative offsets on common lines, are now integrated into a large optimization problem that seeks all geometric transformations. Fortunately, once the pairwise constraints are given, the optimization problem is *separable*, i.e., scales are recovered separately from offsets, which are recovered separately from orientations. The latter is solved using Eq. (13), based on the angles derived in Eq. (19). We are left with the task of recovering $\{M_n\}_{n=1}^N$ and $\mathcal{T} = \{\boldsymbol{\tau}_n^{\mathrm{2D}}\}_{n=1}^N$.

Denote $\bar{M}_n = \log_{\mathrm{e}}(M_n)$, and $\mathcal{M} = \{\bar{M}_n\}_{n=1}^N$. Then

$$\bar{M}_n - \bar{M}_m = \log_{\mathrm{e}}(\hat{M}_{n\to m}), \quad (21)$$

where $\hat{M}_{n\to m}$ is obtained empirically in the pair-wise optimization (19). Compounding constraints as Eq. (21), $\forall n, m$, creates a linear system of equations. Define a data

term which penalizes for deviations from (21)

$$\mathcal{D}^{\mathrm{scale}}(\mathcal{M}) = \sum_{n\neq m} \left|\bar{M}_n - \bar{M}_m - \log_{\mathrm{e}}(\hat{M}_{n\to m})\right|^2. \quad (22)$$

If M_n is inflated to αM_n, $\forall n$ for a constant α, then the log operation and difference in Eqs. (21,22) null the effect of inflation. Hence, $\mathcal{D}^{\mathrm{scale}}(\mathcal{M})$ is invariant to a global scale. To assist the estimation process lock on a typical scale of unity, define a regularization term

$$\mathcal{R}^{\mathrm{scale}}(\mathcal{M}) = \sum_{n=1}^{N} |\bar{M}_n|^2. \quad (23)$$

Using Eqs. (22,23), the minimization

$$\hat{\mathcal{M}} = \arg\min_{\mathcal{M}} \left[\mathcal{D}^{\mathrm{scale}}(\mathcal{M}) + \mathcal{R}^{\mathrm{scale}}(\mathcal{M})\right] \quad (24)$$

is solved linearly. The scale of specimen n is $\hat{M}_n = \exp\hat{\bar{M}}_n$.

The set $\{\boldsymbol{\tau}_n^{\mathrm{2D}}\}_{n=1}^N$ has $2N$ unknowns. The pairwise constraint $\tau_{n\to m}$ is only indicative of the offset component along the common axis between n, m. Using the row vectors defined in Eq. (10,11), the offset [45] satisfies

$$\hat{\mathbf{c}}_{nm}\boldsymbol{\tau}_n^{\mathrm{2D}} - \hat{\mathbf{c}}_{mn}\boldsymbol{\tau}_m^{\mathrm{2D}} = \hat{\tau}_{n\to m}. \quad (25)$$

There are $\binom{N}{2}$, linear constraints as (25), compounding all pairs. A solution is obtained using $N>5$. A data term penalizes deviation from (25)

$$\mathcal{D}^{\mathrm{shift}}(\boldsymbol{\tau}) = \sum_{n\neq m} \left|\hat{\mathbf{c}}_{nm}^{\mathrm{T}}\boldsymbol{\tau}_n - \hat{\mathbf{c}}_{mn}^{\mathrm{T}}\boldsymbol{\tau}_m - \hat{\tau}_{n\to m}\right|^2. \quad (26)$$

A regularization term favoring minimum average shift is

$$\mathcal{R}^{\mathrm{shift}}(\boldsymbol{\tau}) = \sum_{n=1}^{N} \|\boldsymbol{\tau}_n\|_2^2. \quad (27)$$

Once again, we use regularized least-squares minimization,

$$\hat{\boldsymbol{\tau}} = \min_{\boldsymbol{\tau}} \left[\mathcal{D}^{\mathrm{shift}}(\boldsymbol{\tau}) + \mathcal{R}^{\mathrm{shift}}(\boldsymbol{\tau})\right]. \quad (28)$$

Following efficient optimizations (13,24,28), the offsets, relative scales and rotations of the entire set of specimens are directly estimated. This is done without explicit tomographic recovery of β from unknown specimen parameters. This efficiency relies on the availability of the pairwise transformation parameters set $\mathcal{P}$, as discussed next.

4. Coarse-to-Fine Parameter Estimation

The pairwise transformation parameters $\mathcal{P}$ are the root of the estimation process. They are a result of Eq. (19), per pair. The process we use to reach this optimum is described here. The pairwise parameters interact bidirectionally with global (inter-pair) estimation (13,24,28), as illustrated in Fig. 8. The process has the following four elements.

A. Priors: Let $M_{\max}$ be a maximum scale for a specimen, relative to the population's mean specimen and $\bar{M}_{\max} = \log_{\mathrm{e}}(M_{\max})$. Let $\tau_{\max}$ be a maximum offset of an image relative to the optical axis. The values of $M_{\max}, \tau_{\max}$ are thus parameter priors.

B. Grid search: Let us discretize the parameter domains. The angle ψ is in a domain $[-180°, 180°]$, discretized to N_{angle} values. The angle $\psi' \in [0, 180°]$ is discretized to $N_{\text{angle}}/2$ values. The scale logarithm, $\bar{M}$, is in a domain $[-\bar{M}_{\text{max}}, \bar{M}_{\text{max}}]$, discretized to N_{scale} values. The offset τ is in a domain $[-\tau_{\text{max}}, \tau_{\text{max}}]$, discretized to N_{offset} values. On the entire set of N images, exhaustive search of (19) has computational complexity

$$\mathcal{O}(N^2 N_{\text{angle}}^2 N_{\text{scale}} N_{\text{offest}}). \qquad (29)$$

For most specimen, 3D rotation has a strong impact on the projected appearance. Hence, we prioritize angular precision by a large N_{angle} (we used 180) in omnidirectional breadth. Consequently, to keep Eq. (29) manageable, $N_{\text{scale}}, N_{\text{offest}}$ are small (we used 10 grid samples for each). Following exhaustive search by Eqs. (18,19), a crude set of pairwise parameters is obtained, denoted $\mathcal{P}_0$. It serves as initialization to global estimation by (24,28).

C. Global estimation, image correction: Equations (24,28) are fast to compute. The result is an estimated value $\hat{M}_n, \hat{\tau}_n^{\text{2D}}, \forall n$. Each image $\mathbf{y}_n$ is *corrected* for the scale and offset, being shifted by $\hat{\tau}_n^{\text{2D}}$ and magnified by $1/\hat{M}_n$. The corrected image is denoted $\mathbf{y}_n^{\text{correct}}$.

D. Updated fine grid and images: Global corrections are used for a refined pairwise search:
• Search domains are updated:
$\bar{M}_{\text{max}} \leftarrow \max_n \hat{\bar{M}}_n, \qquad \tau_{\text{max}} \leftarrow \max_n \|\hat{\tau}_n^{\text{2D}}\|_2$.
• Instead of $\{\mathbf{y}_n\}_{n=1}^N$, the corrected images $\{\mathbf{y}_n^{\text{correct}}\}_{n=1}^N$ become the input to pairwise correlations (18,19).

Based on this state, step B. Grid search is run again. This time, the scale and offset domains have become narrower, with corresponding finer grids. Hence, this time, the obtained set of pairwise parameters is refined and denoted $\mathcal{P}_1$. Consequently $\mathcal{P}_1$ is used in step C, which leads to step D, and back to B. After several iterations, the process converges to the final scales, offsets and 3D rotations (see Fig. 8 for illustration). In our experiments, we initialized shifts $\hat{\tau}_n^{\text{2D}}$ as the shift between the image center and center-of-mass. All scales $\hat{M}_n$ were initialized to unity.

Computational complexity: traditional SPR comprises pairwise estimation (complexity c_1) and global optimization (complexity c_2). We estimate random scales by coarse-to-fine pairwise search down to resolution $\delta_{\bar{M}}$. An image ensemble with N images has $\mathcal{O}(N^2)$ pairs. Hence, the pairwise and global complexities are $\mathcal{O}\left[c_1 \log(\bar{M}_{\text{max}}/\delta_M)\right]$ and $\mathcal{O}(c_2 + N^4)$, respectively. In practice, 70^3 voxels took 4 minutes to recover by 100 projections on a computer detailed in Sec. 6 (SPR took 0.7 mins).

5. Pruning Outliers

The analysis thus far assumed that all projected specimens are of the same class, and their relative transformation
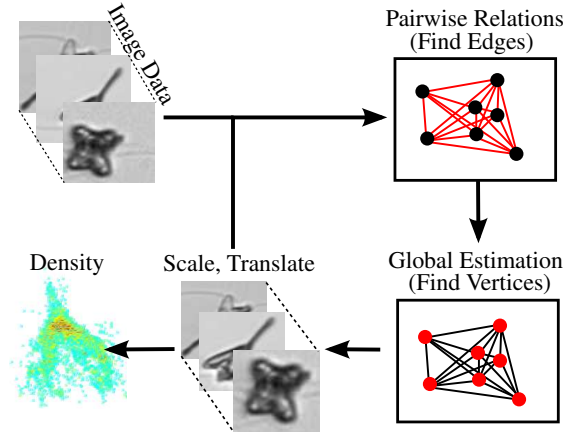

Figure 8. A diagram illustrating the coarse-to-fine parameter estimation. Pairwise interactions are the graph edges and the global parameters are the graph vertices.

is only composed of translation, rotation, and scale (similarity). In practice, the microscope view may be polluted by various specimens belonging to different classes of organisms having widely different morphology. It is undesirable for these outliers to affect the tomographic recovery of β or undermine the estimation $\mathcal{P}$ for the main ensemble class. Even intra-class specimens could have morphological transformations that are non-similar. Images of such specimen are pruned from the data.

There is an indication if at least one image in an image-pair n, m is an outlier. The correlation value

$$\hat{\Sigma}_{nm} = \Sigma_{nm}\left(\hat{\psi}_{n \to m}, \hat{\psi}_{m \to n}, \hat{M}_{n \to m}, \hat{\tau}_{n \to m}\right) \qquad (30)$$

has a low value, despite using the best mutual transformation on a presumed common line. A common line does not exist reliably there. For the entire ensemble, define and affinity matrix $\mathbf{A}$, each of whose element is

$$A_{nm} = \frac{\hat{\Sigma}_{nm}}{\sqrt{\hat{\Sigma}_{nn}}\sqrt{\hat{\Sigma}_{mm}}}. \qquad (31)$$

Note that $\mathbf{A}$ is symmetric with $0 \leq A_{mn} \leq 1$. Using spectral techniques [9], the matrix $\mathbf{A}$ is used to partition the data into clusters. In this work we used a normalized cuts [43] algorithm to prune outliers and extract the main cluster of similar in-class organisms.

6. Test using Experimental Data Rendering

We use our 3D-POP approach, outlined in previous sections, to recover the Hydra polyp (Fig. 6), which was described in Sec. 3. A comparison to SPR recovery is in Fig. 7. Recovered scales are shown in Fig. 9

We quantitively analyze recovery errors (Fig. 6). Rotations are estimated up to a global rotation $\mathbf{O}$, thus the rotational error is defined as

$$\epsilon_{\text{rot}} = \frac{1}{N} \min_{\mathbf{O}} \sum_{n=1}^N \|\mathbf{R}_n - \mathbf{O}\hat{\mathbf{R}}_n\|_2. \qquad (32)$$
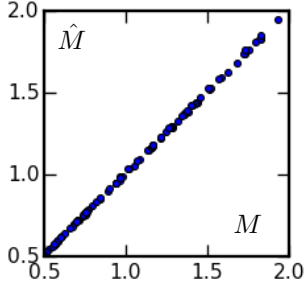
Figure 9. Recovered scales of the Hydra polyp simulation (Fig. 6), compared to the ground truth scales.

The scale and density errors are [3]

$$\epsilon_{\text{scale}} = \frac{1}{N}\|\mathbf{M} - \hat{\mathbf{M}}\|_2, \quad \epsilon_{\text{dens}} = \|\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\|_1/\|\boldsymbol{\beta}\|_1. \quad (33)$$

The resulting errors using 3D-POP were: $\epsilon_{\text{rot}}=2\times10^{-3}$, $\epsilon_{\text{scale}}=3\times10^{-5}$, $\epsilon_{\text{dens}}=0.38$. These errors are much lower than SPR recovery errors, with no scale estimation, $\epsilon_{\text{rot}}=1.96$, $\epsilon_{\text{dens}}=1.52$.

An additional comparison between 3D-POP and SPR used a synthetic random volume that was created by a fluid dynamics model [37]. It is shown in Fig. 10. The volume was projected to 100 random orientations (uniform distribution), each in random scale whose distribution has standard deviation σ_{scale}. In Fig. 10(blue plots), ϵ_{rot} for SPR quickly fails in minor scale variance, while our approach has very small errors, being robust to σ_{scale}. We further assess performance degradation due to non-similarity deformations that compound random scales, translations and orientations. The volume is deformed by $\beta(\boldsymbol{x}) \to \beta(\boldsymbol{x}+\boldsymbol{v}[\boldsymbol{x}])$, where $\boldsymbol{v}=(v_1, v_2, v_3)$ is a random vector field, defined by

$$v_i = \hat{v}_i V L \sin\left[2\pi x_i/(G_i-1) + \phi_i\right], \; i=1,2,3 \quad (34)$$

where $V \sim \mathcal{U}(0, V_{\text{max}})$, $\phi_i \sim \mathcal{U}(-\pi, \pi)$ and $\hat{\mathbf{v}}=(\hat{v}_1, \hat{v}_2, \hat{v}_3)$ is random-sampled on the unit sphere. Here L is the length of the object, in voxels, and $G_1 \times G_2 \times G_3$ is the grid size. In Fig. 10 (red plots), ϵ_{rot} of 3D-POP increases moderately with deformation amplitude V_{max}, and much lower than ϵ_{rot} of SPR in un-deformed (only scaled) objects.

Simulations were run run on a 2.50 GHz Intel Xeon, parallelized on 20 CPU cores. The steps of estimating the rotations, scales and shifts (Sections 3,4) had overall runtime of $\sim$10 minutes. Afterward, the volume density β was recovered in a $150\times150\times150$ voxel domain, at run-time of $\sim$3 minutes, accelerated by an NVIDIA Tesla K40m GPU. For the 3D recovery step we used an *Algebraic Reconstruction Technique* [14,49].

7. Full Microscopy Experimental Recovery

We use 150 images of *Pyramimonas Longicauda* (Fig. 2), a microscopic marine plankton, taken from an annotated dataset [46]. Each image corresponds to a different specimen. These in-situ images $\mathbf{l}_n$ were captured us-
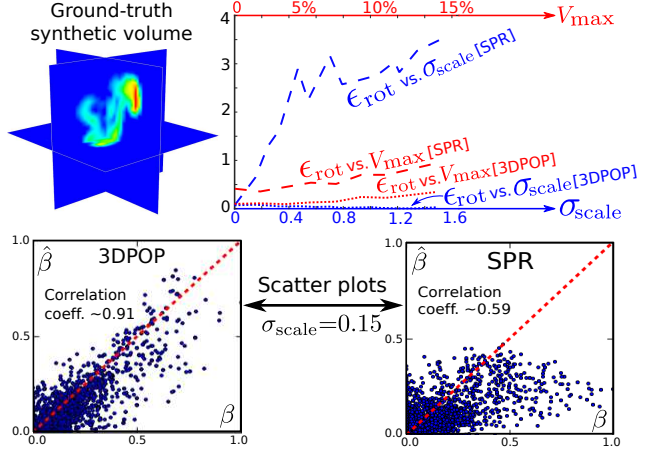

Figure 10. [Top] Ground-truth synthetic volume. ϵ_{rot} as a function of σ_{scale} (red) and V_{max} (blue). The maximal $\sigma_{\text{scale}}=1.6$ corresponds to $M_{\text{max}}=6$, thus, a scale difference of up to $\times36$. [Bottom] Scatter plots for our approach (3D-POP) and SPR.
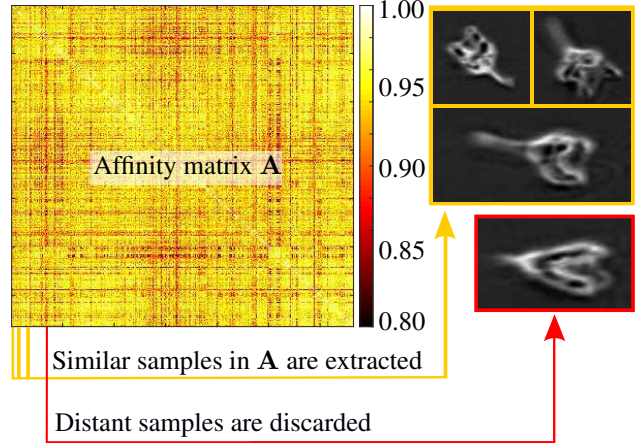

Figure 11. The affinity matrix $\mathbf{A}$ found for the plankton data described in Sec. 7. Note the shades are inverted relative to Fig. 2, due to $-\log_e(\cdot)$. Examples of projections that were found to be similar to extracted cluster are in yellow and an example of a non-similar sample is in red.

ing a submerged instrument [42], in bright-field transmission mode. Hence, each pixel value represents transmissivity. The transmissivity is exponentially related to the optical density. For consistency with the tomographic linear model, the images first undergo a log operation, $\mathbf{y}_n=-\log_e \mathbf{l}_n, \forall n$. Pruning, described in Sec. 5, is illustrated in Fig. 11. It resulted in a cluster of $N=62$ images. The image cluster was used for volumetric recovery. Fig. 12 plots the estimated rotations and volumetric distribution β.

One of the products of the process is the estimated size of each specimen. Normalizing the recovered sizes by a mean cell-size yields a size-distribution of the population. Size distribution parameters are important in research of underlying biological processes [10]. We compute the mean cell
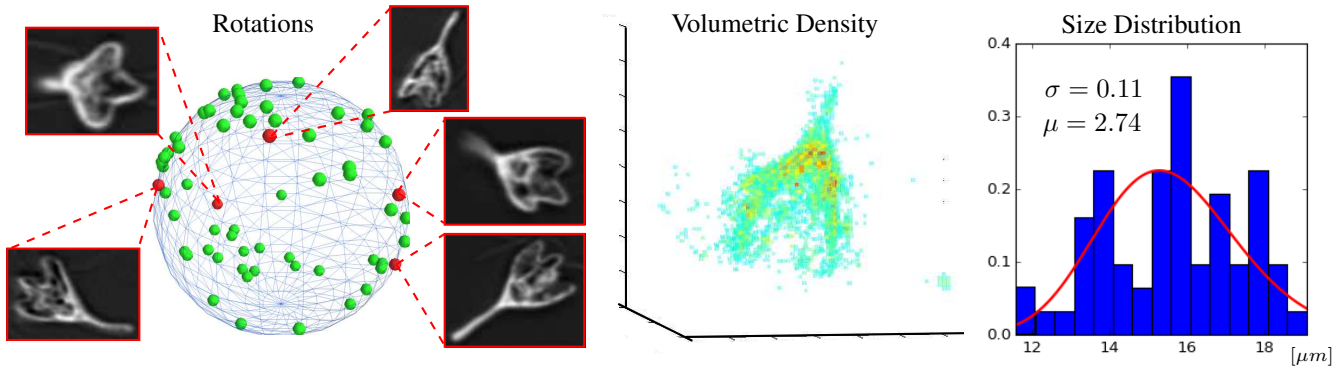
Figure 12. [Left] The recovered orientations and density distribution of the cluster found in the Pyramimonas Longicauda dataset. The corresponding projections of the highlighted (red) orientations are shown. Note the shades are inverted relative to Fig. 2, due to $-\log(\cdot)$. [Right] An estimated normalized size-histogram with a fitted log-normal distribution in red.
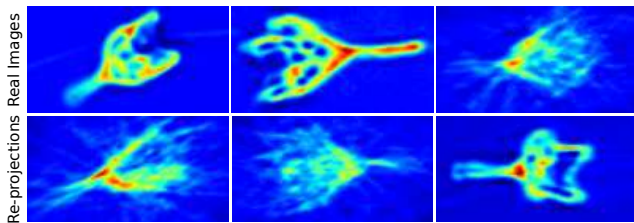

Figure 13. In three different cross-validation tests, a 3D volume is recovered using 61 out of 62 images. The recovered volume is re-projected onto the left-out estimated orientation, to generate a synthetic image (bottom row). Each synthetic image is compared to the left-out input image (top row).
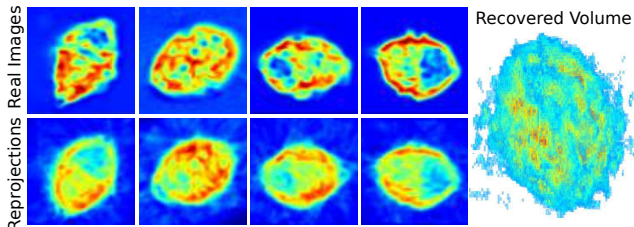

Figure 14. Recovery of a *Heterocapsa triquetra*. [Left] Comparison of real images and re-projections. [Right] Recovered volume.

size as the diameter of the bounding sphere of each specimen. The histogram of the size is shown in Fig. 12, as is the corresponding fit to a log-normal distribution.

In this experiment, we have no ground-truth. Hence we perform cross-validation. A single image is left out of the $N=62$ ensemble, during tomographic recovery of β. The estimated $\hat{\beta}$ is then reprojected to the estimated orientation of the left-out image. Fig. 13 shows re-projections and real images. Fig. 14 shows the re-projections and a recovered 3D density of another specimen from the dataset [46].

8. Discussion

This work derives a framework for a 3D statistical analysis of an ensemble of live organisms, using a single shot per specimen. Existing single-view imaging systems produce data that is utilized for a statistical 3D recovery. A size distribution is obtained, which is of high scientific importance. Our 3D-POP approach relies on Fourier analysis: it works under two conditions: (a) a linear or linearized model (b) line-of-sight integration. The conditions are met in weak refraction and paraxial propagation.

The approach builds upon SPR, with some differences:

I. Cryo-EM suffers from extremely low SNR (each molecule is destroyed by small irradiation).

II. Unlike molecules, organisms within the same class exhibit large variations in 3D structure, particularly scale.

III. Live (unfrozen) specimens are difficult to control.

Currently, our approach does not explicitly recover complex shape variations. However, simulations (Sec. 6) and real data analysis (Sec. 7) indicate recovery robustness. Future work may extend this concept to recovery other morphological variations [26], thus retrieving richer statistics.

Organisms imaged by systems [1,15] in a large volume of water are typically randomly oriented in 3D. However, some systems [42] use fluidic flow which aligns organisms with a high aspect ratio. Rotation about a single axis can reduce the unknown degrees of freedom.

Our analysis could be valuable for taxonomy and classification. Furthermore, in fields where labeling requires experts - unsupervised clustering and size-distribution estimation can improve and expand biological databases.

References

[1] Scripps Plankton Camera. http://spc.ucsd.edu.

[2] S. Agarwal, Y. Furukawa, N. Snavely,I. Simon, B. Curless, S. M. Seitz, R. and Szeliski. Building Rome in a day. *Communications of the ACM*, 54(10), pp.105-112, 2011.

[3] A. Aides, Y. Y. Schechner, V. Holodovsky, M. J. Garay, and A. B. Davis. Multi sky-view 3D aerosol distribution recovery. *Opt. Express*, 21:25820–25833, 2013.

[4] D. Akkaynak, and T. Treibitz. A revised underwater image formation model. *Proc. IEEE CVPR*, 2018.

[5] M. Alterman, Y. Y. Schechner, M. Vo, and S. G. Narasimhan. Passive tomography of turbulence strength. *Proc. ECCV*, pp. 47-60, 2014.

[6] E. Acevedo-Trejos, G. Brandt, J. Bruggeman, and A. Merico. Mechanisms shaping size structure and functional diversity of phytoplankton communities in the ocean. *Scientific Reports*, 5, 2015.

[7] 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. *Proc. ACM TOG*, 27:132, 2008.

[8] M. A. Brubaker, A. Punjani, and D. J. Fleet. Building proteins in a day: Efficient 3D molecular reconstruction. *Proc. IEEE CVPR*, 3099-3108, 2015.

[9] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *PNAS*, 102(21), 7426-7431, 2005.

[10] Z. V. Finkel, J. Beardall, K. J. Flynn, A. Quigg, T. A. V. Rees, and J. A. Raven. Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, 32(1), pp.119-137, 2009.

[11] J. Frank. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*. Oxford University Press, 2006.

[12] J. Frank. Single-particle reconstruction of biological macromolecules in electron microscopy - 30 years. *Quarterly Reviews of Biophysics*, 42(03):139–158, 2009.

[13] J. D. Gaskill. *Linear Systems, Fourier Transforms, and Optics*. John Wiley and Sons, 1978.

[14] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. *Journal of Theoretical Biology*, 29(3):471IN1477–476IN2481, 1970.

[15] G. Gorsky, M. Picheral, and L. Stemmann. Use of the Underwater Video Profiler for the study of aggregate dynamics in the North Mediterranean. *Estuarine, Coastal and Shelf Science*, 50(1), pp.121-128, 2000.

[16] J. Gregson, M. Krimerman, M. B. Hullin, and W. Heidrich. Stochastic tomography and its applications in 3D imaging of mixing fluids. *ACM TOG*, 31(4), 52-1, 2012.

[17] G. C. Hays, A. J. Richardson, and C. Robinson. Climate change and marine plankton. *Trends in Ecology & Evolution*, 20(6):337–344, 2005.

[18] G. T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction From Projections*. Springer Science & Business Media, 2009.

[19] V. Holodovsky, Y. .Y Schechner, A. Levin, A. Levis and A. Aides. In-situ multi-view multi-scattering stochastic tomography. *Proc. IEEE ICCP*, pp.1-12, 2016.

[20] I. Ihrke, K. N. Kutulakos, H. Lensch, M. Magnor, and W. Heidrich. Transparent and specular object reconstruction. *Computer Graphics Forum*, pp. 2400-2426, 2010.

[21] I. Gkioulekas, A. Levin and T. Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. *Proc. ECCV*, 685-701, 2016.

[22] S. Johnsen. *The Optics of Life: A Biologist's Guide to Light in Nature*. Princeton University Press, 2012.

[23] N. Joshi, and M. F. Cohen. Seeing Mt. Rainier: Lucky imaging for multi-image denoising, sharpening, and haze removal. *Proc. IEEE ICCP*, pp. 1-8, 2010.

[24] P. Khungurn, D. Schroeder, S. Zhao, K. Bala and S. Marschner. Matching real fabrics with micro-appearance models. *ACM TOG*, 35(1): 1, 2015.

[25] R. R. Lederman and A. Singer. A representation theory perspective on simultaneous alignment and classification. *arXiv preprint* arXiv:1607.03464, 2016.

[26] R. R. Lederman and A. Singer. Continuously heterogeneous hyper-objects in CryoEM and 3D movies of many temporal dimensions. *arXiv preprint* arXiv:1704.02899, 2017.

[27] S. Lee, K. Kim, A. Mubarok, A. Panduwirawan, K. Lee, S. Lee, H. Park, and Y. Park. High-resolution 3D refractive index tomography and 2D synthetic aperture imaging of live phytoplankton. *J. of the Optical Society of Korea*, 18(6):691–697, 2014.

[28] A. Levis, Y. Y. Schechner, A. Aides and A. B. Davis. Airborne three-dimensional cloud tomography. *Proc. IEEE ICCV*, 2015.

[29] A. Levis, Y. Y. Schechner and A.B. Davis. Multiple-scattering microphysics tomography. *Proc. IEEE CVPR*, 2017

[30] Q. Liu and N. Nguyen. An accurate algorithm for nonuniform fast Fourier transforms (NUFFTs). *IEEE Microwave and Guided Wave Letters*, 8(1):18–20, 1998.

[31] A. Livshits, L. Shani-Zerbib, Y. Maroudas-Sacks, E. Braun, and K. Keren. Structural inheritance of the actin cytoskeletal organization determines the body axis in regenerating Hydra. *Cell Reports*, 18(6):1410–1421, 2017.

[32] S. Lombardi, and K. Nishino. Reflectance and natural illumination from a single image. *Proc. ECCV*, 582-595, 2012.

[33] C. Ma, X. Lin, J. Suo, Q. Dai, and G. Wetzstein. Transparent object reconstruction via coded transport of intensity. *Proc. IEEE CVPR*, pp. 3238-3245, 2014.

[34] S. P. Mallick, S. Agarwal, D. J. Kriegman, S. J. Belongie, B. Carragher, and C. S. Potter. Structure and view estimation for tomographic reconstruction: A Bayesian approach. *Proc. IEEE CVPR*, 2253-2260, 2006.

[35] D. Martinec, and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. *Proc IEEE CVPR*, pp. 1-8, 2007.

[36] S. Matej, J. A. Fessler, and I. G. Kazantsev. Iterative tomographic image reconstruction using Fourier-based forward and back-projectors. *IEEE Trans. on Medical Imaging*, 23(4):401–412, 2004.

[37] G. Matheou and D. Chung. Large-eddy simulation of stratified turbulence. Part II: Application of the stretched-vortex model to the atmospheric boundary layer. *J. Atmos. Sci.*, 71:4439–4460, 2014.

[38] E. D. McCauley. *The Estimation of the Abundance and Biomass of Zooplankton in Samples.* A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters, pp.228-265, 1984.

[39] A. D. Mullen, T. Treibitz, P. L. Roberts, E. L. Kelly, R. Horwitz, J. E. Smith, and J. S. Jaffe. Underwater microscopy for in situ studies of benthic ecosystems. *Nature Communications*, 7, 2016.

[40] A. D. Mullen, T. Treibitz, P. L. Roberts, and J. S. Jaffe. An underwater microscope for in situ imaging of seafloor organism. *OSA Optics in the Life Sciences Congress*, 2017.

[41] F. Natterer. *The Mathematics of Computerized Tomography.* SIAM, 2001.

[42] R. J. Olson and H. M. Sosik. A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging flowcytobot. *Limnology and Oceanography: Methods*, 5(6):195–203, 2007.

[43] J. Shi, and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22(8), pp.888-905, 2000.

[44] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-em by eigenvectors and semidefinite programming. *SIAM Journal on Imaging Sciences*, 4(2):543–572, 2011.

[45] A. Singer and Y. Shkolnisky. Center of mass operators for Cryo-EM theory and implementation. In *Modeling Nanoscale Imaging in Electron Microscopy*, pp 147–177. Springer, 2012.

[46] H. Sosik, E. Peacock, and E. Brownlee. Annotated plankton images data set for developing and evaluating classification methods, 2015. https://darchive.mblwhoilibrary.org/handle/1912/7341

[47] P. P. Srinivasan, R. Ng, and R. Ramamoorthi. Light Field Blind Motion Deblurring. *Proc IEEE CVPR*, 2017.

[48] S. Su, and W. Heidrich. Rolling shutter motion deblurring. *Proc IEEE CVPR*, pp. 1529-1537, 2015.

[49] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers. Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics Express*, 24(22):25129–25147, 2016.

[50] M. Van Heel. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21(2):111–123, 1987.

[51] M. Sheinin, and Y. Y. Schechner. The next best underwater view. *Proc IEEE CVPR*, pp. 3764-3773, 2016.

[52] L. Wang, A. Singer, and Z. Wen. Orientation determination of Cryo-em images using least unsquared deviations. *SIAM Journal on Imaging Sciences*, 6(4):2450–2483, 2013.