Assessment of Imaging Models for Volumetric Tomography of Fluid Flows

Abstract

Volumetric tomography is a powerful tool that employs 2D projections to reconstruct unknown physical quantities in 3D fluid flows. A mathematical model of the imaging process is required to recover the desired volumetric fields. Errors in the imaging model can corrupt the reconstructions, so an accurate model is critical. This work reports the first systematic assessment of nine imaging models in terms of accuracy, computational cost, and range of applicability. A sample matrix method is developed to assess and improve the model's accuracy. A flame chemiluminescence tomography experiment and a synthetic tomographic particle image velocimetry test were conducted. For both luminosity field reconstruction and velocity estimation, the Voxel Spread Function (VSF) model is the most accurate, but the computational cost is tens or hundreds of times higher than other models. The worst models are the Ray-length and VC Direct models. Their error increases when the spatial frequency or seeding density of the object field increases.

Keywords: Volumetric tomography, imaging model, flow field diagnostics

1. Introduction

Spatially-resolved measurements are indispensable for the study of fluid flow. With the continuous development of optical devices, such as camera sensors and flexible fiber bundles, as well as a surge in computational power, optical-based volumetric tomography (VT) has evolved into a mature flow diagnostic that can produce instantaneous and time-resolved three-dimensional (3D) estimates of both scalar and vector fields of key physical quantities [1, 2]. Continuous scalars of interest include the emission source field produced by chemiluminescence [3-6], laser-induced fluorescence [7-11], or laser-induced incandescence [12]. Moreover, a pair of scalar fields recorded in quick succession can be used to estimate velocity fields by cross-correlation [13] or optical flow [14, 15]. Apart from these optical-based VT, there are also other tomographic modalities which need more complicated forward models based on the sensors adopted and the physics of the measurement process, and have large differences compared with optical-based tomography. These cases are out of the scope of this work. The present work only focuses on the optical-based VT problems.

The essence of optical-based VT is to reconstruct a scalar field within the volume of interest (VoI) from one or more two-dimensional (2D) projections of the flow. In VT, the continuous scalar field in the VoI is discretized as a weighted sum of a series of basis functions, $\sum_j^N f_j \varphi_j$, where N is the number of basis functions, and f_j is the coefficient of jth function. The most typical basis functions are cubic voxels within each of which the scalar is assumed to be uniform. Once the basis functions are chosen, the scalar field could be represented by the vector $\mathbf{f} = \{f_j\}$. The 2D projections recorded by image sensors can be represented by a vector $\mathbf{p} = \{p_i\}$. High-fidelity recovery of $\mathbf{f}$ from measured $\mathbf{p}$ requires an *imaging model*, which mimics the way projections are captured by an imaging system (*i.e.*, a camera or set of cameras as well as the required optics), and an *inversion algorithm*, which inverts this model with a set of projection data to estimate the corresponding 3D field. The imaging model describes the response of the imaging system to all the basis functions and is represented by a weight matrix $\mathbf{W}$. Therefore, the projections $\mathbf{p}$ equal $\mathbf{Wf}$ and $\mathbf{f}$ is solved

using the inversion algorithm. An accurate reconstruction of $\mathbf{f}$ largely depends on both the imaging model and the inversion algorithm, since the former determines the correctness of the equation, and the latter controls the accuracy of the solution.

Due to space constraints in actual measurements, there are usually only limited projections. This leads to an underdetermined system of equations. So, previous efforts have been largely devoted to developing an inversion algorithm that solves the 3D field more accurately from a rank-deficient equation system. Commonly used algorithms include the algebraic reconstruction technique (ART) [16, 17], the multiplicative ART (MART) [18, 19], Landweber iteration [20, 21], maximum likelihood expectation maximization (MLEM) [22, 23] and deep learning algorithms [24, 25], among others. Several studies also assessed the proper choice of a reconstruction algorithm [13, 26-28] in some specific applications, including systematic comparisons in the context of tomographic absorption spectroscopy [26] as well as tomographic particle image velocimetry (TPIV) [10, 22, 23].

However, imaging models have obtained far less attention compared with the inversion algorithms. The first major problem faced when investigating imaging models is that few previous studies provide a clear description of the mathematical representation of the imaging model (*i.e.* how to calculate the weight matrix $\mathbf{W}$). One of the earliest clear descriptions of the imaging model in a fluid flow application is from Elsinga *et al.* [13] and was used in TPIV. In this study, a bundle of light rays emitting into the VoI is traced backwards from a pixel in the image plane. The intersection of the ray bundle and the voxel was considered to be the weight factor. However, the accurate calculation of this intersection was quite challenging considering the ray bundle with quadrilateral cross-sections, cubic voxels, and varying intersection angles. Thus, the ray bundles and voxels were approximated as cylinders and spheres, respectively. This approximation leads to a much easier calculation neglecting the influence of intersection angles [29]. The intersection was usually approximated by various fitting functions [30] and was widely adopted in later TPIV studies. The same approximation and weight matrix calculation method has also been used in flame chemiluminescence tomography (FCT). Worth and Dawson [31] employed a Gaussian function to set up the weight matrix and reconstruct the 3D OH* chemiluminescence from two interacting turbulent flames. Wiseman *et al.* [29] took a similar approximation to make the calculation independent of the intersection angles. The above approximation simplifies the calculation, but introduces errors at the same time. To solve this problem, Floyd *et al.* [32] developed a "subvoxel" method to determine the intersection volume between a ray bundle and a cubic voxel. In this method, the voxels that may intersect with the ray bundle from a pixel are divided into subvoxels. The contribution of a specific voxel to the pixel equals the number of subvoxels that intersect with the ray bundle. The above imaging models all attempt to "see" from the image plane and then check the intersections of the ray bundles and voxels in the VoI. A series of more recently developed imaging models trace rays in the forward direction, *e.g.* from the VoI to the image plane. Schanz *et al.* [33] developed several weighting functions to mimic the projection of a single voxel (the so-called Voxel Spread Function (VSF)) in TPIV. Yu, *et al.* [34] introduced a more accurate imaging model to calculate the VSF directly, where numerous point light sources are assumed to be randomly positioned within the voxel and the summation of their projections forms the VSF. Liu *et al.* [35] also adopted this imaging model in TPIV experiments. Ruan *et al.* [36] employed it to reconstruct a turbulent flame in a gas turbine model combustor.

Apart from the sometimes vague description of imaging models, their performance against

each other is also unclear. These imaging models differ in accuracy and computational cost due to their distinct approximation of the true physical imaging process. Due to the reasons like limited computer memory [19], geometrical drifts [37] or setup adjustments, the imaging model should be recalculated for different cases or even different frames of the same target flow. Therefore, cheap rather than expensive reconstruction cost is also an important aspect of a good imaging model. The applicability of these models for different tomographic applications is of great interest, yet there are limited studies on this topic.The applicability of these models for different tomographic applications is of great interest, yet there are limited studies on this topic. One relevant study by Thomas *et al.* [30] qualitatively assessed the imaging models that are only commonly applied in TPIV. However, the aim of TPIV is to extract the velocity field from the discrete scattered light fields at two consecutive time instants. Therefore, the correlation between the two reconstructions will also affect the final velocity estimation. Thus, the applicability of those imaging models to other VT modalities such as FCT, Volumetric Laser-Induced Fluorescence (VLIF), or Volumetric Laser-Induced Incandescence (VLII) should be carefully assessed as the reconstruction accuracy is critical in inferring quantities such as flame curvature [38], temperature [39] and soot volume fraction [40, 41].

To conclude, an accurate imaging model is essential for an accurate VT. However, there is neither a clear summary of these imaging models nor an assessment of their performances. This work aims to address these two issues. First, the work presents a clear definition of the commonly used imaging models and their classification. Next, a sample matrix assessment method is proposed for a quantitative assessment and optimization of these imaging models. Then, the models' influence on the accuracy of reconstructions under different VT applications is discussed in light of experimental and numerical studies. Finally, the conclusions and their potential applications are summarized.

2. Mathematical Background

2.1 Volumetric reconstruction

The mathematical formulation of VT is briefly summarized here for the reader's convenience. To begin, consider a 3D distribution of the quantity of interest (QoI), denoted $f(\mathbf{x}_\mathrm{w})$, where the vector $\mathbf{x}_\mathrm{w} = [x_\mathrm{w}, y_\mathrm{w}, z_\mathrm{w}]^\mathrm{T}$ indicates a location in the VoI. Light emitted within the VoI is captured by the imaging system, and a *projection function* (Ψ) yields the ultimate location on the sensor, $\mathbf{x}_\mathrm{s} = [x_\mathrm{s}, y_\mathrm{s}]^\mathrm{T}$, of a ray of light emitted at $\mathbf{x}_\mathrm{w}$,

$$\mathbf{x}_\mathrm{s} = \Psi(\mathbf{x}_\mathrm{w}). \tag{1}$$

This function is specified through the use of a camera calibration procedure, as discussed in Sect. 2.2. The sensor is usually divided into pixels, and the signal measured at each pixel is a function of the light transmitted to the sensor,

$$R_i \propto \int_{\chi_i} \int_0^\infty I[\Psi^{-1}(\mathbf{x}_\mathrm{s}, l)]\mathrm{d}l\,\mathrm{d}\mathbf{x}_\mathrm{s}. \tag{2}$$

Here, R_i is the signal recorded by the ith pixel, which corresponds to the region χ_i on the sensor; the scalar l indicates a distance along the ray, beginning at the camera and extending out through the VoI; and I is an intensity source term aligned with the *back-projection function*, $\Psi^{-1}(\mathbf{x}_\mathrm{s}, l)$, which is the inverse of Ψ. The back-projection function returns the world point, $\mathbf{x}_\mathrm{w}$, that is a distance l from the camera along the line of sight (LoS) corresponding to a particular sensor location, $\mathbf{x}_\mathrm{s}$. In VT, some part of the light at $\mathbf{x}_\mathrm{w}$ is related to the QoI (*e.g.*, a local intensity source

term that corresponds to chemiluminescence or scattered light) and a post-processed signal, p_i, is expressed as a function of the QoI along each LoS,

$$p_i = \iint_{\chi_i} \int_{l_{\text{in}}}^{l_{\text{out}}} f\,[\Psi^{-1}(\mathbf{x}_{\text{s}}, l)]\,\mathrm{d}l\,\mathrm{d}\mathbf{x}_{\text{s}}, \tag{3}$$

where l_{in} and l_{out} are the distance along a LoS at which a ray enters and exits the VoI, respectively. The goal of VT is to recover f from a set of projection data, which is called tomographic reconstruction.

In VT, the QoI is typically represented via the finite element method; the basis, $\Phi = \{\varphi_j\}$, comprises a set of basis functions, $\varphi_j(\mathbf{x}_{\text{w}})$ (subscript j denoting the jth basis function), which collectively span the VoI and are usually normalized. The most common discretization scheme is the voxel basis, wherein the domain is divided into N small cubic regions (voxels) within which f is assumed to be constant. For instance, the function φ_j is unity inside the jth voxel and zero outside. Using this representation, Eq. (3) can be approximated by a summation,

$$p_i \approx \sum_{j=1}^{N} W_{i,j} f_j, \tag{4}$$

where f_j is the value of f in the jth voxel and the weight:

$$W_{i,j} = \iint_{\chi_i} \int_{l_{\text{in}}}^{l_{\text{out}}} \varphi_j[\Psi^{-1}(\mathbf{x}_{\text{s}}, l)]\,\mathrm{d}l\,\mathrm{d}\mathbf{x}_{\text{s}}, \tag{5}$$

is the contribution of f_j to the signal at the ith pixel, p_i. This approximation can be expressed as a matrix system,

$$\mathbf{p} = \mathbf{W}\mathbf{f}, \tag{6}$$

where $\mathbf{p} = \{p_i\}$ is an $M \times 1$ vector that contains the projection data, $\mathbf{f} = \{f_j\}$ is an $N \times 1$ vector that describes the value of f throughout the domain in terms of the basis Φ, and $\mathbf{W}$ is a $M \times N$ matrix, called the weight matrix, sensitivity matrix, or ray-sum matrix, which relates a distribution of the QoI to the corresponding set of projection data. In other words, $\mathbf{W}$ is the imaging model, which is specified by one's choice of basis and numerical approximation of Eq. (5), and reconstruction consists of estimating $\mathbf{f}$ given $\mathbf{W}$ and $\mathbf{p}$.

The accuracy of reconstructions is governed by

1. the ability of the basis to adequately represent f, *i.e.*, the validity of the approximation in Eq. (4);
2. the fidelity of the numerical implementation of Eq. (5) used to construct $\mathbf{W}$; and
3. the degree to which the reconstruction algorithm can recover $\mathbf{f}$ from $\mathbf{p}$, which depends on the arrangement of cameras as well as the choice of algorithm.

Imaging models differ with respect to the implementation of Eq. (5), and this work assesses the effect of imaging models on reconstructions in VT.

2.2 Pinhole camera model with polynomial distortion

Before establishing an imaging model, it is necessary to specify a projection function, Ψ. In this function, a point in the VoI, $\mathbf{x}_{\text{w}}$, is rotated and translated to obtain a point in the camera's local coordinate system, $\mathbf{x}_{\text{c}} = [x_{\text{c}}, y_{\text{c}}, z_{\text{c}}]^{\text{T}}$,

$$\begin{bmatrix} x_{\text{c}} \\ y_{\text{c}} \\ z_{\text{c}} \end{bmatrix} = \mathbf{R}_{\text{c}} \begin{bmatrix} x_{\text{w}} \\ y_{\text{w}} \\ z_{\text{w}} \end{bmatrix} + \mathbf{t}_{\text{c}}, \tag{7}$$

where $\mathbf{R}_{\text{c}}$ and $\mathbf{t}_{\text{c}} = -\mathbf{R}_{\text{c}}\mathbf{c}$ are the camera's rotation matrix and translation vector and $\mathbf{c}$ is the location of the camera in world coordinates. The normalized pinhole projection indicates the non-

dimensional location of $\mathbf{x}_\mathrm{c}$ on the sensor,

$$\begin{bmatrix} x_\mathrm{n} \\ y_\mathrm{n} \end{bmatrix} = \begin{bmatrix} x_\mathrm{c}/z_\mathrm{c} \\ y_\mathrm{c}/z_\mathrm{c} \end{bmatrix}. \tag{8}$$

Radial lens distortions are usually modeled with a polynomial function, the corrected location of the projection $\mathbf{x}_\mathrm{d} = [x_\mathrm{d}, y_\mathrm{d}]^\mathrm{T}$ is denoted as:

$$\begin{bmatrix} x_\mathrm{d} \\ y_\mathrm{d} \end{bmatrix} = \left(1 + k_{\mathrm{r},1} r_\mathrm{n}^2 + k_{\mathrm{r},2} r_\mathrm{n}^4 + k_{\mathrm{r},3} r_\mathrm{n}^6\right) \begin{bmatrix} x_\mathrm{n} \\ y_\mathrm{n} \end{bmatrix} + \begin{bmatrix} d_\mathrm{x} \\ d_\mathrm{y} \end{bmatrix}, \tag{9}$$

where $r_\mathrm{n}^2 = x_\mathrm{n}^2 + y_\mathrm{n}^2$; $k_{\mathrm{r},1}$, $k_{\mathrm{r},2}$, and $k_{\mathrm{r},3}$ are radial distortion coefficients (the order of this polynomial can be expanded or reduced as needed); and d_x and d_y are tangential distortions. Tangential distortions are given by another polynomial function,

$$\begin{bmatrix} d_\mathrm{x} \\ d_\mathrm{y} \end{bmatrix} = \begin{bmatrix} 2k_{\mathrm{d},1} x_\mathrm{n} y_\mathrm{n} + k_{\mathrm{d},2}(r_\mathrm{n}^2 + 2x_\mathrm{n}^2) \\ k_{\mathrm{d},1}(r_\mathrm{n}^2 + 2y_\mathrm{n}^2) + 2k_{\mathrm{d},2} x_\mathrm{n} y_\mathrm{n} \end{bmatrix}, \tag{10}$$

where $k_{\mathrm{d},1}$ and $k_{\mathrm{d},2}$ are the tangential distortion coefficients. Finally, the ultimate location $\mathbf{x}_\mathrm{s}$ of light on the sensor is obtained by a matrix transformation,

$$\begin{bmatrix} \mathbf{x}_\mathrm{s} \\ 1 \end{bmatrix} = \begin{bmatrix} x_\mathrm{s} \\ y_\mathrm{s} \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_\mathrm{d} \\ y_\mathrm{d} \\ 1 \end{bmatrix}, \tag{11}$$

where $\mathbf{K}$ is the camera's intrinsic matrix,

$$\mathbf{K} = \begin{bmatrix} F_\mathrm{x} & s_\mathrm{sk} & x_0 \\ 0 & F_\mathrm{y} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

In this matrix, F_x and F_y are given by the camera's focal length divided by the width and height of a pixel, respectively; (x_0, y_0) is the camera's principal point; and s_sk is a skew parameter.

A camera calibration procedure is required to determine the unknown parameters in Ψ. Pictures of a calibration target are taken by each camera, and a calibration algorithm is used to estimate $\mathbf{R}_\mathrm{c}$, $\mathbf{t}_\mathrm{c}$, $\mathbf{K}$, and the distortion parameters. We used Zhang's [42] method of camera calibration, as implemented by Bouget [43].

Many imaging models require the back-projection function, Ψ^{-1}, which is obtained by inverting Ψ. Inversion of the pinhole model is conducted in three steps. First, Eq. (11) is inverted to calculate the normalized location of each pixel (*i.e.*, following distortion in the forward procedure),

$$\begin{bmatrix} x_\mathrm{d} \\ y_\mathrm{d} \\ 1 \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} x_\mathrm{s} \\ y_\mathrm{s} \\ 1 \end{bmatrix}. \tag{13}$$

Second, the undistorted positions must be calculated by inverting the polynomial expressions in Eqs. (9) and (10). We use a gradient-based nonlinear solver to obtain $\mathbf{x}_\mathrm{n}$, starting from an initial guess of $\mathbf{x}_\mathrm{d}$; convergence is normally achieved within 20 iterations. Third, the LoS vector originating at $\mathbf{x}_\mathrm{n}$ is determined by inverting the transformation in Eq. (7),

$$\mathbf{v} = \mathbf{R}^\mathrm{T} \begin{bmatrix} x_\mathrm{n} \\ y_\mathrm{n} \\ 1 \end{bmatrix} + \mathbf{c}. \tag{14}$$

The back projection function is simply

$$\Psi^{-1}(\mathbf{x}_s, l) = l\frac{\mathbf{v}}{\|\mathbf{v}\|}, \tag{15}$$

where $\mathbf{v}$ depends on $\mathbf{x}_s$.

3. Analysis of Imaging Models

3.1 Taxonomy

Since the object under investigation is the emission or scattering field in a flame or a flow fluid, the scope of the imaging models discussed in this work includes those used in the VT modalities such as FCT and TPIV. The target field of other VT methods the Background Oriented Schlieren technique for example is not a luminosity field, so specific imaging models for this technique are not included here. Also, as mentioned in the introduction, many studies only provide quite vague descriptions of the imaging models. Thus, we only summarize the imaging models which are clearly described in the literature. The most fundamental physics behind all the models discussed is that light travels in straight lines. Depending on the direction in which the light rays are traced, imaging models for VT can be divided into two categories: pixel-centric (PC) models and voxel-centric (VC) models. According to this physical principle, it is obvious that if we trace the light ray from the pixel to the reconstruction volume, the voxels intersecting with the ray contribute to the signal of this specific pixel. How much each intersecting voxel contributes is approximated differently within the pixel-centric models. The common principle for all the approximations is that the shorter the distance between the voxel and the central axis of the ray bundle is, the smaller intersecting volume will be. As for the VC models, it assumes the light intensity is uniform within the voxel, they trace lights emitting in the voxel to formulate a reasonable voxel spread function i.e. the intensity distribution on the imaging plane. The detailed information of all the models we conclude from the literature are introduced in the following text.

3.1.1 PC models

In a PC model, the bundle of rays that reaches a pixel is traced backwards through the VoI using Ψ^{-1}, and the volumetric intersection of this bundle with each voxel is computed to build $\mathbf{W}$. These intersection volumes represent the fraction of light emitted from a voxel that reaches a given pixel. Two crucial tasks arise when constructing the weight matrix. The first task is to identify the set of voxels transected by each ray bundle, called the ray's *weighting scope*. The second task is to calculate $W_{i,j}$ for all these intersections (*i.e.*, approximating Eq. (5)). Elements of $\mathbf{W}$ that do not correspond to an intersection are zero. We discuss four PC models, visualized in Fig. 1:

1. Ray-length [44],
2. PC Linear [45],
3. PC Gaussian [13], and
4. Subvoxel [32] models.

In the Ray-length model [44] (Fig. 1(a)), the ray bundle that reaches a pixel is approximated by the pixel's primary ray, *i.e.*, the ray that intersects the center of the pixel. Intersections are computed using this 1D line ("thin ray") such that $W_{i,j}$ is the chord length of the ith LoS through the jth voxel. This model can be implemented with an efficient, analytical ray-box intersection algorithm [46]

In the PC Linear model [45] (Fig. 1(b)), voxels are represented using spheres, and the ray bundles are represented using cylinders to simplify the calculation of the intersections. The volume of the voxels and the area of the pixels are kept constant by using a voxel radius of $(3/(4\pi))^{1/3}l_{\mathrm{vx}}$,

where l_{vx} is the width of a voxel, and pixel radius of $(1/\pi)^{1/2} l_{\mathrm{px}}$, where l_{px} is the width of a pixel. Voxel-cylinder intersections occur when the minimum distance from the central axis of a cylinder to a voxel center, d_{v}, satisfies

$$d_{\mathrm{v}} < \left(\frac{3}{4\pi}\right)^{\frac{1}{3}} l_{\mathrm{vx}} + M \left(\frac{1}{\pi}\right)^{\frac{1}{2}} l_{\mathrm{px}}, \tag{16}$$

as shown in Fig. 1(b). Here, *M* represents the pixel-to-voxel ratio. Weights are determined using a linear ramp that peaks at the cylinder's central axis (*i.e.*, the LoS corresponding to the Ray-length model).

$$W_{i,j} = 1 - k * \frac{d_{\mathrm{v},i,j}}{l_{\mathrm{vx}}}, \tag{17}$$

where $d_{\mathrm{v},i,j}$ is the minimum distance between the ith cylinder axis and the center of the jth voxel. k is an empirically defined value based on the pixel-to-voxel ratio. These weights are normalized by l_{vx} for numerical stability. However, it should be noted that doing so means that reconstructed vectors $\mathbf{f}$ are scaled by the same factor.

The PC Gaussian model [13] (Fig. 1(b)) is akin to the PC Linear model except that the weight factor tapers off according to a Gaussian distribution instead of a linear ramp. Voxel-cylinder intersections are computed by Eq. (16), as before, but the intersection volume is approximated in terms of a Gaussian function:

$$W_{i,j} = -\exp\left(-\frac{d_{\mathrm{v},i,j}}{2l_{\mathrm{vx}}^2 \sigma^2}\right), \tag{18}$$

where the standard deviation σ is an empirically defined value.

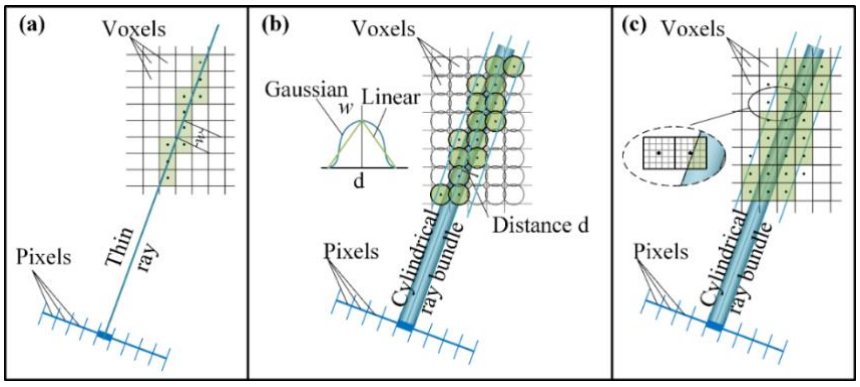

Fig. 1. Schematic of PC imaging models. (a) Weights in the Ray-length model are the chord length of a thin line, corresponding to the center of a pixel, through a voxel. (b) Linear and Gaussian weights represent the intersection of a cylindrical ray bundle with a spherical voxel. (c) Subvoxel weights represent the ratio of "subvoxels" transected by a cylindrical ray bundle to the total number of subvoxels in a given voxel.

Lastly, in the subvoxel model [32] (Fig. 1(c)), the ray bundle is represented by a cylinder as in the previous two models, but a new equation:

$$d_{\mathrm{v}} < l_{\mathrm{vx}} + M \left(\frac{1}{\pi}\right)^{\frac{1}{2}} l_{\mathrm{px}}, \tag{19}$$

is used to identify voxels which may intersect with the cylinder. For these voxels, a higher resolution sub-grid is defined, where each subvoxel is regarded as a point and is assessed using the strict intersection condition $d_V < M(1/\pi)^{1/2} l_{\mathrm{px}}$. The sum of the intersected subvoxels is then used to approximate the intersection volume between the voxel and the cylindrical ray bundle. Floyd [32] found that using a 10×10×10 grid of subvoxels produced a good approximation to the intersection

volume.

3.1.2 VC models

VC imaging models trace rays emitted from the VoI to the cameras using their projection function, as depicted in Fig. 2(a). Similar to the PC imaging models, models also differ in two aspects: the *weighting scope*, *i.e.*, the set of pixels that a voxel will influence, and the function used for weight calculation. We tested five VC models:

1. The voxel spread function (VSF) [34],
2. VC Direct [30],
3. VC Gaussian [33],
4. VC Bilinear [33], and
5. Disc-Intersection [44] models.

VSFs [34] (Fig. 2(b)) represent the response of an imaging system to uniform radiation from a voxel, similar to a point spread function. Light emitted at a point emanates outward, spherically, and the cone of light that transects the lens is projected onto the camera sensor, as shown in Fig. 2(b). The light emitted at some point $\mathbf{x}_\mathrm{w}$ in the VoI forms a point at $\mathbf{x}_\mathrm{s} = \Psi(\mathbf{x}_\mathrm{w})$, as discussed above. Analytically determining the sensor's response to a volume of light is not feasible, but this response can be approximated by Monte Carlo simulation, which is done by repeatedly sampling points in a voxel to generate a cluster of points on each of the camera sensors. Positions in the voxel are sampled from a uniform distribution, and the weight of each pixel for a given voxel ($W_{i,j}$) is simply the ratio of points in that pixel to the total number of samples.

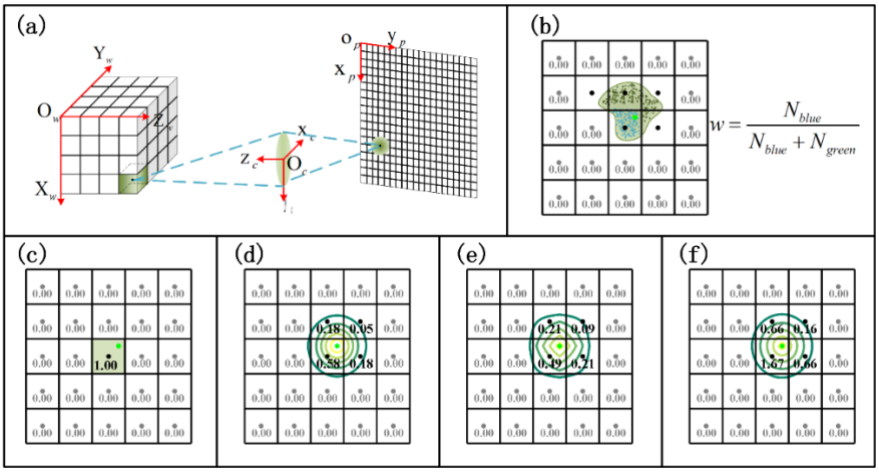

Fig. 2. VC imaging models are based on (a) the response of a camera to uniform radiation from a voxel. Subfigures (b)–(f) depict VC models sampled on a 5×5 pixel grid, where the green dots indicate the sensor location corresponding to light emitted at the voxel center. VC models include the (b) VSF model, (c) VC direct model, (d) VC Gaussian model, (e) VC Bilinear model, and (f) Disc-intersection model.

Unlike the VSF technique, the remaining VC models utilize a single-point source, positioned at the center of a voxel, to approximate the sensor's response to volumetric radiation from a voxel.

In the VC direct model [30] (Fig. 2(c)), the weight of the pixel that contains $\Psi(\mathbf{x}_\mathrm{w})$ is set to unity and all other pixels have zero weight for that voxel.

The VC Gaussian model [33] (Fig. 2(d)) assigns weights using a Gaussian decay that is parameterized by the distance from pixel centroids to the sensor location of a point source positioned at the center of a voxel. The distance from the ith pixel centroid to the sensor location reached by the jth voxel center is denoted $d_{\mathrm{s},i,j}$. According to [44], when $l_\mathrm{px} \approx l_\mathrm{vx}$, the weights are modeled

as

$$W_{i,j} = \exp\left(-d_{\mathrm{s},i,j}^2 log20\right). \tag{20}$$

The weighting scope can be determined by directly defining an upper limit for $d_{\mathrm{s},i,j}$ according to the pixel-to-voxel ratio or a low limit for $W_{i,j}$, such as 0.01 as a cut-off threshold.

The VC Bilinear model [33] (Fig. 2(e)) is similar to the VC Gaussian model except that the weights are approximated as a bilinear function of the x_{s}- and y_{s}-components of $d_{\mathrm{s},i,j}$: $d_{\mathrm{x},i,j}$ and $d_{\mathrm{y},i,j}$, respectively. Bilinear weights are given by

$$W_{i,j} = \left(1 - k\frac{d_{\mathrm{x},i,j}}{l_{\mathrm{px}}}\right)\left(1 - k\frac{d_{\mathrm{y},i,j}}{l_{\mathrm{px}}}\right), \tag{21}$$

where k is a coefficient that controls the decay as well as the weighting scope. Eq. (21) also presumes that $l_{\mathrm{px}} \approx l_{\mathrm{vx}}$.

In the Disc-Intersection model [44] (Fig. 2(f)), weights correspond to the intersection area of radiation from a voxel projected onto the sensor with a pixel. The projected radiation and pixels are approximated as circles in [44] and, assuming a pixel-to-voxel ratio of one, the resulting weights are

$$W_{i,j} = 2\,\mathrm{acos}\left(\frac{d_{\mathrm{s},i,j}}{d_{\mathrm{disc}}}\right) - 2\frac{d_{\mathrm{p},i,j}}{d_{\mathrm{disc}}}\left[1 - \left(\frac{d_{\mathrm{s},i,j}}{d_{\mathrm{disc}}}\right)^2\right]^{1/2}, \tag{22}$$

where $d_{\mathrm{disc}} = 2(\pi)^{-1/2} l_{\mathrm{px}}$ is the diameter of the disc.

For the convenience of comparison and to fully-utilize information recorded by each pixel, we set the pixel-to-voxel ratio to 1 in our analysis.

3.2 Imaging model assessment method

Since the true distribution of QoI is unknown, the best imaging model cannot be identified through a comparison between the reconstructions and the ground truth. As the mathematical approximation of an imaging process is represented by the weight matrix, a reasonable way to assess these models is to compare their corresponding weight matrices. How well they mimic the true imaging process determines how accurate they are. Thus, a weight matrix made up of weights calculated from Eq. (5) is an ideal benchmark. It is non-trivial to calculate the whole weight matrix considering the varying cross-section of the ray bundle as well as the intersection angle. But if we only focus on a few voxels (*i.e.*, several columns in **W**), the computational cost is affordable. So, a sample matrix is developed, as shown in Fig. 3. A set of uniformly distributed voxels are sampled from the whole VoI, and their corresponding columns form the sample matrix. The benchmark matrix is the sample matrix calculated from Eq. (5), and therefore is a good reference case. The weights (*i.e.*, intersections) are calculated by Monte Carlo simulation [47]. A million random points are uniformly generated in a voxel and assessed one by one to see whether a point is inside the ray bundle. The intersection equals the number of points inside the ray bundle divided by the total number of points.

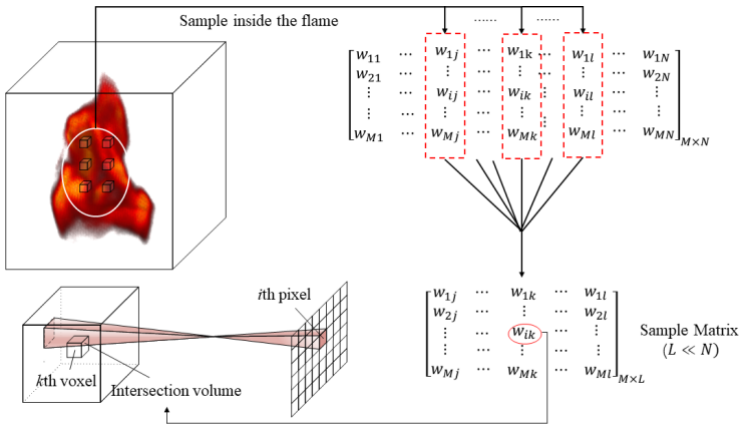

Fig. 3. The process of building the benchmark matrix.

To evaluate the sample matrices from different imaging models, several assessment indices are proposed. The first index is the similarity of the columns in the sample matrix (SoSM for short). Each column represents the VSF of a voxel. The right hand side of Fig. 4 is an example of the VSFs of a voxel calculated from the benchmark and a certain imaging model (only pixels in the weighting scope are shown) respectively. The similarity between these two VSFs is represented by their correlation coefficient. The SoSM equals the mean correlation coefficient of all columns in the sample matrix. The second index is the error of the sample matrix (EoSM for short), which is defined as the mean error of all weights in the sample matrix, calculated as:

$$\mathrm{EoSM}_{im} = \sum_{M\times L}\left|\mathrm{w}_{ij}^{im} - \mathrm{w}_{ij}^{b}\right|/(M \times L), \tag{23}$$

where w_{ij}^{im} and w_{ij}^{b} denote the weight calculated from a particular imaging model and that from the benchmark matrix respectively.

The third index is the standard deviation of the voxel volume (SDoVV for short). In each view, a voxel intersects with a set of ray bundles from different pixels, and these ray bundles do not intersect with each other, so the sum of these intersections equals the volume of a voxel. For ease of comparison between different imaging models, the voxel volume in each model is normalized to one. Thus, $\sum_{j\in view_k} w_{ij} = 1$ is an implicit requirement for a correct imaging model. Here, subscript *i* and *j* specify the voxel and the pixel respectively, and $view_k$ indicates the *k*th view. To evaluate how well an imaging model meets this requirement, the SDoVV is defined as the standard deviation of each voxel volume in each view.

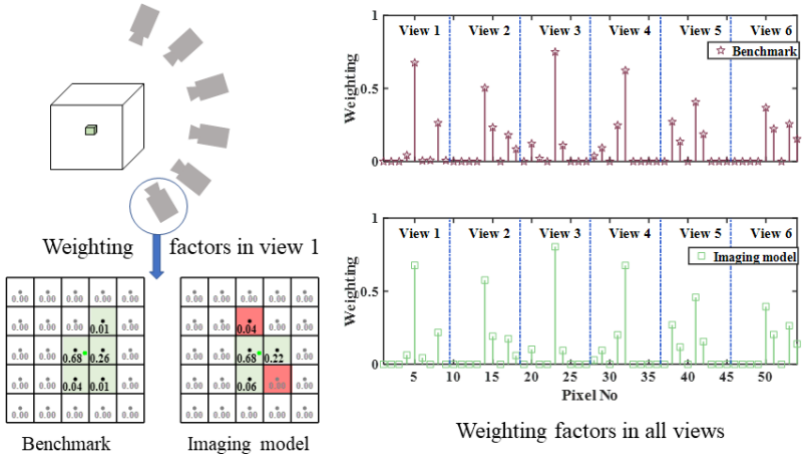

Fig. 4. Illustration of the SoSM and EoSM indices. The projection of a voxel is recorded from several views. In each

view, the voxel corresponds to a few pixels and each voxel-pixel pair owns a weight. The weights in all the views form a column in the sample matrix. The weights calculated from an imaging model are different from that of the benchmark matrix. The SoSM refers to the mean correlation coefficient of the columns calculated from the benchmark and an imaging model. The EoSM refers to the mean errors of all the weights in the sample matrix.

4. Comparison of imaging models in a FCT system

4.1 Experimental and computational setup

Section 3.2 introduces a sample matrix to evaluate the accuracy of different models. To prove its effectiveness, a FCT experimental system was set up. The system mainly consisted of a high-speed camera (Photron FASTCAM Mini AX100) and a customized fiber bundle (Fig. 5). The flame in the experiment was located at the center of an approximate arc formed by six lenses (f=24mm, f/2.8), and the viewing angles are listed in Table 1. No filter was used. Radiation from a swirling diffusion flame was collected and transferred to the output of a fiber bundle, and a high-speed camera then recorded the flame projections from all viewing angles in a single image. The details of the burner used can be found in [48].

Table 1 Viewing angles for all six lenses

View#	Yaw(°)	Pitch(°)	Roll(°)
1	22.89	8.98	90.32
2	54.09	-1.72	87.88
3	100.25	3.49	91.24
4	128.31	23.43	90.98
5	142.07	18.91	89.65
6	178.64	-0.15	93.73

The flame was contained within a $60 \times 20 \times 20\ mm^3$ reconstruction volume, which was discretized into $216 \times 72 \times 72$ voxels. Each voxel has a size of $0.27\ mm$, which is fine enough for the flame in our experiment. The pixel resolution of the camera is 1024×1024 and each view covers an area of 230×100 pixels on the camera. The pixel size of the camera is $20\ \mu m$. The six projections after background removal are shown in Fig. 6. At the selected time instant the flame structure is shaped as a continuous smooth-contoured arc, which decreases in thickness with downstream distance. A second lower intensity curved flame branch can be seen behind the main branch in views 1, 2, and 4. The ART algorithm with a relaxation factor of 1.0 was used for reconstruction. Two termination criteria were adopted in this work: (1) the relative change of the solutions calculated from two adjacent iterations is smaller than 10^{-5}; and (2) the number of iterations is larger than 50. The reconstruction program was run on a workstation equipped with 128 GB of memory and two E5-2670 CPUs.

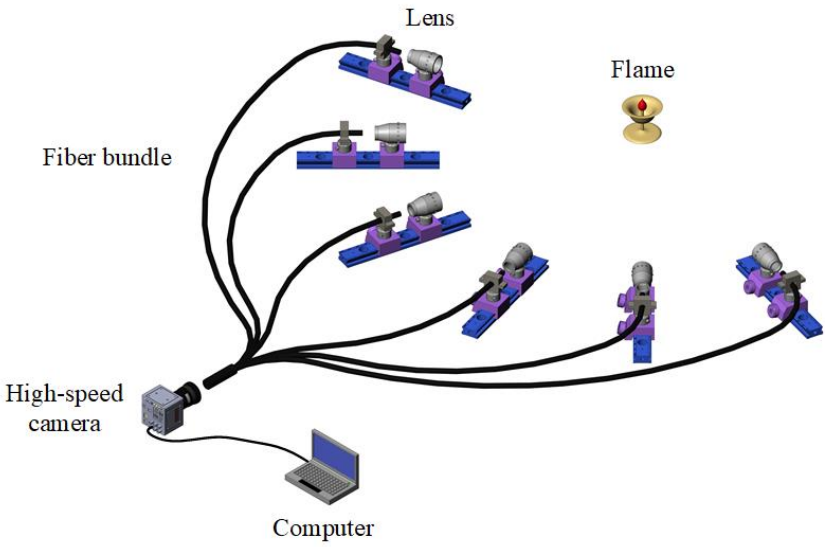

Fig. 5. Illustration of the FCT experiment setup.

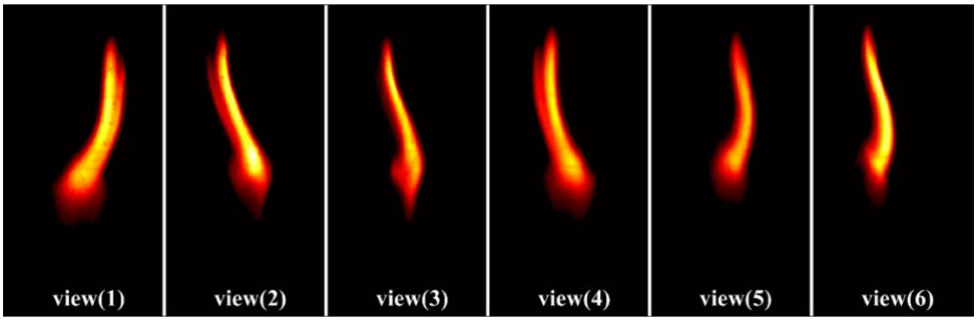

Fig. 6. Flame projection images after background removal.

4.2 Application of sample matrices in the evaluation of experimental results

Since the true chemiluminescence distribution is unknown, the comparison starts from the weight matrices of different imaging models. The size of the weight matrix is 138000×1119744, considering $216 \times 72 \times 72$ voxels and six 230×100 pixel projections. A total of 27 voxels were selected to form the sample matrix. Then all the imaging models introduced in Section 3.1 were compared against the benchmark matrix in terms of the assessment indices introduced in Section 3.2. The results are listed in Table 2. The first two indices, *i.e.*, SoSM and EoSM can measure to what degree the imaging model can approximate the real imaging process. The closer the SoSM of an imaging model is to one (or the closer the EoSM of an imaging model is to zero), the more similar a sample matrix calculated by this imaging mode is to the benchmark matrix. The VSF model has the best performance in terms of both indices. The VSF model traces all light emitted from a voxel to the image plane. Since no approximation is adopted in the tracing process, the intersections calculated by this model have nearly no differences with the benchmark. The subvoxel model has the second smallest EoSM, for only this model considers the variation of intersection angles when calculating the intersection volume. The SoSM and EoSM of the VC Bilinear, Disc-Intersection, PC Gaussian, VC Gaussian, and PC Linear models are at the third accuracy level. These models calculate the weights according to a function (Eqs. (17,18,20-22)) of either d_v in the 3D volume or d_s in the image plane, failing to consider the variation of intersection volume when the intersection angle changes. The differences in the function for weights calculation and the

weighting scope lead to the different SoSM and EoSM of each model. The Ray-length and VC direct models perform worst as quantified by the SoSM and EoSM, for they take the roughest approximation when calculating the intersection volume. They approximate either a voxel as a point or a ray bundle as a 1D line. The third index (SDoVV) evaluates how the imaging model meets the constraint $\sum_{j\in view_k} w_{ij} = 1$ (an implicit requirement for a correct imaging model). Among the imaging models analyzed in this work, not all of them meet this requirement. Notably, the Ray-length model has a quite large SDoVV compared with the others, which inevitably influences the reconstruction accuracy.

Table 2 Comparison of different imaging models in terms of assessment indices

Imaging model	SoSM	EoSM	SDoVV
VSF	0.999	0.001	0.000
Subvoxel	0.999	0.005	0.082
VC Bilinear	0.999	0.006	0.000
Disc-Intersection	0.998	0.008	0.036
PC Gaussian	0.996	0.011	0.097
VC Gaussian	0.994	0.014	0.078
PC Linear	0.991	0.018	0.093
Ray-length	0.886	0.078	0.230
VC Direct	0.832	0.097	0.000

The imaging models have been evaluated before reconstruction by the sample matrices. Whether this assessment is reasonable can then be confirmed experimentally. According to the previous analysis, the VSF imaging model made no approximation and had the best assessment indices, so the reconstruction calculated by the VSF imaging model was used as the benchmark. The reconstruction accuracy of the remaining imaging models was quantified through the correlation coefficient with the benchmark, defined as:

$$\gamma = \sum\left(\boldsymbol{f}_r - \overline{\boldsymbol{f}_r}\right)\left(\boldsymbol{f}_b - \overline{\boldsymbol{f}_b}\right) / \left(\sum\left(\boldsymbol{f}_r - \overline{\boldsymbol{f}_r}\right)^2 \sum\left(\boldsymbol{f}_b - \overline{\boldsymbol{f}_b}\right)^2\right)^{1/2}, \tag{24}$$

where $\boldsymbol{f}_r$ and $\boldsymbol{f}_b$ represent the reconstruction and the benchmark, respectively.

The benchmark is shown in Fig. 7(a) and the reconstructions using different imaging models are shown in the remaining panels of Fig. 7. Except for panels (h) and (i), all the other panels look very similar to the benchmark. In panels (a) to (h), the second lower intensity curved flame branch (circled in blue) can be clearly seen beside the main branch, which agrees well with the projections. The reconstruction accuracy of different imaging models shows large dependence on the SoSM and EoSM. As shown in Fig. 8, the obvious rule is that a larger SoSM or a smaller EoSM corresponds to a higher reconstruction accuracy for all models except the Ray-length model. The reconstruction from the Subvoxel model is closest to the benchmark calculated from the VSF model. The reconstruction accuracy of the VC Bilinear, Disc-Intersection, PC Gaussian, VC Gaussian, PC Linear, and VC Direct models decreases in order. This matches well with the reconstruction accuracy predicted by the EoSM index. These results indicate that both SoSM and EoSM are reasonable indices that reflect the accuracy of an imaging model.

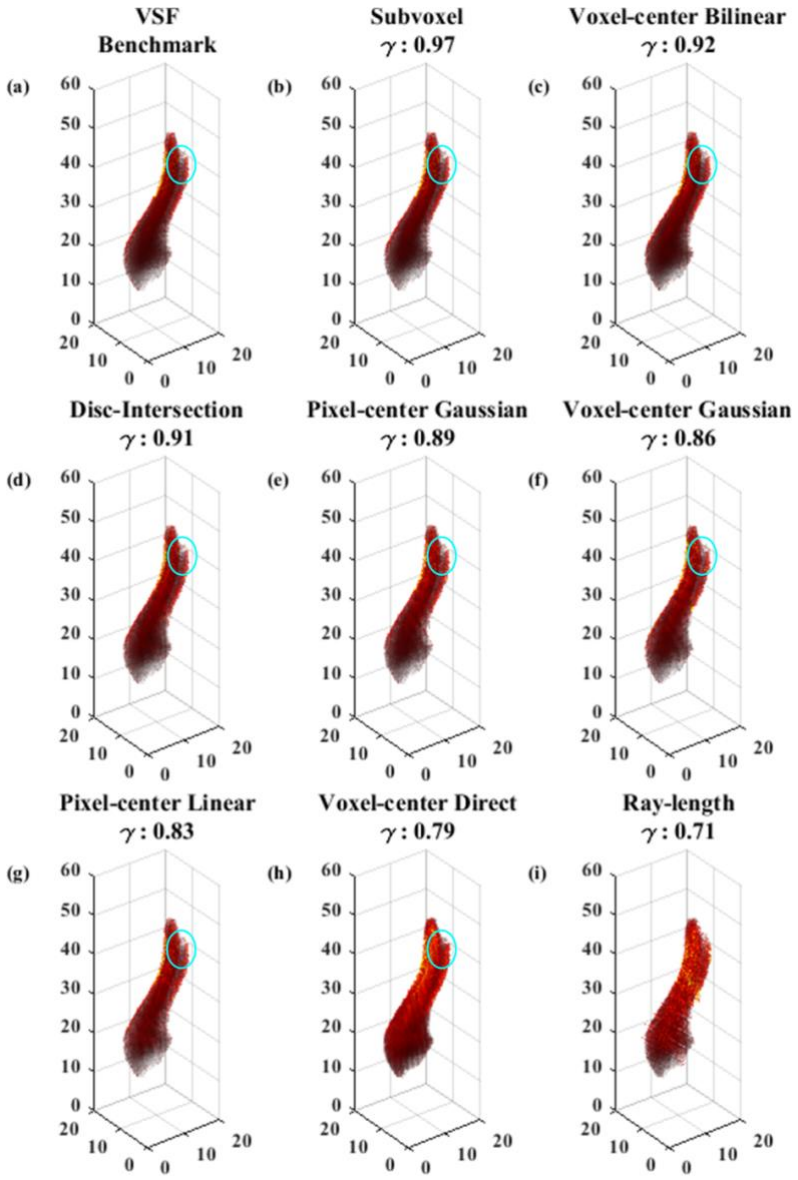

Fig. 7. The reconstructions using different imaging models.

However, only using these indices cannot fully explain the worst performance of the Ray-length model. The low accuracy of the Ray-length model was reflected by the much larger SDoVV than other models as shown in Table 2. The voxel volumes of all selected 27 voxels in the six views calculated from the Ray-length model are represented by yellow squares in Fig. 9. The blue band represents the maximum variation range of the voxel volumes for the other eight models. Compared with other imaging models, the deviations of voxel volume in the Ray-length model are significantly larger. For example, the 25th voxel in view 6 (circled in green) has a volume more than twice of the correct volume. The 26th voxel equals to zero in view 3, which means that this voxel does not have a corresponding pixel at all in this view. However, it does intersect with pyramid rays from four pixels in view 3. The Ray-length model fails to capture these intersections because pixels are approximated as discrete points. Once the thin ray from the pixel center does not intersect with a voxel, the corresponding weight is zero. This is an inherent disadvantage of the Ray-length model.

To conclude, the reconstruction accuracy of an imaging model can be inferred from the assessment indices of its sample matrix. Since the effectiveness of the SoSM and EoSM is the same,

considering either one of them is sufficient. In addition, the SDoVV also requires attention. For a reasonable imaging model, the typical SDoVV should be below 0.1. This sample matrix evaluation method allows us to assess the accuracy of an imaging model before reconstruction and can be broadly applied to many other similar setups. Once other new models are developed, the sample matrix method could be applied to assess the accuracy of the weight matrix and then predict the possible reconstruction accuracy of the new models.

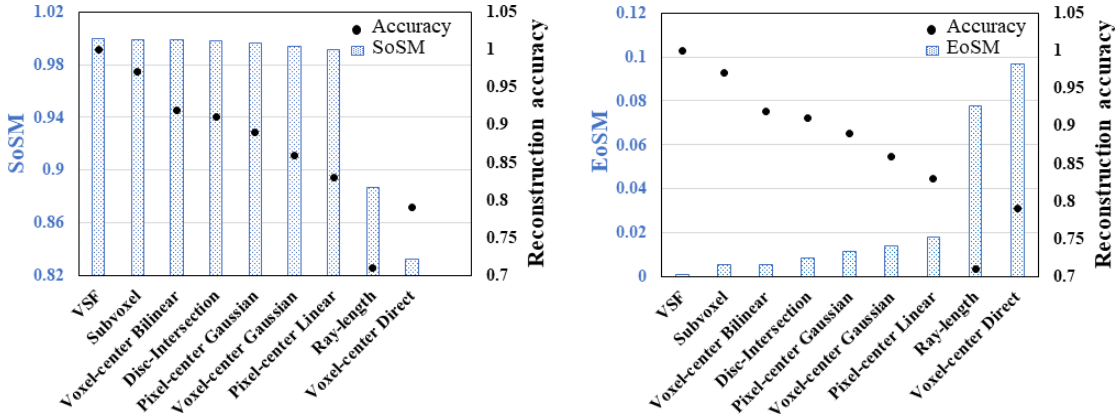

Fig. 8. The relationship between the reconstruction accuracy and assessment indices (SoSM and EoSM). The bars represent the SoSM or EoSM of the models and the dots describe the reconstruction accuracy.

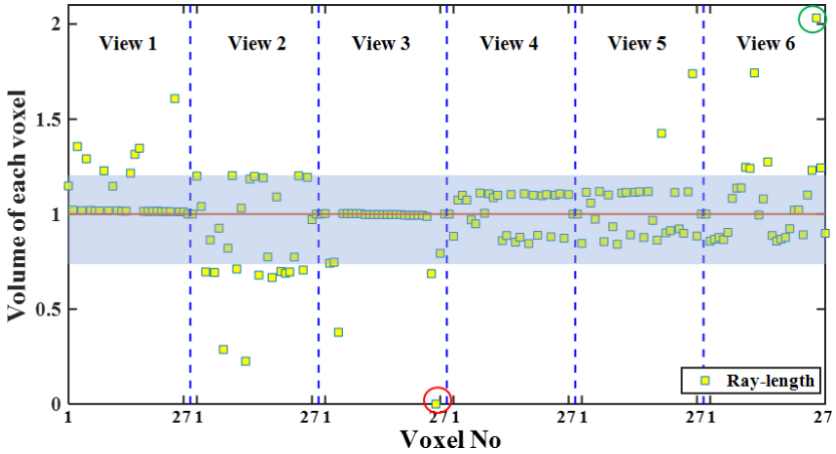

Fig. 9. The volumes of 27 selected voxels from the six views. The yellow squares represent the voxel volumes for the Ray-length model. The blue band represents the maximum variation range of the voxel volumes for the other eight models. The volume of the 26th voxel in view 3 equals zero circled in red. The volume of the 25th voxel in view 6 equals 2.2 circled in green.

4.3 Comparison of computational cost

The previous section compares the imaging models in terms of accuracy. This section focuses on the computational cost. The computational time of the whole weight matrix using different imaging models is presented in Table 3. The models in the third and fourth accuracy levels have little difference in computational cost even though their reconstruction accuracies are different. Among these two levels, the VC Bilinear imaging model is the most efficient. The Subvoxel model and VSF model have higher reconstruction accuracy, and need much more time. When computational cost is not a concern, VSF is the best choice for a high-accuracy reconstruction. Otherwise, the cheaper models such as Subvoxel and VC Bilinear are more suitable at the cost of compromised reconstruction accuracy.

Table 3 Comparison of different imaging models in terms of reconstruction accuracy and computational cost.

Imaging model	Reconstruction accuracy	Computational cost (s)
VSF	1	7121.23
Subvoxel	0.97	841.17
VC Bilinear	0.92	55.41
Disc-Intersection	0.91	56.42
PC Gaussian	0.89	58.83
VC Gaussian	0.86	56.98
PC Linear	0.83	58.09
Ray-length	0.71	44.44
VC Direct	0.79	54.09

4.4 Optimization of an imaging model

The accuracy of an imaging model can be judged upon the assessment indices of its sample matrix. Thus, these indices can also be used to optimize an imaging model. Here, the PC Gaussian model is taken as an example. As introduced in Section 2, the model accuracy is related to two factors: weighting scope and the function to determine each weight. A typical example is shown in the left bottom corner of Fig. 4. The pixels in red (defined as mismatched pixels) indicate the error caused by the wrong scope. The difference in values in green pixels indicates the error caused by the different weight calculation methods. Thus, these two factors should be considered simultaneously for optimization. Firstly, the best scope is determined by finding the scope radius that corresponds to the fewest mismatches, shown in Fig. 10(a). The next step is to determine the optimal function for weight calculation. In this PC Gaussian model, the function for the weight calculation is in the form presented in Eq. (18). The aim is to find the best σ that minimizes the EoSM and SDoVV of the sample matrix. However, the minimum EoSM and SDoVV do not always appear at the same σ as indicated in Fig. 11(a). The minimum EoSM appears at $\sigma = 0.44$, and the minimum SDoVV is at $\sigma = 0.49$. The SDoVV provides an *a priori* reasonable range of σ. From the red dotted line in Fig. 11(a), when σ is in the range of 0.42~0.7, the SDoVV is less than 0.1. According to the results in Sections 4.2 and 4.3, the accuracy of an imaging model is largely dependent on EoSM when the SDoVV is less than 0.1. So, σ is determined to be 0.44 for EoSM minimization. Through this optimization process, the accuracy of the PC Gaussian model improved from 0.89 to 0.92.

To verify the effectiveness of this optimization process, the impact of the scope radius and σ on the model accuracy is shown in Fig. 10(b) and Fig. 11(b), respectively. In Fig. 10(b), when the function for weights calculation is optimal, the optimal scope radius (1.19 voxels), inferred from the minimum of the total number of mismatched pixels, exactly corresponds to the most accurate reconstruction. When the scope is fixed at the optimum, the most accurate result appears at $\sigma = 0.47$ as shown in Fig. 11(b). This is slightly different from the inferred value of 0.44 because $\sigma = 0.44$ is determined only by the EoSM without considering SDoVV. Actually, the comprehensive influences of EoSM and SDoVV make the accuracy increase rate on the left side higher than the decrease rate on the right side and an optimal σ lies between these two turning points (marked by red stars in Fig. 11(a)). As it is hard to quantify the interaction between these indices on reconstruction accuracy, this optimization process could not always find the best parameters. But it is still a useful method to improve the accuracy of an imaging model to a relatively good level.

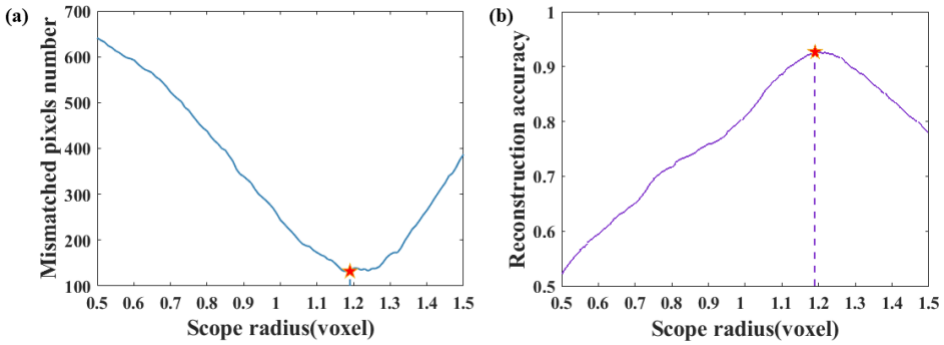

Fig. 10. The optimization of the weighting scope. (a) The total number of mismatched pixels as a function of scope radius. Here, σ is randomly set as 0.4. (b) The reconstruction accuracy as a function of scope radius. σ is set as the optimal value of 0.47.

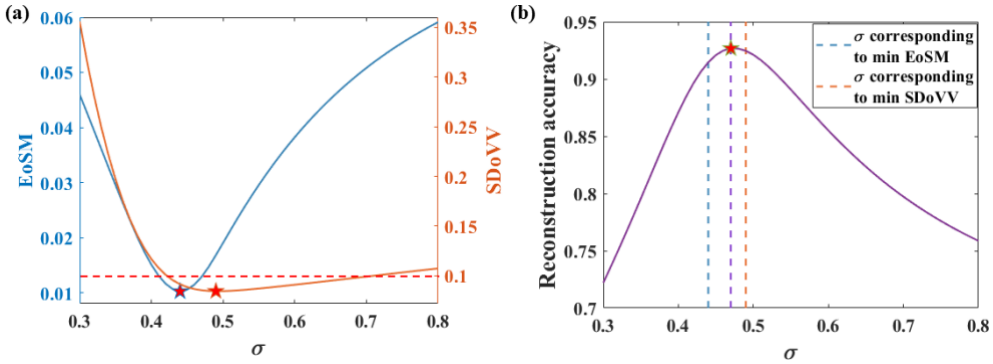

Fig. 11. The optimization of the function for weight calculation. (a) The EoSM and SDoVV as a function of scope radius. Here, the scope radius is set as the optimal value of 1.19 voxels. (b) The reconstruction accuracy as a function of σ when the scope radius is set as the optimal value of 1.19 voxels.

5. Impact of the spatial frequency of the target phantom

The previous section demonstrated the reconstruction of a real flame and evaluated the performance of nine imaging models. The flame used was smooth and continuous, and can be considered as an object with a low spatial frequency. However, objects with high spatial frequencies could also exist. How these imaging models perform for objects with different spatial frequencies is the topic of this section. It is not easy to generate objects with different spatial frequencies practically, so simulation studies are adopted. Eighty radially advancing sinusoidal phantoms are generated in a reconstruction volume of $40 \times 80 \times 80$ voxels, using Eqs. (25-27):

$$q = 0.5 \sin\left(\frac{2\pi}{\lambda} d\right) d(d_c < 40), \tag{25}$$

$$d = \sqrt{(X_{voxel} - X)^2 + (Y_{voxel} - Y)^2 + (Z_{voxel} - Z)^2}, \tag{26}$$

$$d_c = \sqrt{(X_{voxel} - X)^2 + (Y_{voxel} - Y)^2 + (Z_{voxel} - Z)^2}, \tag{27}$$

where λ is the wavelength of the sinusoidal function, X_{voxel}, Y_{voxel} and Z_{voxel} are the world coordinates of a certain voxel. (X, Y, Z) is a randomly generated position in world coordinates. The origin of the world coordinate system is at the upper left corner of the QoI. Ten sets of (X, Y, Z) were generated in our simulation. For each (X, Y, Z), the wavelength changes from ∞ to $10/3$ voxels to formulate 8 phantoms. Phantoms 1 to 8 represent one of the phantom sets as shown in Fig. 12. Three layers are plotted to illustrate the structure of each phantom. In conclusion, eighty phantoms were generated in total, and at each frequency, there were ten different phantoms. These

phantoms are then reconstructed with nine views equally spaced throughout π radians. The rankings of the models in terms of reconstruction accuracy for all phantoms are summarized statistically. We found that all the models follow the same ranking order as concluded in Section 4. The reconstructions of one of the phantoms are shown in Fig. 13 as an example. It can be noted that the reconstruction of layers near the edge may suffer extremely large errors when using the voxel-center Direct and Ray-length models. This proves that errors in the image model can corrupt the reconstructions.

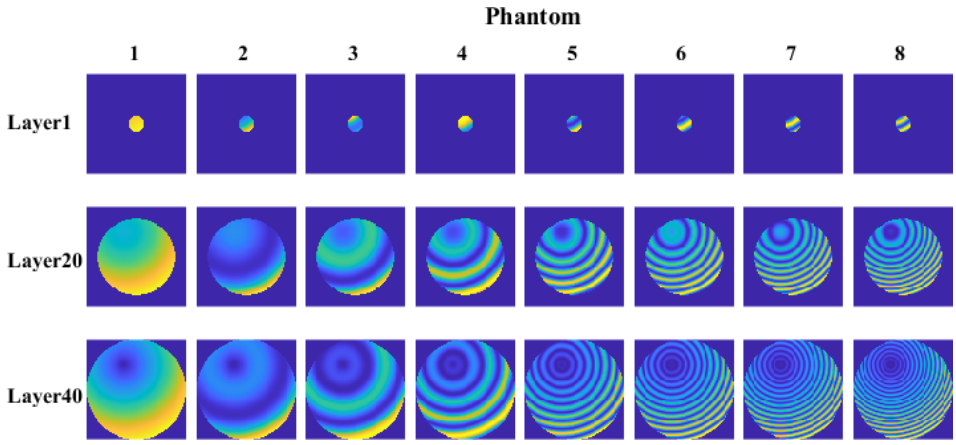

Fig. 12. Phantoms with different spatial frequencies. The wavelengths of phantom 1-8 are in order as follows: ∞, 160/3, 80/3, 40/3, 20/3, 5, 4, and 10/3 voxels.

For an imaging system with a certain view number and distribution, the reconstruction accuracy of an imaging model is a function of spatial frequency, *i.e.*:

$$\gamma = R(f), \tag{28}$$

where γ is the correlation coefficient between the reconstruction and the ground truth, which represents the reconstruction accuracy. f is the spatial frequency, which equals the reciprocal of the wavelength. The function R of all the nine imaging models are shown in Fig. 14. For the VSF model, no errors exist in the imaging model, but the reconstruction accuracy still decreases at high frequencies. This is caused by the rank deficiency of the underdetermined equation system $\mathbf{p} = \mathbf{W}\mathbf{f}$, in which the number of unknowns is greater than the equations. For other models, the error from the imaging model could be coupled with the inversion error. So, the reconstruction accuracy of VSF and other models are compared in the same subgraph to observe the error increase caused by different imaging models. The blue asterisks in Fig. 14 indicate the accuracy of each model. For the first six imaging models, errors are not magnified during the inversion for any of the spatial frequencies tested. For the VC Direct model and Ray-length model, the reconstruction accuracy decreased more rapidly than the VSF model at high frequencies which indicates that the error from the imaging model will be amplified in the inversion process when the frequencies increase. So, when using these models, objects with higher spatial frequencies will suffer larger errors. To verify the generality of the above conclusions, ten sets of phantoms (each set has eight phantoms with the same (X, Y, Z) but different spatial frequencies) were used totally. The rest nine sets indicate the same conclusion drawn above.

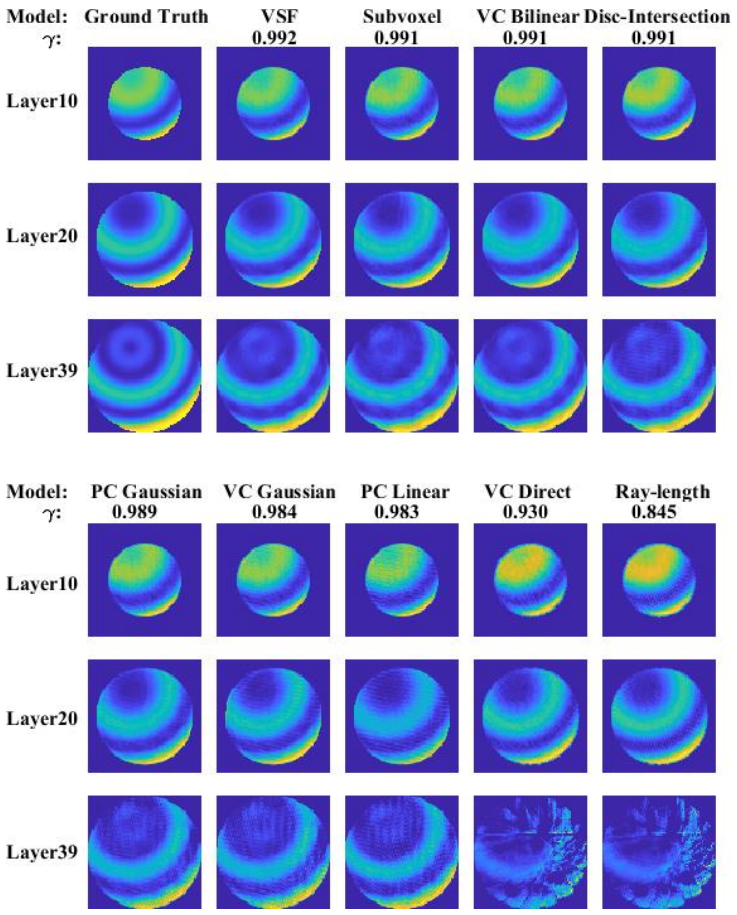

Fig. 13. Reconstruction of phantom 4 using nine views equally spaced throughout π radians.

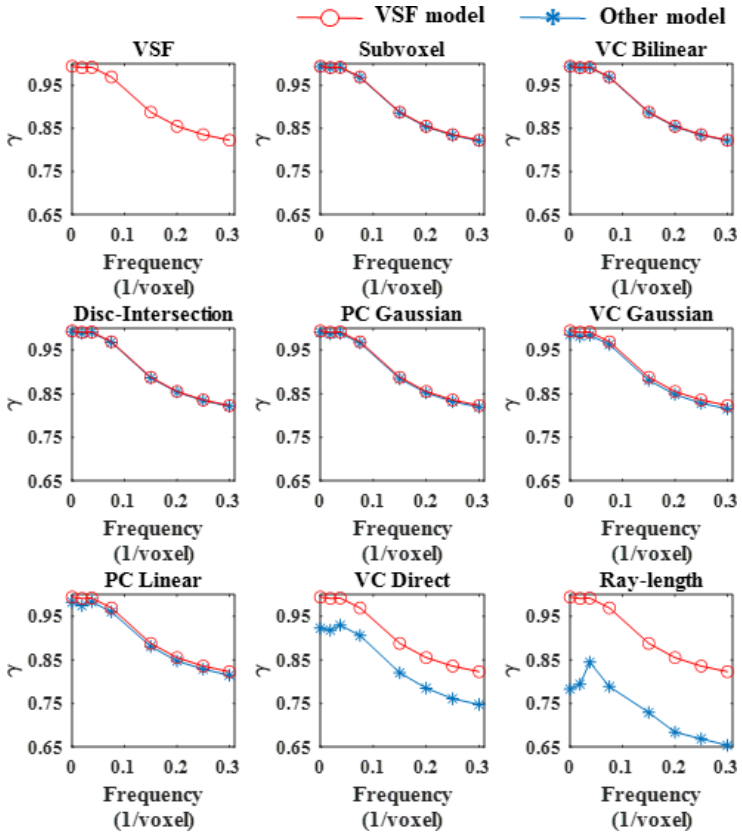

Fig. 14. The relationship between the reconstruction accuracy and the spatial frequency of the phantom.

6. Comparison of imaging models for velocity field reconstruction

The previous results suggest that there are non-negligible accuracy differences between these models when reconstructing a continuous scalar field. But it should be noticed that in a TPIV experiment, the reconstructed object consists of discrete scattered light fields and an additional correlation step is needed to calculate the velocity field based on these. Thus, this section aims to further assess the performance of these imaging models for velocity field reconstruction via simulations.

The VoI and the camera system were assumed the same as those used in the FCT experiment with zero calibration error. The velocity field used is an ABC flow [44] defined as:

$$u(\boldsymbol{X},t) = A(\sin kZ) + C(\cos kY), \tag{29}$$

$$v(\boldsymbol{X},t) = B(\sin kX) + A(\cos kY), \tag{30}$$

$$w(\boldsymbol{X},t) = C(\sin kY) + B(\cos kX), \tag{31}$$

with $\mathrm{A} = \sqrt{3}, \mathrm{B} = \sqrt{2}, \mathrm{C} = 1$ and $k = 0.012$. The VoI was seeded with 3 voxel diameter Gaussian particles at a seeding density of $N_{ppp} = 0.0512$. Here, N_{ppp} represents the number of particles per pixel (ppp). Then each particle was displaced according to the velocity field, for a time step of 5/18 s, producing displacements of 1.8–2.9 voxels. Artificial projections at six views were simulated with the VSF model at two adjacent time instants and projections at the first time instant are shown in Fig. 15. Here, four representative imaging models are selected from different accuracy levels to reconstruct the scattered light intensity fields. These models are the VSF, Subvoxel, PC Gaussian, and VC Direct. The reconstructions at the first instant are shown in Fig. 16. The reconstruction accuracy here is represented by the correlation coefficient between the reconstructions and the ground truth. Reconstructions from the first three models look similar, but actually have a decreasing reconstruction accuracy. The VC Direct model has a visible difference in terms of particle intensity. Further, the velocity field is calculated with a single-pass cross-correlation algorithm [49]. The interrogation window size was $28 \times 28 \times 28$ voxels, with an overlap of 75%. The mean error of velocity estimation (E_v) is calculated as:

$$E_v = \sqrt{(u_c - u_t)^2 + (v_c - v_t)^2 + (w_c - w_t)^2}/n, \tag{32}$$

where u_c, v_c, w_c are the velocities in three directions estimated from cross-correlation and the u_t, v_t, w_t are the true velocities. n is the total number of estimated velocity vectors. The velocity estimation errors of VSF, Subvoxel, PC Gaussian and VC Direct model are 0.162, 0.170, 0.168 and 0.174 voxels. VSF is still the best model, and VC is the worst. Although the Subvoxel model has a higher reconstruction accuracy for the scattered light reconstruction than the PC Gaussian model, the final error in velocity estimation for this model is larger than that of the PC Gaussian model. This may be attributed to the cross-correlation process, the accuracy of which is related to the similarity of the scattered light distributions at the two consecutive frames rather than the absolute accuracy of the reconstruction at a single frame. In other words, the cross-correlation process relaxes the requirement for reconstruction accuracy.

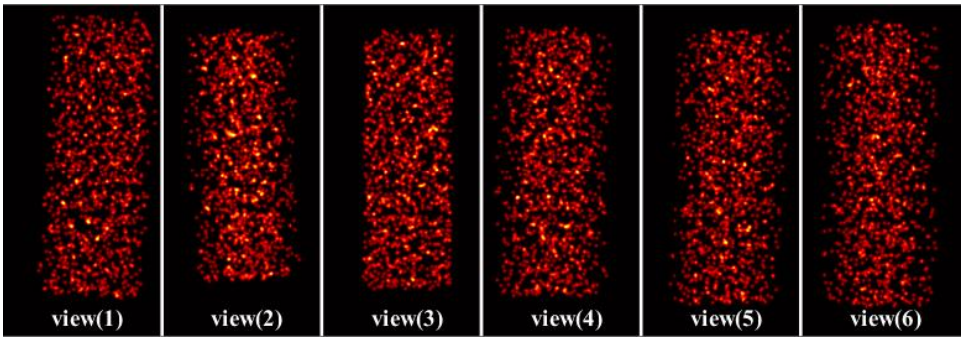

Fig. 15. The simulated projections at six views.

The above conclusion is made at a seeding density of 0.0512 ppp. What if the seeding density increases which is similar to the spatial frequency increase for a continuous phantom? To study this problem, a set of scattered light fields with increasing seeding density are generated to estimate the same velocity field. Seven seeding densities ranging from 0.05 to 0.16 ppp (a typical density range for TPIV) were simulated. At each seeding density, ten cases were generated to ensure the reliability of the results. Figure 17 shows the velocity estimation error as a function of seeding density using different imaging models, and each point represents the average value of E_v from ten cases, the error bar represents the standard deviation of the ten cases. First, when seeding density changes from 0.05 to 0.16 ppp, the VSF model is always the best model and the VC Direct is the worst. The remaining two models produce similar error levels which may be explained by the effect of the cross-correlation process. Second, compared with the other three models, the velocity estimation error using the VC Direct model increases significantly with increasing seeding density. This conclusion is similar to that drawn from the continuous field with different spatial frequencies. A less accurate model like VC Direct model performs worse at higher seeding density.

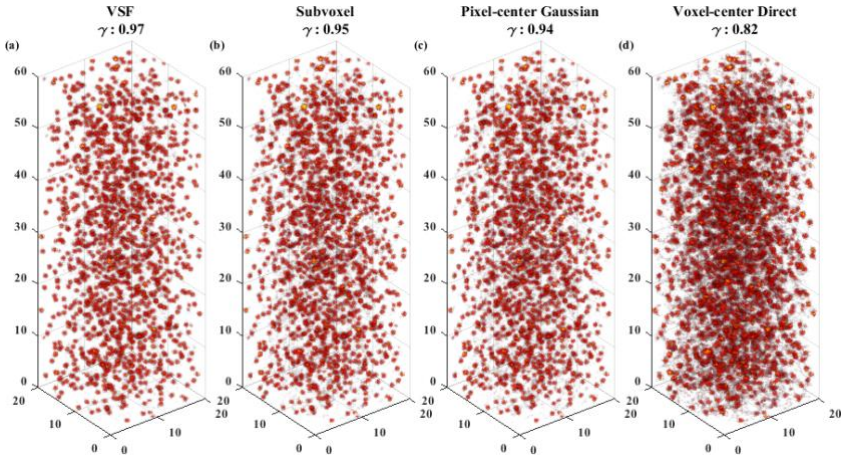

Fig. 16. The reconstruction of discrete scattered light fields using the selected imaging models.

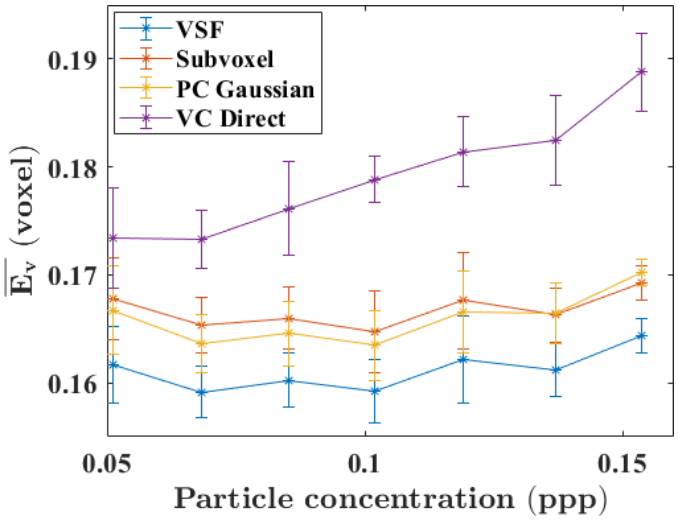

Fig. 17. Velocity estimation error as a function of seeding density using different imaging models.

7. Conclusions

This work focuses on the comparison of imaging models in VT modalities applied to flow diagnostics. First, nine commonly adopted imaging models were introduced and divided into two categories: pixel-centric and voxel-centric imaging models. Then a sample matrix assessment method was developed to evaluate the accuracy of the weight matrix, predict the reconstruction accuracy of the imaging model, and optimize some imaging models that have adjustable parameters in a FCT experiment. After that, the impact of increasing spatial frequency in the target field on the reconstruction accuracy of different imaging models is analyzed by a simulation study. Finally, the effectiveness of the models when applied to TPIV was investigated. The main conclusions are summarized as follows.

The accuracy of an imaging model could be judged by three indices *i.e.* SoSM, EoSM and SDoVV. An increase in the EoSM or a decrease in SoSM means a decrease in the reconstruction accuracy. The SDoVV index will have an obvious impact when it is much larger than zero (more than 0.1). In addition, these sample matrix indices can be used to optimize the performance of an imaging model including empirical parameters. In the flame reconstruction case detailed in Section 4. The reconstruction accuracy of PC Gaussian model improved from 0.89 to 0.92.

The accuracy and computational cost of the nine imaging models introduced are first evaluated in a FCT experiment and then verified in the simulation studies with phantoms of different spatial frequencies. The VSF is the most accurate model among all the imaging models investigated here, as it applies no approximation to the size or shape of the voxels or pixels during weight matrix calculation. But it also posed the greatest computational cost. The Subvoxel model is not far behind in terms of accuracy, since it only approximates the pixel as a sphere, and the variation of intersecting angles is partly considered. This model requires 1/10 of the computational cost in comparison with the VSF model. Then the VC Bilinear, Disc-Intersection, PC Gaussian, VC Gaussian, and PC Linear models neglect the influence of intersecting angle, so these models all belong to the third accuracy level. The accuracy of these models is largely dependent on their weighting scope and the function to calculate the weights. The worst models are the VC Direct and Ray-length models, for either the pixels or voxels are approximated as single points. The computational costs for models in third and fourth accuracy levels are similar and are about 6% of the Subvoxel model in the flame reconstruction case in this study. The accuracy orders of the nine

imaging models are the same in both FCT experiment and the simulations for spatial frequency. So, these conclusions could be applied to other similar VT modalities for 3D luminosity field reconstruction.

The increase in spatial frequency of the target field has an impact on the reconstruction accuracy of both VC Direct and Ray-length models. The error caused by the VC Direct model and Ray-length model will increase at higher spatial frequencies. The errors of other imaging models do not change obviously for phantoms with wavelengths ranging from 10/3 to ∞.

In the TPIV simulation cases, it is found that the most accurate model, VSF, corresponds to the smallest error in the final velocity estimation in a seeding density ranging from 0.05 to 0.16 ppp. The models in the second and third accuracy levels may have little difference in velocity estimation due to the cross-correlation process. The model belonging to the fourth accuracy level has the largest error in velocity estimation and this error will increase from 0.175 to 0.195 voxels with the increase of seeding density.

The conclusions drawn here have some practical significance. The sample matrix method provides some simple assessment indices to compare the accuracy of a newly developed imaging model with the existing ones. Furthermore, the existing model which has parameters estimated based on experience in the past could now be optimized by minimizing the EoSM in the sample matrix. The assessment of the nine imaging models demonstrates that an accurate imaging model like VSF should be the best choice for both scalar field reconstruction and velocity estimation when there is no limitation on computational cost. Otherwise, Subvoxel and VC Bilinear are good alternatives for scalar field reconstruction. The PC Gaussian model may be suitable for TPIV, since it can achieve similar velocity estimation accuracy with a lower computational cost compared with the Subvoxel model.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (Grant No. 51976122, 52061135108), Foundation of Science and Technology on Combustion and Explosion Laboratory (Grant No. 6142603200508) and National Science and Technology Major Project (Grant No. 2017-III-0007-0033). We would like to thank Prof. Samuel Grauer for his help with this work.

References

[1] Bennett, K. and R.L. Byer, *Optical tomography: experimental verification of noise theory.* Opt Lett, 1984. **9**(7): p. 270-2.

[2] Hertz, H.M. and G.W. Faris, *Emission tomography of flame radicals.* Opt Lett, 1988. **13**(5): p. 351-3.

[3] Wang, J., et al., *High spatial resolution computed tomography of chemiluminescence with densely sampled parallel projections.* Opt Express, 2019. **27**(15): p. 21050-21068.

[4] Floyd, J., *A 3D Time Resolved Sensor for Turbulent Combustion*, in *Department of Mechanical Engineering*. 2009, Imperial College London: London.

[5] Cai, W., et al., *Numerical and experimental validation of a three-dimensional combustion diagnostic based on tomographic chemiluminescence.* Optics Express, 2013. **21**(6).

[6] Jin, Y., et al., *Three-dimensional dynamic measurements of CH* and C2* concentrations in flame using simultaneous chemiluminescence tomography.* Opt Express, 2017. **25**(5): p. 4640-4654.

[7] Halls, B.R., et al., *4D spatiotemporal evolution of combustion intermediates in turbulent flames using burst-mode volumetric laser-induced fluorescence.* Opt Lett, 2017. **42**(14): p. 2830-2833.

[8] Halls, B.R., et al., *Single-shot, volumetrically illuminated, three-dimensional, tomographic laser-induced-fluorescence imaging in a gaseous free jet.* Optics Express, 2016. **24**(9).

[9] Li, T., et al., *Tomographic imaging of OH laser-induced fluorescence in laminar and turbulent jet flames.* Measurement Science and Technology, 2018. **29**(1).

[10] Wang, Q., et al., *Three-dimensional concentration field imaging in a swirling flame via endoscopic volumetric laser-induced fluorescence at 10-kHz-rate.* Science China Technological Sciences, 2020. **63**(10): p. 2163-2168.

[11] Gomez, M., et al., *Time-resolved Volumetric (4D) Laser Induced Fluorescence Imaging of Primary Spray Breakup*, in *AIAA SCITECH 2022 Forum*. 2022.

[12] Meyer, T.R., et al., *High-speed, three-dimensional tomographic laser-induced incandescence imaging of soot volume fraction in turbulent flames.* Opt Express, 2016. **24**(26): p. 29547-29555.

[13] Elsinga, G.E., et al., *Tomographic particle image velocimetry.* Experiments in Fluids, 2006. **41**(6): p. 933-947.

[14] Kumashiro, K., A.M. Steinberg, and M. Yano, *High Spatial Resolution 3D Fluid Velocimetry by Tomographic Particle Flow Velocimetry*, in *AIAA Scitech 2019 Forum*. 2019.

[15] Liu, H., Z. Yang, and W. Cai, *Application of three-dimensional diagnostics on the direct-current electric-field assisted combustion.* Aerospace Science and Technology, 2021. **112**.

[16] Gordon, R., R. Bender, and G.T. Herman, *Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography.* Journal of Theoretical Biology, 1970. **29**(3): p. 471-481.

[17] Gordon, R., *A tutorial on art (algebraic reconstruction techniques).* IEEE Transactions on Nuclear Science, 1974. **21**(3): p. 78-93.

[18] Herman, G.T. and A. Lent, *Iterative reconstruction algorithms.* Computers in Biology and Medicine, 1976. **6**(4): p. 273-294.

[19] Worth, N.A. and T.B. Nickels, *Acceleration of Tomo-PIV by estimating the initial volume intensity distribution.* Experiments in Fluids, 2008. **45**(5): p. 847-856.

[20] Yang, W.Q., et al., *An image-reconstruction algorithm based on Landweber's iteration method for electrical-capacitance tomography.* Measurement Science and Technology, 1999. **10**(11): p. 1065-1069.

[21] Rossberg, A. and H. Funke, *Determining the radial pair distribution function from X-ray absorption spectra by use of the Landweber iteration method.* J Synchrotron Radiat, 2010. **17**(2): p. 280-8.

[22] Szlávecz, Á., et al., *GPU-based acceleration of the MLEM algorithm for SPECT parallel imaging with attenuation correction and compensation for detector response.* IFAC Proceedings Volumes, 2011. **44**(1): p. 6195-6200.

[23] Gaitanis, A., et al., *PET image reconstruction: A stopping rule for the MLEM algorithm based on properties of the updating coefficients.* Comput Med Imaging Graph, 2010. **34**(2): p. 131-41.

[24] Huang, J., et al., *Limited-projection volumetric tomography for time-resolved turbulent combustion diagnostics via deep learning.* Aerospace Science and Technology, 2020. **106**.

[25] Wang, G., J.C. Ye, and B. De Man, *Deep learning for tomographic image reconstruction.* Nature Machine Intelligence, 2020. **2**(12): p. 737-748.

[26] Yu, T. and W. Cai, *Benchmark evaluation of inversion algorithms for tomographic absorption spectroscopy.* Appl Opt, 2017. **56**(8): p. 2183-2194.

[27] Atkinson, C.H. and J. Soria, *Algebraic reconstruction techniques for tomographic particle image velocimetry*, in *Proc. 16th Australasian Fluid Mechanics Conf*. 2007: Gold Coast, Australia. p. 191–

8.

[28] Petra, S., et al., *Tomographic Image Reconstruction in Experimental Fluid Dynamics: Synopsis and Problems*, in *Mathematical Modelling of Environmental and Life Sciences Problems*, R.A. Bucuresti, Editor. 2008.

[29] Wiseman, S.M., et al., *Measurements from flame chemiluminescence tomography of forced laminar premixed propane flames.* Combustion and Flame, 2017. **183**: p. 1-14.

[30] Thomas, L., et al. *Influence of geometric parameters and image preprocessing on tomo-PIV results*. in *15th International Symposium on Applications of Laser Techniques to Fluid Mechanics*. 2010. Lisbon, Portugal.

[31] Worth, N.A. and J.R. Dawson, *Tomographic reconstruction of OH* chemiluminescence in two interacting turbulent flames.* Measurement Science and Technology, 2013. **24**(2).

[32] Floyd, J., P. Geipel, and A.M. Kempf, *Computed Tomography of Chemiluminescence (CTC): Instantaneous 3D measurements and Phantom studies of a turbulent opposed jet flame.* Combustion and Flame, 2011. **158**(2): p. 376-391.

[33] Schanz, D., et al., *Non-uniform optical transfer functions in particle imaging: calibration and application to tomographic reconstruction.* Measurement Science and Technology, 2013. **24**(2).

[34] Yu, T., H. Liu, and W. Cai, *On the quantification of spatial resolution for three-dimensional computed tomography of chemiluminescence.* Opt Express, 2017. **25**(20): p. 24093-24108.

[35] Liu, N. and L. Ma, *Regularized tomographic PIV for incompressible flows based on conservation of mass.* Appl Opt, 2020. **59**(6): p. 1667-1677.

[36] Ruan, C., et al., *Experimental characterization of the spatiotemporal dynamics of a turbulent flame in a gas turbine model combustor using computed tomography of chemiluminescence.* Energy, 2019. **170**: p. 744-751.

[37] Cornic, P., et al., *Another look at volume self-calibration: calibration and self-calibration within a pinhole model of Scheimpflug cameras.* Measurement Science and Technology, 2016. **27**(9).

[38] Yu, T., et al., *A quantitative evaluation method of 3D flame curvature from reconstructed flame structure.* Experiments in Fluids, 2020. **61**(2).

[39] Halls, B.R., et al., *Two-color volumetric laser-induced fluorescence for 3D OH and temperature fields in turbulent reacting flows.* Opt Lett, 2018. **43**(12): p. 2961-2964.

[40] Shui, C.Y., H.C. Liu, and W.W. Cai, *Benchmark evaluation of tomographic algorithms for simultaneous reconstruction of temperature and volume fraction fields of soot and metal-oxide nanoparticles in non-uniform flames.* Science China-Technological Sciences, 2021. **64**(2): p. 237-250.

[41] Liu, D., et al., *Simultaneous Measurement of Three-Dimensional Soot Temperature and Volume Fraction Fields in Axisymmetric or Asymmetric Small Unconfined Flames With CCD Cameras.* Journal of Heat Transfer, 2010. **132**(6).

[42] Zhang, Z., *A flexible new technique for camera calibration.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(11): p. 1330-1334.

[43] Bouguet, J.-Y. *"Camera Calibration Toolbox for Matlab"*.

[44] Thomas, L., B. Tremblais, and L. David, *Optimization of the volume reconstruction for classical Tomo-PIV algorithms (MART, BIMART and SMART): synthetic and experimental studies.* Measurement Science and Technology, 2014. **25**(3).

[45] Bosbach, J., M. Kühn, and C. Wagner, *Large scale particle image velocimetry with helium filled*

soap bubbles. Experiments in Fluids, 2008. **46**(3): p. 539-547.

[46] Williams, A., et al., *An efficient and robust ray-box intersection algorithm*, in *ACM SIGGRAPH 2005 Courses*. 2005, Association for Computing Machinery: Los Angeles, California. p. 9–es.

[47] Xu, Y., et al., *Metropolis Monte Carlo simulation scheme for fast scattered X-ray photon calculation in CT.* Opt Express, 2019. **27**(2): p. 1262-1275.

[48] Han, X., et al., *The Effect of Stratification Ratio on the Macrostructure of Stratified Swirl Flames: Experimental and Numerical Study.* Journal of Engineering for Gas Turbines and Power, 2018. **140**(12).

[49] Ben Gida, H., Gurka, R., and Liberzon, A. . *"OpenPIV-MATLAB — An open-source software for particle image velocimetry; test case: Birds' aerodynamics," SoftwareX, 12, 100585, .* 2020; Available from: https://doi.org/10.1016/j.softx.2020.100585.