Transmittance-based Extinction and Viewpoint Optimization

Paul Himmler [ID] and Tobias Günther [ID]

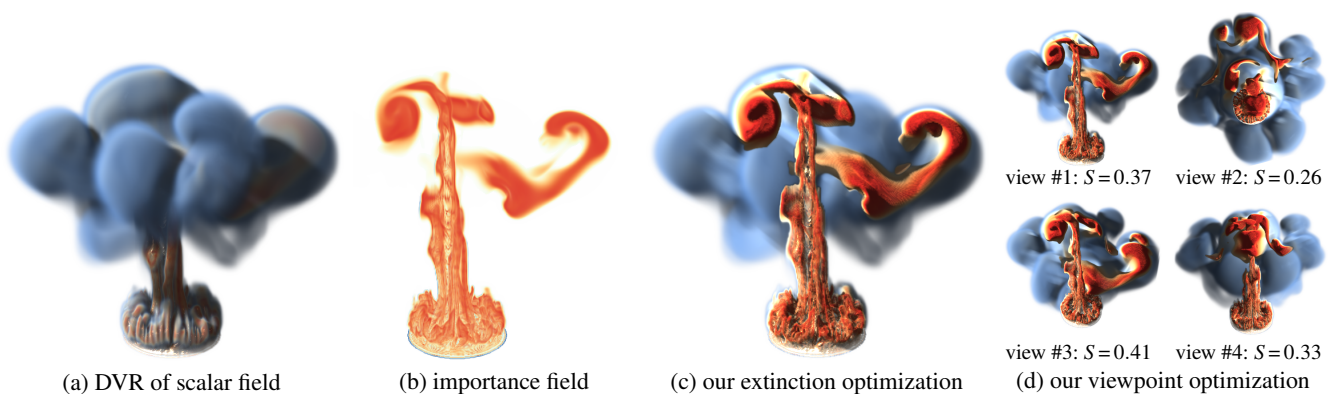Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Figure 1: *Conventional direct volume rendering (DVR) in (a) is often prone to occlusions. Given the importance field in (b), we propose a transmittance-based extinction optimization in (c) and a concurrent viewpoint selection in (d), which together reveal interesting structures. A visibility score S is listed for each of the exemplary views in (d), indicating that view #3 conveyed most of the important regions.*

Abstract
A long-standing challenge in volume visualization is the effective communication of relevant spatial structures that might be hidden due to occlusions. Given a scalar field that indicates the importance of every point in the domain, previous work synthesized volume visualizations by weighted averaging of samples along view rays or by optimizing a spatially-varying extinction field through an energy minimization. This energy minimization, however, did not directly measure the contribution of an individual sample to the final pixel color. In this paper, we measure the visibility of relevant structures directly by incorporating the transmittance into a non-linear energy minimization. For the first time, we not only perform a transmittance-based extinction optimization, we concurrently optimize the camera position to find ideal viewpoints. We derive the partial derivatives for the gradient-based optimization symbolically, which makes the application of automatic differentiation methods unnecessary. The transmittance-based formulation gives a direct visibility measure that is communicated to the user in order to make aware of potentially overlooked relevant structures. Our approach is compatible with any measure of importance and its versatility is demonstrated in multiple data sets.

CCS Concepts
*• **Human-centered computing** → **Scientific visualization;** Visualization techniques; • **Computing methodologies** → **Visibility;***

1. Introduction

Scientific visualization is frequently concerned with the visualization of three-dimensional spatial data. Depending on the application, different parts of a spatial domain might be more or less relevant to the user. Unfortunately, many of those interesting regions might be occluded and thereby hidden. In the literature, this is known as the *occlusion problem* [VG05, MMD10, GRT13, AZD17]. To reduce occlusions two strategies are available: either the transparency of less relevant structures is increased to clear the view on meaningful structures, or a viewpoint is chosen in which less occlusion occurs. Either way, a measure is needed that determines how much of the relevant information is visible on the screen. Previous opacity and

extinction optimization approaches [GRT13, GRT14, GSME*14, GTG17, AZD17, BRGG20] all utilized a metric that assessed the visibility of important information *indirectly*. That is, an energy was defined that rises when occlusions occur, but how well the resulting structures are actually visible on screen was not calculated directly. This indirect optimization had the advantage that a solution could be found analytically with a subsequent smoothing step [AZD17, GTG17], which made this approach simple and efficient. The indirect nature, however, made the adjustment of energy weights difficult and there was no direct feedback on how much relevant information is seen. In this paper, we address this challenge by employing a quality metric that directly measures the impact that each voxel has on the screen, which is expressed by the so-called *transmittance*. Having such a direct quality metric not only allows us to phrase the adjustment of the transparency in an optimization-based manner, for the first time, we are also able to optimize the viewpoint simultaneously. Fig. 1 shows results of our extinction and viewpoint optimization. We present a variational formulation that leads to a non-linear optimization problem, for which we provide an efficient parallelized algorithm that is implemented on the GPU. In analogy to the volume rendering integral, we derive integral equations for the required partial derivatives that can be computed directly in a ray marching loop. The implementation of an automatic differentiation algorithm is therefore not necessary. In summary, we make the following contributions:

- In contrast to the previous extinction optimization [AZD17], we utilize a *transmittance-based* quality metric for the assessment of how well relevant information is visible on screen.
- We concurrently optimize for the extinction field *and* for the viewpoint that communicate the relevant information best.
- Using variational calculus, we derive closed-form equations for the required derivatives, which can be calculated by ray marching.
- We show to the user how much relevant information is visible.

In the following, we first summarize existing work on transfer function, extinction, and viewpoint optimization, which is followed by a brief description of direct volume rendering. Afterwards, the transmittance-based optimization of extinction and viewpoint is introduced. The approach is evaluated on several data sets, covering application domains such as fluid dynamics, geophysics, and medicine. The paper is concluded with an outlook on future work.

2. Related Work

2.1. Transfer Function Optimization

Direct volume rendering uses transfer functions to map from a scalar field to optical properties such as color and transparency [JSYR14]. How well information is perceived strongly depends on the choice of transfer function. To avoid tedious trial and error exploration, He et al. [HHKP96] stochastically mutated transfer functions from an initial pool based on interactive user preferences and quality metrics such as entropy, histogram variance, and edge energy. Using the scalar field and its first and second-order directional derivatives along the gradient, Kindlmann and Durkin [KD98] semi-automatically adjusted transfer functions to reveal material boundaries. Wu and Qu [WQ07] let users choose individual features from multiple direct volume renderings that have been made with different transfer functions, and then synthesized a view that shows

or removes selected features. Correa and Ma [CM11] augmented a transfer function editor with a visibility histogram to communicate how well certain value ranges are represented in the image. Further, they represented the opacity transfer function with a Gaussian mixture model and optimized for an improved visibility. To minimize the informational divergence between data and view, Ruiz et al. [RBB*11] minimized the Kullback-Leibler distance between the viewed visibility distribution and a user-defined target distribution. All of the methods above modeled transfer functions as functions of the scalar value (and its derivatives). To address occlusions the transfer function was varied spatially, which is discussed next.

2.2. Opacity and Extinction Optimization

Several approaches varied the opacity spatially along view rays in order to remove occlusions [VG05]. Early smart visibility approaches [VKG04, VKG05] displayed per pixel the object with maximum importance along the ray or switched to sparser object representations to support windowing. The opacity has also been adjusted in a context-dependent manner [BGKG05]. Further lines of research avoided occlusions through exploded views [APH*03] and spatial deformation [CSC07]. A finite-difference based optimization of discrete opacity values along rays was performed by Chan et al. [CWM*09], who improved the perception of direct volume renderings by considering visibility, structural shape and image variations, and perception theory. Instead, we utilize a continuous measure for pixel contribution and derive derivatives symbolically. Maximizing the contribution of the farthest relevant sample point on a ray, Marchesin et al. [MMD10] derived a blending weight for the samples. Ament et al. [AZD17] extended a linear visibility optimization from semi-transparent geometry [GRT13, GRT14, GSME*14] to volume data and realized that the optimization has a closed-form solution if smoothing is done separately in a post-process. This idea has later been picked up for geometric data [GTG17, BRGG20, ZRPD20]. The problem is that this formulation reduces the visibility of a voxel *indirectly* based on how much importance is gathered in front or behind it, which does not measure how visible a certain voxel truly is. In this paper, we measure the actual contribution of a fragment to the final pixel color [CM11]. Unlike Correa and Ma [CM11], we optimize for an extinction field and not for a parametric transfer function. Further, Ament et al. [AZD17] performed a separate extinction optimization from the light source to illuminate relevant objects, which can lead to inconsistent lighting. In this paper, we optimize the extinction from camera and light source concurrently to obtain an illumination that is consistent with the visible volume.

2.3. Viewpoint Optimization

As scientific data sets increase in resolution quickly, the automatic generation of meaningful preview images or animation paths became early on a research interest. By decomposing a volume data set into a set of components and by using isosurface properties as feature descriptor, a globally optimal viewpoint has been searched for as a compromise between locally optimal viewpoints for each component [TFTN05]. Another optimization target was an even opacity distribution, an even distribution of salient features, and more perceived curvatures [SJ06]. The opacity entropy [SJ06] has later been applied in a differentiable volume renderer to find best

viewpoints [WW22]. To balance between global structures or local details two descriptors were introduced [TLB*09]: a shape view descriptor which evaluates the overall feature orientation, and a detail view descriptor which measures the amount of visible details on boundary structures by estimating variances. A first information theoretic approach based on entropy measures accounted for the transfer function, the data distribution, the voxel visibility, view-likelihood and view-stability [BS05]. Several following methods likewise expressed the quality of a viewpoint based on information theoretic measures [CJ10], including multi-scale entropy [VMN08]. The information theoretic approach was also applied to judge the image quality after a streamline generation in flow visualization [XLS10, LMSC11, TMWS13]. To guide users towards viewpoints that were chosen by experts, image similarity metrics have been evaluated by majority voting [TWC*16] and have been estimated by deep learning [YLLY19]. In this paper, we optimize for the extinction in the volume by measuring the amount of visible important information, as given by the user, and we select a suitable viewpoint simultaneously. None of the methods above attempted to address both problems at the same time.

2.4. Direct Volume Rendering

Direct volume rendering is a popular approach for the visualization of three-dimensional scalar fields [JSYR14]. For this, the scalar field $s(\mathbf{x}) : \mathcal{D} \to \mathbb{R}$, which is defined in a spatial domain $\mathcal{D} \subseteq \mathbb{R}^3$, is mapped via transfer functions to optical properties. Based on those, a participating medium is constructed, which can be rendered from a given viewpoint using light transport simulation methods [Jar08]. These computer graphics techniques may include illumination effects such as reflections and shadows to improve the spatial perception. The volume rendering equation determines the radiance $L(\mathbf{x}_0 \leftarrow \omega)$ seen at point $\mathbf{x}_0$ coming in from direction ω:

$$L(\mathbf{x}_0 \leftarrow \omega) = \int_{\mathbf{x}_0}^{\mathbf{x}_1} T(\mathbf{x}_0 \leftrightarrow \mathbf{x}) \cdot \mu_s(\mathbf{x}) \cdot L_i(\mathbf{x} \leftarrow \omega) \, d\mathbf{x} \qquad (1)$$

The radiance is collected along the view ray $\mathbf{x}(t) = \mathbf{x}_0 + t \cdot \omega$, consisting of three terms: the view ray transmittance $T(\mathbf{x}_0 \leftarrow \mathbf{x})$, the scattering coefficient $\mu_s(\mathbf{x})$, and the incident radiance $L_i(\mathbf{x} \leftarrow \omega)$. The subscript i in L_i distinguishes the *incident* radiance from *emitted* radiance L_e, which is introduced later.

Transmittance $T(\mathbf{x}_1 \leftrightarrow \mathbf{x}_2)$ describes the fraction of light that begins its journey at point $\mathbf{x}_1$ and reaches point $\mathbf{x}_2$ without scattering:

$$T(\mathbf{x}_1 \leftrightarrow \mathbf{x}_2) = e^{-\int_{\mathbf{x}_1}^{\mathbf{x}_2} \mu_t(\mathbf{x}) \, d\mathbf{x}} \qquad (2)$$

The extinction coefficient $\mu_t(\mathbf{x}) = \mu_s(\mathbf{x}) + \mu_a(\mathbf{x})$ probabilistically describes the fraction of light that is either scattered (μ_s) or absorbed (μ_a) per *unit distance*. It is therefore a number that depends on the extent of the data set. This coefficient is used to define the optical thickness of a medium, i.e., setting a small value makes a voxel invisible, while a high value makes it appear opaque.

Scattering Coefficient $\mu_s(\mathbf{x})$ determines the visual appearance of a location in a participating medium by defining the fraction of photons that scatter per unit distance. In direct volume rendering, it is common to apply phenomenological scattering models [JSYR14]:

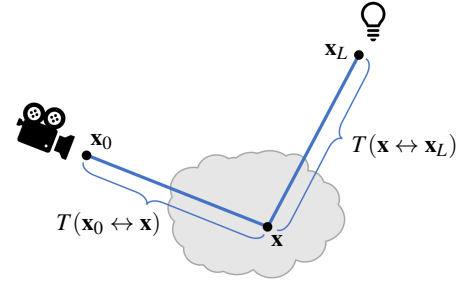$$\mu_s(\mathbf{x}) = \mu_t(\mathbf{x}) \cdot \mathbf{c}(\mathbf{x}) \qquad (3)$$

Figure 2: *Single-scattering light transport accumulates all incoming light that gets reflected along the view ray towards the camera. The camera is place at $\mathbf{x}_0$, the light source at $\mathbf{x}_L$ and one exemplary scattering point $\mathbf{x}$ is shown. The fraction of light arriving from the light source at the scattering point is the transmittance $T(\mathbf{x} \leftrightarrow \mathbf{x}_L)$ and the fraction of light reaching the camera from the scattering point is denoted by the transmittance $T(\mathbf{x}_0 \leftrightarrow \mathbf{x})$.*

where the color $\mathbf{c}(\mathbf{x}) : \mathcal{D} \to [0,1]^3$ specifies per color channel the fraction of extinct light that is lost due to scattering (μ_s/μ_t). A low value means that light is absorbed (black), while a high value indicates scattering.

Incident Radiance $L_i(\mathbf{x} \leftarrow \omega)$ determines the amount of light incoming at point $\mathbf{x}$ and that is scattered in direction ω. We apply a single-scattering model, which directly connects the scattering point $\mathbf{x}$ to a light source at point $\mathbf{x}_L$ that emits a radiance L_e:

$$L_i(\mathbf{x} \leftarrow \omega) = f(\mathbf{x}, \omega_L \to \omega) \cdot T(\mathbf{x} \leftrightarrow \mathbf{x}_L) \cdot L_e \qquad (4)$$

where the light ray transmittance $T(\mathbf{x} \leftrightarrow \mathbf{x}_L)$ denotes the fraction of light reaching the scattering point from the light source. Thus, as shown in Fig. 2, two transmittance terms are considered: one from the camera in Eq. (1) and one from the light source in Eq. (4). The phase function $f(\mathbf{x}, \omega_L \to \omega)$ describes how much incoming light from direction ω_L is being scattered into the outgoing direction ω. We apply a shading similar to Ament et al. [AZD17], which contains in our case normalized Blinn-Phong shading $f_s(\mathbf{x}, \omega_L, \omega)$ and Henyey-Greenstein forward-scattering $f_v(\mathbf{x}, \omega_L, \omega)$ (with anisotropy $a = 0.6$) [HG41].

$$
\begin{aligned}
f(\mathbf{x}, \omega_L \to \omega) = & \; k_v(\mathbf{x}) \cdot f_v(\mathbf{x}, \omega_L, \omega) \\
& + k_s(\mathbf{x}) \cdot f_s(\mathbf{x}, \omega_L, \omega) \cdot (\omega_L \cdot \mathbf{n}(\mathbf{x}))_+
\end{aligned}
\qquad (5)
$$

with the normal $\mathbf{n}(\mathbf{x}) = \frac{\nabla \bar{\mu}(\mathbf{x})}{\|\nabla \bar{\mu}(\mathbf{x})\|}$ and a clamping of negative values to zero $(\cdot)_+$. Following Ament et al. [AZD17], we later set $k_v(\mathbf{x}) = 1$ and $k_s(\mathbf{x}) = \|\nabla \bar{\mu}(\mathbf{x})\|$.

Differentiable Volume Rendering. Inverse rendering refers to the reconstruction of scene parameters from a given image [KBM*20, RRN*20], which can be phrased as an optimization problem to which iterative gradient-based solvers can be applied. It is therefore necessary to determine how much a perturbation of a scene parameter affects the final image, i.e., a gradient of the output image with respect to the scene parameters is needed. In the field of direct volume rendering, reverse automatic differentiation with analytic inversion of blending functions has been proposed [WW22] and generative networks have been trained [BLL19]. For global illumination simulations, differentiable Monte Carlo renderers [NDVZJ19] have been

proposed that approximate gradients [ZSGJ21] or calculate them in two passes [VSJ21]. To address visibility discontinuities, edge sampling towards silhouettes [LADL18], a bidirectional generation of boundary paths that connect sensors and emitters [ZMY*20], and a transformation from boundary to area integrals [BLD20] have been proposed. The path space formulation has recently been extended to optimize the boundary geometry of translucent volumetric objects [YZM*23]. In this paper, we model single-scattering light transport on differentiable scalar fields $s(\mathbf{x})$, which allows us to compute the necessary derivatives symbolically.

3. Method

In volume visualization, visibility optimization of relevant information is a long-standing problem [VG05, CWM*09, CM11, AZD17]. In this paper, we introduce a concurrent optimization of the extinction field and the camera viewpoint to clear the view onto relevant structures. To assess the visibility of relevant information, we measure the transmittance of relevant objects in the scene directly. In the following section, we first formally introduce the variational formulation that is used to derive the required partial derivatives.

3.1. Problem Statement

Given is a 3D scalar field $s(\mathbf{x}) : \mathcal{D} \to \mathbb{R}$ in the spatial domain $\mathcal{D} \subseteq \mathbb{R}^3$ that we want to visualize. Further, a differentiable 3D importance field $g(\mathbf{x}) : \mathcal{D} \to [0,1]$ defines how important each region in the domain is. Lastly, transfer functions are provided to compute color $\mathbf{c}(\mathbf{x}) : \mathcal{D} \to [0,1]^3$ and base extinction $\bar{\mu}_t(\mathbf{x}) : \mathcal{D} \to [0,\widetilde{\mu}_t]$ from the scalar field $s(\mathbf{x})$ with $\widetilde{\mu}_t$ being the majorant extinction, i.e., the highest possible extinction coefficient. The base extinction $\bar{\mu}_t(\mathbf{x})$ is the extinction coefficient that our optimization will strive towards if a voxel does not occlude any relevant information. Thus, if no notion of relevance is given, our approach becomes a standard direct volume rendering with $\bar{\mu}_t(\mathbf{x})$ being the extinction coefficient.

Our goal is to find a new extinction field $\mu_t(\mathbf{x}) : \mathcal{D} \to [0,\widetilde{\mu}_t]$ as well as a perspective camera transformation $\mathbf{g}_V(\mathbf{x}) : \mathcal{D} \to \mathcal{Y}$, which maps from world space $\mathcal{D}$ to view space $\mathcal{Y}$, such that both together produce a direct volume visualization with single-scattering light transport in which the important regions are shown. This entails the adaption of the extinction to reveal important structures and an optimization of the camera position.

Differentiable Volume Rendering. In order to optimize for the unknown functions $\mu_t(\mathbf{x})$ and $\mathbf{g}_V(\mathbf{x})$, a differentiable volume renderer is needed. Inserting Eqs. (3)–(4) into the volume rendering equation in Eq. (1) gives a single-scattering radiance integral, which is differentiable and contains both unknowns, the extinction coefficient $\mu_t(\mathbf{x})$ and the camera transformation $\mathbf{g}_V(\mathbf{x})$:

$$L = \int_{\mathbf{x}_0}^{\mathbf{x}_1} \mu_t(\mathbf{x}) \cdot \mathbf{c}(\mathbf{x}) \cdot T_V(\mathbf{g}_V(\mathbf{x})) \cdot f(\mathbf{x}, \omega_L \to \omega) \cdot T_L(\mathbf{g}_L(\mathbf{x})) \cdot L_e \, d\mathbf{x} \tag{6}$$

In this formulation, $\mathbf{x}_0$ and $\mathbf{x}_1$ are the entry and exit of a view ray. The maps $\mathbf{g}_V(\mathbf{x}) : \mathcal{D} \to \mathcal{Y}$ and $\mathbf{g}_L(\mathbf{x}) : \mathcal{D} \to \mathcal{Z}$ define perspective transformations from world space $\mathcal{D}$ into view space $\mathcal{Y}$ and light

space $\mathcal{Z}$, respectively. Later, we introduce the view space transmittance $T_V(\mathbf{y}) = T(\mathbf{x}_0 \leftrightarrow \mathbf{g}_V^{-1}(\mathbf{y}))$ and light space transmittance $T_L(\mathbf{z}) = T(\mathbf{g}_L^{-1}(\mathbf{y}) \leftrightarrow \mathbf{x}_L)$, which depend on the extinction $\mu_t(\mathbf{x})$.

Variational Approach. We express the search for the optimal extinction coefficient $\mu_t(\mathbf{x})$ and the camera transformation $\mathbf{g}_V(\mathbf{x})$ as variational minimization problem. Variational calculus is a mathematical toolbox that allows for the optimization of *functionals*, i.e., functions which receive other functions as input [GF63]. The functional $\mathcal{F}$ that we will optimize later integrates a Lagrangian $\mathcal{L}$ over the world space domain $\mathcal{D}$:

$$\mathcal{F}[\mu_t(\mathbf{x}), \mathbf{g}_V(\mathbf{x})] = \int_{\mathcal{D}} \mathcal{L}\left(\mathbf{x}, \mu_t(\mathbf{x}), \mathbf{g}_V(\mathbf{x}), \frac{\partial \mu_t(\mathbf{x})}{\partial \mathbf{x}}, \frac{\partial \mathbf{g}_V(\mathbf{x})}{\partial \mathbf{x}}\right) d\mathbf{x} \tag{7}$$

In general, a necessary condition for the minimium of such functional is given by the Euler-Lagrange equations [GF63]:

$$\frac{\delta \mathcal{F}}{\delta \mu_t(\mathbf{x})} = \frac{\partial \mathcal{L}}{\partial \mu_t(\mathbf{x})} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left(\frac{\partial \mathcal{L}}{\partial \frac{\partial \mu_t(\mathbf{x})}{\partial x_j}} \right) = 0 \tag{8}$$

$$\frac{\delta \mathcal{F}}{\delta \mathbf{g}_V(\mathbf{x})} = \frac{\partial \mathcal{L}}{\partial \mathbf{g}_V(\mathbf{x})} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left(\frac{\partial \mathcal{L}}{\partial \frac{\partial \mathbf{g}_V(\mathbf{x})}{\partial x_j}} \right) = \mathbf{0} \tag{9}$$

with $\mathbf{x} = (x_1, x_2, x_3)$. These differential equations hold for every point in the domain. Note that $\frac{\delta \mathcal{F}}{\delta \mu_t(\mathbf{x})}$ in Eq. (8) is scalar-valued and that $\frac{\delta \mathcal{F}}{\delta \mathbf{g}_V(\mathbf{x})}$ in Eq. (9) is vector-valued. In the following, we describe the ingredients needed to define the functional that we minimize.

3.2. Camera Transformation

The first unknown of our optimization is the camera transformation $\mathbf{g}_V(\mathbf{x})$. In the following, we describe its degrees of freedom and provide partial derivatives with respect to the degrees of freedoms.

Degrees of Freedom. For the efficient generation of view rays, a camera transformation is often defined by three basis vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ and by a camera location $\mathbf{o} \in \mathcal{D}$ in world space, as shown in Fig. 3. Instead of optimizing these vectors, we set the camera to a fixed location and apply a world space rotation $\mathbf{R}(\alpha_1, \alpha_2) \in SO(3)$ to the scene, which contains the unknown azimuthal and polar angles α_1 and α_2, keeping the camera upright:

$$\mathbf{R}(\alpha_1, \alpha_2) = \begin{pmatrix} \cos(\alpha_1) & \sin(\alpha_1)\sin(\alpha_2) & \sin(\alpha_1)\cos(\alpha_2) \\ 0 & \cos(\alpha_2) & -\sin(\alpha_2) \\ -\sin(\alpha_1) & \cos(\alpha_1)\sin(\alpha_2) & \cos(\alpha_1)\cos(\alpha_2) \end{pmatrix} \tag{10}$$

Thus, the camera position is restricted to a sphere that encloses the data set. Without loss of generality, the pivot point around which the camera rotates is set to the origin of the world coordinate system. During volume rendering, view rays are cast through the image coordinates $(y_1, y_2) \in [-1,1]^2$ via:

$$\mathbf{x}(t) = \mathbf{R} \cdot \left(\frac{\mathbf{u} \cdot y_1 + \mathbf{v} \cdot y_2 + \mathbf{w}}{\|\mathbf{u} \cdot y_1 + \mathbf{v} \cdot y_2 + \mathbf{w}\|} t + \mathbf{o} \right) \tag{11}$$
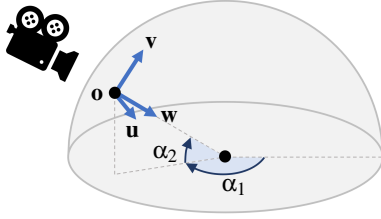
Figure 3: *The camera orientation is defined by a* **u**, **v**, **w** *frame, and the camera position is* **o**. *When optimizing for the optimal viewpoint, we rotate the world space by the azimuthal and polar angles* α_1, α_2.

Perspective Mapping. The perspective mapping $\mathbf{g}_V(\mathbf{x})$ transforms a world space coordinate $\mathbf{x} \in \mathcal{D}$ to the camera space $\mathbf{y} \in \mathcal{Y} = [-1,1] \times [-1,1] \times [0,1]$ as follows:

$$\mathbf{g}_V(\mathbf{x}) = \left((\mathbf{v} \times \mathbf{w})^{\mathrm{T}} \mathbf{k}, \quad (\mathbf{w} \times \mathbf{u})^{\mathrm{T}} \mathbf{k}, \quad \frac{(\mathbf{u} \times \mathbf{v})^{\mathrm{T}} (\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o} - \mathbf{w} \cdot z_n)}{(\mathbf{u} \times \mathbf{v})^{\mathrm{T}} \mathbf{w} \cdot (z_f - z_n)} \right)^{\mathrm{T}}$$
(12)

with $\mathbf{k} = \frac{\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o}}{(\mathbf{u} \times \mathbf{v})^{\mathrm{T}} (\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o})}$. Under this perspective mapping, the world space depth range $[z_n, z_f]$ along the $\mathbf{w}$ axis is mapped to a linear depth in $[0,1]$. The inverse transformation $\mathbf{g}_V^{-1}(\mathbf{y})$ is:

$$\mathbf{g}_V^{-1}(\mathbf{y}) = \mathbf{R} \cdot \left((\mathbf{u} \cdot y_1 + \mathbf{v} \cdot y_2 + \mathbf{w}) \cdot (z_n + (z_f - z_n) \cdot y_3) + \mathbf{o} \right) \quad (13)$$

Similarly, the transformation to the light space is given by a perspective mapping $\mathbf{g}_L(\mathbf{x}) : \mathcal{D} \to \mathcal{Z}$. We chose to model a point light, although directional lights would be easily supported, as well.

Derivatives. The partial derivative of the transformation $\mathbf{g}_V(\mathbf{x})$ with respect to the rotation angles α_i is computed symbolically:

$$\frac{\partial \mathbf{g}_V(\mathbf{x})}{\partial \alpha_i} = \left((\mathbf{v} \times \mathbf{w})^{\mathrm{T}} \frac{\partial \mathbf{k}}{\partial \alpha_i}, \quad (\mathbf{w} \times \mathbf{u})^{\mathrm{T}} \frac{\partial \mathbf{k}}{\partial \alpha_i}, \quad \frac{(\mathbf{u} \times \mathbf{v})(\frac{\partial \mathbf{R}}{\partial \alpha_i}^{\mathrm{T}} \mathbf{x})}{(\mathbf{u} \times \mathbf{v})^{\mathrm{T}} \mathbf{w} \cdot (z_f - z_n))} \right)$$
(14)

with $\frac{\partial \mathbf{k}}{\partial \alpha_i} = \frac{(\frac{\partial \mathbf{R}}{\partial \alpha_i}^{\mathrm{T}} \mathbf{x}) \cdot (\mathbf{u} \times \mathbf{v})^{\mathrm{T}} (\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o}) - (\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o}) \cdot (\mathbf{u} \times \mathbf{v})^{\mathrm{T}} (\frac{\partial \mathbf{R}}{\partial \alpha_i}^{\mathrm{T}} \mathbf{x})}{\|(\mathbf{u} \times \mathbf{v})^{\mathrm{T}} (\mathbf{R}^{\mathrm{T}} \mathbf{x} - \mathbf{o})\|^2}$. Using this derivative, we can express how much the camera transformation changes when varying the azimuthal and polar angle of the rotation.

3.3. View Space and Light Space Quantities

The second unknown is the extinction $\mu_t(\mathbf{x})$, which appears in the transmittance in Eq. (2). The volume rendering equation in Eq. (6) contains two transmittance terms: one is measured along view rays (T_V), the other along light rays (T_L). To accelerate computations we discretize both onto perspective grids, as shown in Fig. 4.

Transmittance. We describe the view space transmittance $T_V(\mathbf{y}) : \mathcal{Y} \to [0,1]$ and the light space transmittance $T_L(\mathbf{z}) : \mathcal{Z} \to [0,1]$ in their respective view space and light space coordinates using the view space transformation $\mathbf{g}_V(\mathbf{x}) : \mathcal{D} \to \mathcal{Y}$ and the light space transformation $\mathbf{g}_L(\mathbf{x}) : \mathcal{D} \to \mathcal{Z}$:

$$T_V(\mathbf{y}) = e^{-\int_{y_0}^{\mathbf{y}} \mu_t(\mathbf{g}_V^{-1}(\mathbf{y}')) \, \mathrm{d}\mathbf{y}'} \quad (15)$$

$$T_L(\mathbf{z}) = e^{-\int_{z_0}^{\mathbf{z}} \mu_t(\mathbf{g}_L^{-1}(\mathbf{z}')) \, \mathrm{d}\mathbf{z}'} \quad (16)$$
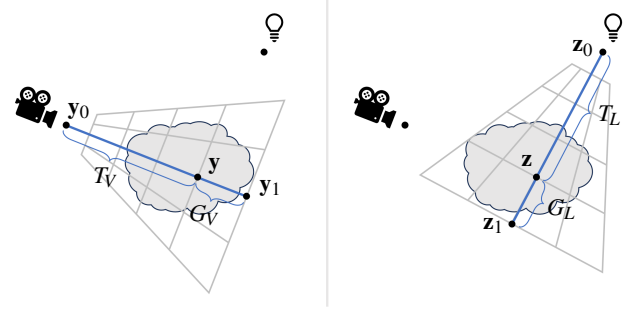

Figure 4: *To efficiently sample the transmittance up to a certain point and the amount of visible information behind that point, we discretize the transmittance and visibility integrals onto perspective grids, shown here for the view space (left) and the light space (right).*

with the camera being placed at $\mathbf{y}_0$ and the light source being placed at $\mathbf{z}_0$. The fields are discretized onto perspective grids with a resolution of $X \times Y \times Z$ by looping once over each ray:

$$T_V(\mathbf{y}_{i,j,k}) \approx \prod_{k'=1}^{k-1} e^{-\mu_t(\mathbf{g}_V^{-1}(\mathbf{y}_{i,j,k'})) \cdot h} \quad (17)$$

$$T_L(\mathbf{z}_{i,j,k}) \approx \prod_{k'=1}^{k-1} e^{-\mu_t(\mathbf{g}_L^{-1}(\mathbf{z}_{i,j,k'})) \cdot h} \quad (18)$$

where h is the grid spacing in view direction. Afterwards, the transmittance can be looked up in world space by transforming to the grid coordinates and by trilinearly interpolating the transmittance value. For light rays, this is known as a shadow volume [AZD17]. In computer graphics, the exponential is often first-order Taylor approximated, i.e., $e^{-\mu_t(\mathbf{x}_i) \cdot h} = e^{-\alpha_i} \approx 1 - \alpha_i$, which results in the common alpha blending equations [BM08].

Visibility Integrals. How much an individual voxel strives to be seen depends on the amount of relevant visible information behind it. Visibility is thereby characterized by high extinction and transmittance. Thus, we compute visibility-weighted importance integrals *behind* a given point in view space (G_V) and light space (G_L) as:

$$G_V(\mathbf{y}) = \int_{\mathbf{y}}^{\mathbf{y}_1} g(\mathbf{g}_V^{-1}(\mathbf{y}')) \cdot \mu_t(\mathbf{g}_V^{-1}(\mathbf{y}')) \cdot T_V(\mathbf{y}') \, \mathrm{d}\mathbf{y}' \quad (19)$$

$$G_L(\mathbf{z}) = \int_{\mathbf{z}}^{\mathbf{z}_1} g(\mathbf{g}_L^{-1}(\mathbf{z}')) \cdot \mu_t(\mathbf{g}_V^{-1}(\mathbf{z}')) \cdot T_L(\mathbf{z}') \, \mathrm{d}\mathbf{z}' \quad (20)$$

which are computed on a $X \times Y \times Z$ grid by traversing the rays back-to-front, whenever the camera changes its position.

$$G_V(\mathbf{y}_{i,j,k}) \approx \sum_{k'=k+1}^{Z} g(\mathbf{g}_V^{-1}(\mathbf{y}_{i,j,k'})) \cdot \mu_t(\mathbf{g}_V^{-1}(\mathbf{y}_{i,j,k'})) \cdot T_V(\mathbf{y}_{i,j,k'}) \cdot h$$

$$G_L(\mathbf{z}_{i,j,k}) \approx \sum_{k'=k+1}^{Z} g(\mathbf{g}_L^{-1}(\mathbf{z}'_{i,j,k'})) \cdot \mu_t(\mathbf{g}_L^{-1}(\mathbf{z}_{i,j,k'})) \cdot T_L(\mathbf{z}_{i,j,k'}) \cdot h$$

with h being again the step size. With this, the amount of occluded relevant information can be looked up at any world space position.

3.4. Optimization

Now we have all the ingredients needed to define the functional that we aim to minimize. Afterwards, we provide its functional derivatives and a gradient descent update for the optimization.

Energy. Similar to the opacity optimization energy [AZD17, GTG17], we decompose our variational energy into multiple terms:

$$\mathcal{F}[\mu_t(\mathbf{x}), \mathbf{g}_V(\mathbf{x})] = \mathcal{F}_p + \mathcal{F}_q + \mathcal{F}_r \tag{21}$$

with each individual term serving a specific task:

$$\mathcal{F}_p = \frac{p}{2} \int_{\mathcal{D}} (\mu_t(\mathbf{x}) - \overline{\mu}_t(\mathbf{x}))^2 \, d\mathbf{x} \tag{22}$$

$$\mathcal{F}_q = -q \int_{\mathcal{D}} g(\mathbf{x}) \cdot \mu_t(\mathbf{x}) \cdot T_V(\mathbf{g}_V(\mathbf{x})) \, d\mathbf{x} \tag{23}$$

$$\mathcal{F}_r = -r \int_{\mathcal{D}} g(\mathbf{x}) \cdot \mu_t(\mathbf{x}) \cdot T_L(\mathbf{g}_L(\mathbf{x})) \, d\mathbf{x} \tag{24}$$

where the extinction $\mu_t(\mathbf{x})$ is subject to the constraint:

$$0 \le \mu_t(\mathbf{x}) \le \widetilde{\mu}_t : \forall \mathbf{x} \in \mathcal{D} \tag{25}$$

$\mathcal{F}_p$ in Eq. (22) is a regularization term that – in the absence of all other terms – makes the unknown extinction $\mu_t(\mathbf{x})$ become equal to the given base extinction $\overline{\mu}_t(\mathbf{x})$. Thus, if no extinction optimization is applied, we obtain a regular volume rendering.

$\mathcal{F}_q$ in Eq. (23) maximizes the extinction $\mu_t(\mathbf{x})$ of important structures ($g(\mathbf{x})$ is high) such that they scatter more light and the transmittance $T_V(\mathbf{x})$ is kept high, which makes sure that no object occludes the view up to $\mathbf{x}$. Phrased as a minimization, the term is negated.

$\mathcal{F}_r$ in Eq. (24) does the same from the perspective of the light source. This ensures that light is reaching the relevant structures.

Since $\mu_t(\mathbf{x})$ in Eqs. (23)–(24) is unbounded, the constraint in Eq. (25) is added. We set the upper bound $\widetilde{\mu}_t$ to the maximum extinction that is present in the transfer function that produced $\overline{\mu}_t(\mathbf{x})$. Thereby, the volume does not become more optically dense than the base volume rendering.

We set the same energy weights in all data sets $p = 10^{-3}$, $q = 10$, $r = 1$ unless mentioned otherwise. A parameter study is presented in the supplemental material, where other choices are demonstrated.

With our approach it is possible to optimize for a single extinction field $\mu_t(\mathbf{x})$ that fulfills its visibility constraints from both the camera and light perspective. Alternatively, we may also optimize for two separate extinction fields: one for the camera constraint, and one for the light constraint. The latter approach follows Ament et al. [AZD17], which, however, may lead to an inconsistent lighting. Later, in Section 4.2, we demonstrate and compare both approaches.

Partial Derivatives. Differentiation of Eqs. (22)–(24) according to the Euler-Lagrange equation in Eq. (8) with respect to the extinction $\mu_t(\mathbf{x})$ results in the following functional derivatives:

$$\frac{\delta \mathcal{F}_p}{\delta \mu_t} = p \cdot (\mu_t(\mathbf{x}) - \overline{\mu}_t(\mathbf{x})) \tag{26}$$

$$\frac{\delta \mathcal{F}_q}{\delta \mu_t} = q \cdot (G_V(\mathbf{g}_V(\mathbf{x})) - g(\mathbf{x}) \cdot T_V(\mathbf{g}_V(\mathbf{x}))) \tag{27}$$

$$\frac{\delta \mathcal{F}_r}{\delta \mu_t} = r \cdot (G_L(\mathbf{g}_L(\mathbf{x})) - g(\mathbf{x}) \cdot T_L(\mathbf{g}_L(\mathbf{x}))) \tag{28}$$

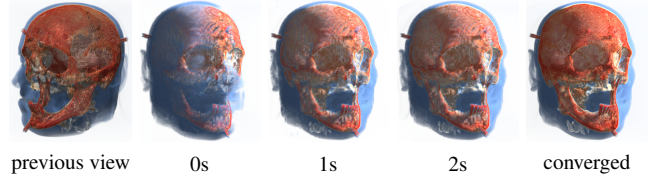| previous view | 0s | 1s | 2s | converged |

Figure 5: *To reach interactivity, the extinction optimization is carried out over the course of multiple frames. Here, the delay of the visibility adjustment after interactive camera adjustment is shown. The grid resolution is adapted automatically to support interactivity.*

Since only $\mathcal{F}_q$ depends on the camera transformation $\mathbf{g}_V(\mathbf{x})$, and thus the angles α_i, there is only one functional derivative for the camera optimization:

$$\frac{\delta \mathcal{F}_q}{\delta \alpha_i} = -q \int_{\mathcal{D}} g(\mathbf{x}) \cdot \mu_t(\mathbf{x}) \cdot \frac{\delta T_V(\mathbf{g}_V(\mathbf{x}))}{\delta \alpha_i} \, d\mathbf{x} \tag{29}$$

Using $\mu_t(\mathbf{x})$ from the extinction optimization in Eq. (29) results in a coupled optimization, for which the convergence requires careful balancing of learning rates. Using a view-independent proxy $\mu_t(\mathbf{x}) := \widetilde{\mu}_t \cdot g(\mathbf{x})$ in Eq. (29) instead is more stable and empirically gave similar optimal camera positions. The functional derivatives are derived in the additional material. We estimate the energy derivative $\frac{\delta \mathcal{F}_q}{\delta \alpha_i}$ with second-order central differences.

Gradient Descent. To compute the extinction field $\mu_t(\mathbf{x})$ that fulfills the Euler-Lagrange equation in Eq. (8), we discretize the field onto a world space grid with extinction values $\mu_{t,i}$, which are optimized via gradient descent, starting at an initial guess $\mu_{t,i}^{(0)} := \overline{\mu}_t(\mathbf{x}_i)$. Thus, if the optimization is disabled ($q = r = 0$), we already see the regular volume rendering. Likewise, to find the camera rotation angles α_i that fulfill the Euler-Lagrange equation in Eq. (9), we apply the chain rule using Eq. (14), integrate contributions from all points in the domain, and perform a gradient descent with step size h, starting from an initial guess $\alpha_i^{(0)} = 0$:

$$\mu_{t,i}^{(n+1)} \approx \mu_{t,i}^{(n)} - h \cdot \frac{\delta \mathcal{F}}{\delta \mu_t(\mathbf{x})} \tag{30}$$

$$\alpha_i^{(n+1)} \approx \alpha_i^{(n)} - h \cdot \frac{\delta \mathcal{F}_q}{\delta \alpha_i} \tag{31}$$

To satisfy the constraint in Eq. (25), the extinctions $\mu_{t,i}$ are clamped to the valid range after each gradient descent step. Since our functional is non-linear, the result depends on the initial condition, which is discussed later in Section 4.6.

3.5. Visualization System

We implemented our extinction and viewpoint optimization using CUDA on the GPU, setting ourselves the following requirements:

(R1): Support camera navigation during extinction optimization.
(R2): Provide feedback on how much relevant information is seen.

Interactivity. When disabling the viewpoint optimization, the user is able to change the camera transformation interactively in order to explore the data set. To reach interactive frame rates (R1), the iteration in Eq. (30) is deferred over multiple frames. Thus, during
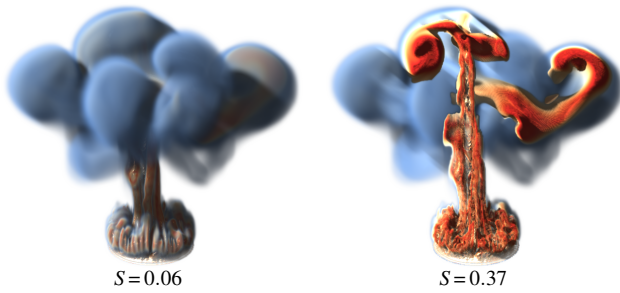
$S = 0.06$ $S = 0.37$

Figure 6: *To judge the visibility of relevant structures, we introduce a visibility score S. Here, we see the improvement achieved over direct volume rendering (left) compared to our method (right).*

interactive camera navigation the optimization needs a few seconds to adjust the visibility to the new viewpoint, see Fig. 5, which was similarly the case in early optimization algorithms [GRT13].

Feedback. Our visibility optimizations incorporate the transmittance of each object. To convey how much relevant information is seen (R2), we accumulate a visibility score $S \in \mathbb{R}_+$ over the domain:

$$S = \int_{\mathcal{D}} g(\mathbf{x}) \cdot \mu_t(\mathbf{x}) \cdot T_V(\mathbf{g}_V(\mathbf{x})) \, d\mathbf{x} \tag{32}$$

The score is high if a voxel is important ($g(\mathbf{x})$ is high), the voxel is scattering light ($\mu_t(\mathbf{x})$ is high), and no other objects occluded the view ($T_V(\mathbf{g}_V(\mathbf{x}))$ is high). Fig. 6 shows quantitative measurements of the visibility score for different views.

Multi-Resolution Solver on GPU. To meet the above requirements, we implemented the gradient-based optimization with a multi-resolution update of the grids of Section 3.3 on the GPU. Following 32 iterations on quarter resolution, the result is upscaled to half resolution, and after 64 iterations, the result is upscaled to the full resolution. The number of iterations per resolution level were chosen empirically. Each optimization step includes the following:

1. calculate derivatives of energies with respect to the unknowns
2. gradient step on extinction field and on camera parameters
3. recalculate the view space and light space grids, cf. Section 3.3
4. calculate energies and metrics after the optimization step

The gradient descent steps are carried out with adaptive estimates of lower-order moments by using the Adam method [KB17]. The lower resolution grids that we start the optimization from do not need additional memory, since they are stored at every second or fourth grid point in the full resolution grid. Upscaling to a higher resolution thereby becomes a simple interpolation problem.

4. Results

In the following, we apply our algorithm to a number of data sets from fluid dynamics, geophysics, and medicine. After introducing the data sets, we discuss the separate and joint extinction optimization, we compare with previous work, study the impact of the grid resolutions, provide performance measurements, and discuss limitations. We refer to the supplemental material for an informal user study, a parameter study of the weights q and r, as well as additional comparisons with related work for a different camera view.

4.1. Test Data

The ROTATING MIXER contains a numerical simulation of a liquid in a cylindrical container that is stirred into motion by three rotating paddles. We show the vorticity [GT18] of this flow, highlighting regions with exceptionally high vortical motion.

The EARTH MANTLE simulation resolves the sinking of cold, dense material from the crust, and the rising of hot plumes from the core [GBT21]. We visualize the temperature anomaly, which shows the difference in temperature to the average temperature at a certain depth and we highlight cold slabs with high importance.

The VISIBLE HUMAN data set (version 2.0) contains high-resolution CT and MR scans of the head and neck of a male human [RHGJ03]. We visualize the CT data set and assigned bone structures a high importance.

The HEPTANE FLAME data set contains a single time step from a combustion simulation of a jet of heptane gas. The importance is set to reproduce the visualizations of Ament et al. [AZD17].

4.2. Separate vs Joint Extinction Optimization

When minimizing Eqs. (22)–(24), we may either solve for one extinction grid $\mu_t(\mathbf{x})$ that clears the view from the camera and the light, or we solve for two separate extinction grids by once setting $q = 0$ and once setting $r = 0$. A joint solver has a lighting that is consistent with the visible extinction field. In contrast, the separate solution displays the context around the region of interest independent of the illumination direction. Results for both approaches can be seen in Fig. 7, where the nose and forehead have been removed in the joint optimization to clear the path for the light. Unlike Ament et al. [AZD17], we have the freedom to choose between the two options, while they had to solve for both views separately.

4.3. Comparisons

Extinction Optimization. We compare our approach with three extinction optimization algorithms in Fig. 8. Methods that do not optimize the shadows are compared without rendering shadows, i.e., $T_L = 1$. Viola et al. [VKG04] segmented the domain into multiple objects and determined per object a sparseness level, which allows for a smooth transition from fully transparent to fully opaque object representations. We compare with their cylindrical maximum importance projection, which determines along the ray the most relevant object and then clips everything in front of this object. In our reimplementation, we consider objects to be determined by the $g(\mathbf{x}) = 0.5$ level set, and start the regular ray marching at the first level set intersection of the object with highest value along the view ray. Being based on a maximum intensity projection, the algorithm cannot retain depth information. This is particularly noticeable in the VISIBLE HUMAN data set. Marchesin et al. [MMD10] computed an importance-weighted color average of samples along the view ray. Accordingly, the visibility score was equal-weighted along the ray. We determined the transmittance for the blending with the background by calculating the transmittance of the base extinction field $\bar{\mu}(\mathbf{x})$. Due to the order-independence of the compositing, depth information cannot be maintained. While adding detail in some regions, less relevant structures receive too much weight, which is
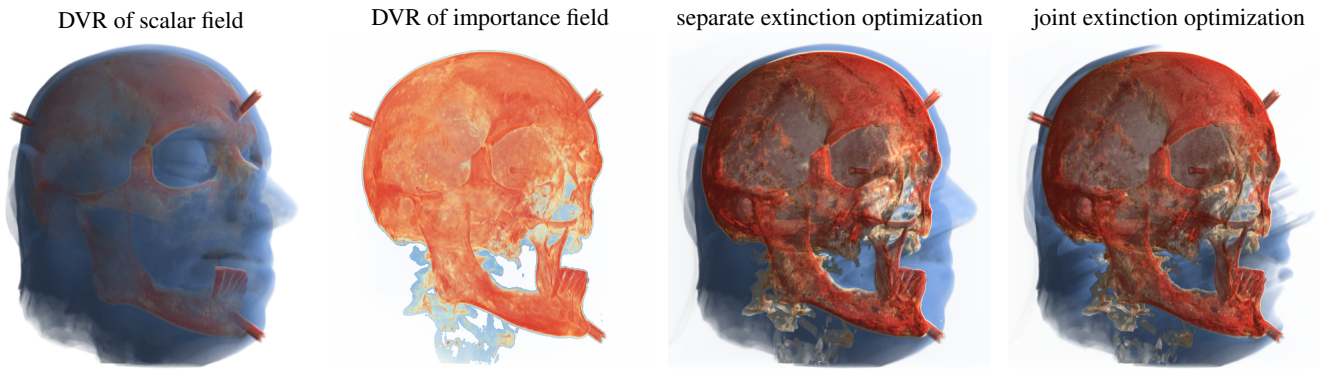
Paul Himmler & Tobias Günther / Transmittance-based Extinction and Viewpoint Optimization

DVR of scalar field	DVR of importance field	separate extinction optimization	joint extinction optimization

Figure 7: *With our approach we can either optimize for two separate extinction fields for the view from the camera and from the light, which preserves more context, or the two views are optimized jointly, which produces a more consistent shadowing, but removes nose and forehead.*
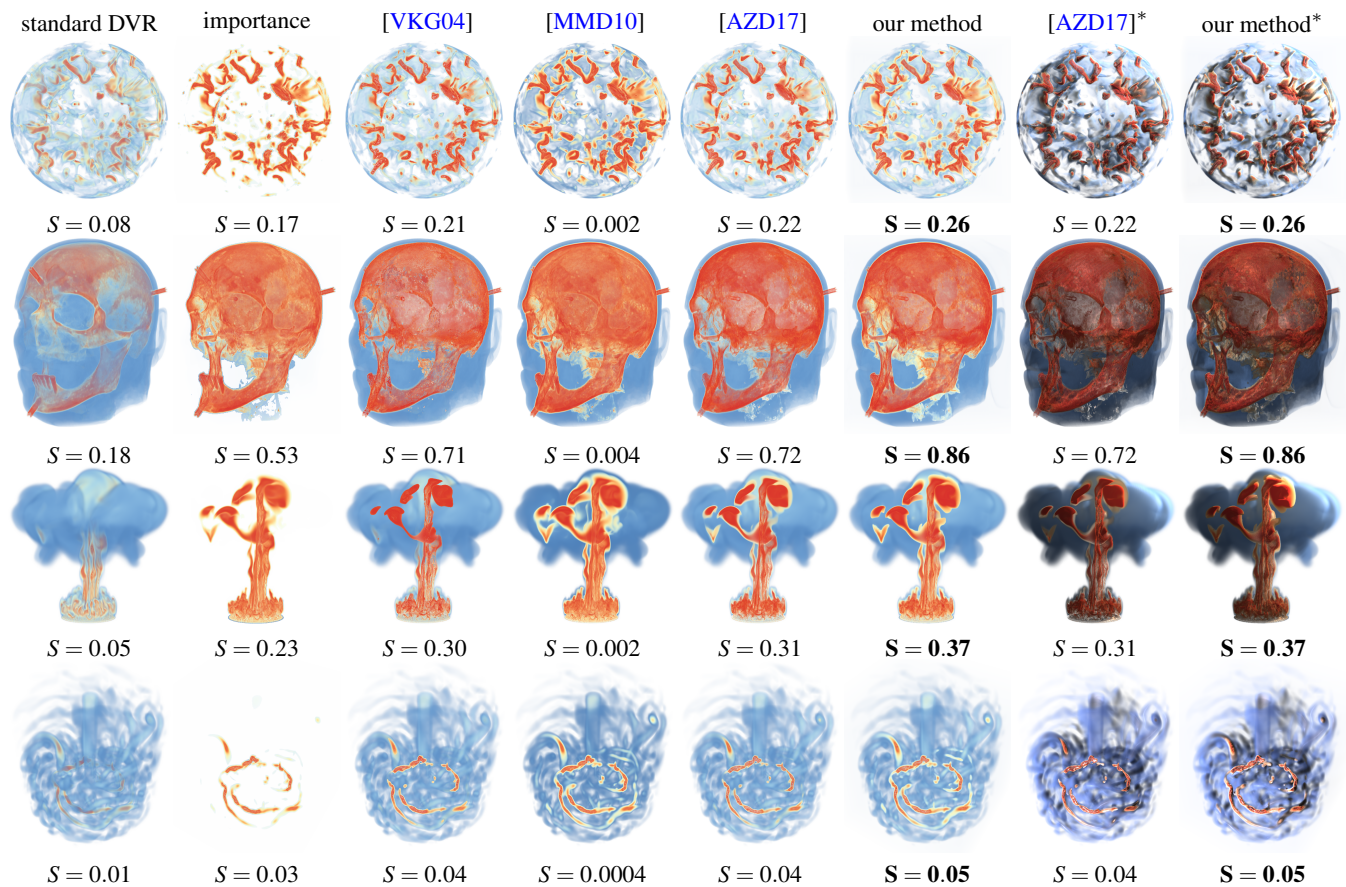
standard DVR	importance	[VKG04]	[MMD10]	[AZD17]	our method	[AZD17]*	our method*

$S = 0.08$	$S = 0.17$	$S = 0.21$	$S = 0.002$	$S = 0.22$	$\mathbf{S = 0.26}$	$S = 0.22$	$\mathbf{S = 0.26}$
$S = 0.18$	$S = 0.53$	$S = 0.71$	$S = 0.004$	$S = 0.72$	$\mathbf{S = 0.86}$	$S = 0.72$	$\mathbf{S = 0.86}$
$S = 0.05$	$S = 0.23$	$S = 0.30$	$S = 0.002$	$S = 0.31$	$\mathbf{S = 0.37}$	$S = 0.31$	$\mathbf{S = 0.37}$
$S = 0.01$	$S = 0.03$	$S = 0.04$	$S = 0.0004$	$S = 0.04$	$\mathbf{S = 0.05}$	$S = 0.04$	$\mathbf{S = 0.05}$

Figure 8: *Comparison of extinction optimization methods on the* EARTH MANTLE, VISIBLE HUMAN, HEPTANE FLAME, *and* ROTATING MIXER *data set. Viola et al. [VKG04] and Marchesin et al. [MMD10] did not include shadow optimizations and are therefore compared without shadowing. Two additional columns (marked by *) compare Ament et al. [AZD17] with our method. Previous approaches [VKG04, MMD10] provide limited depth perception. Ament et al. [AZD17] has global parameters, making it difficult to bring out differently important structures throughout the domain. Some parts might even be lost entirely. In contrast, our method maximizes the visibility successfully everywhere.*

particularly noticeable in the HEPTANE FLAME data set, as well as in some parts of the EARTH MANTLE. Lastly, we compare with the extinction optimized volume illumination by Ament et al. [AZD17], which adjusts the extinction of each sample point by considering the integral of squared importance in front and behind the sample point.

Here, the results are shown with and without shadowing. Ament et al. has global parameters that bring out structures with different importance, depending on how the weights are tuned. Some parts of the neck are missing in the VISIBLE HUMAN and some vortex structures in the wake of the paddles are missing in the ROTATING
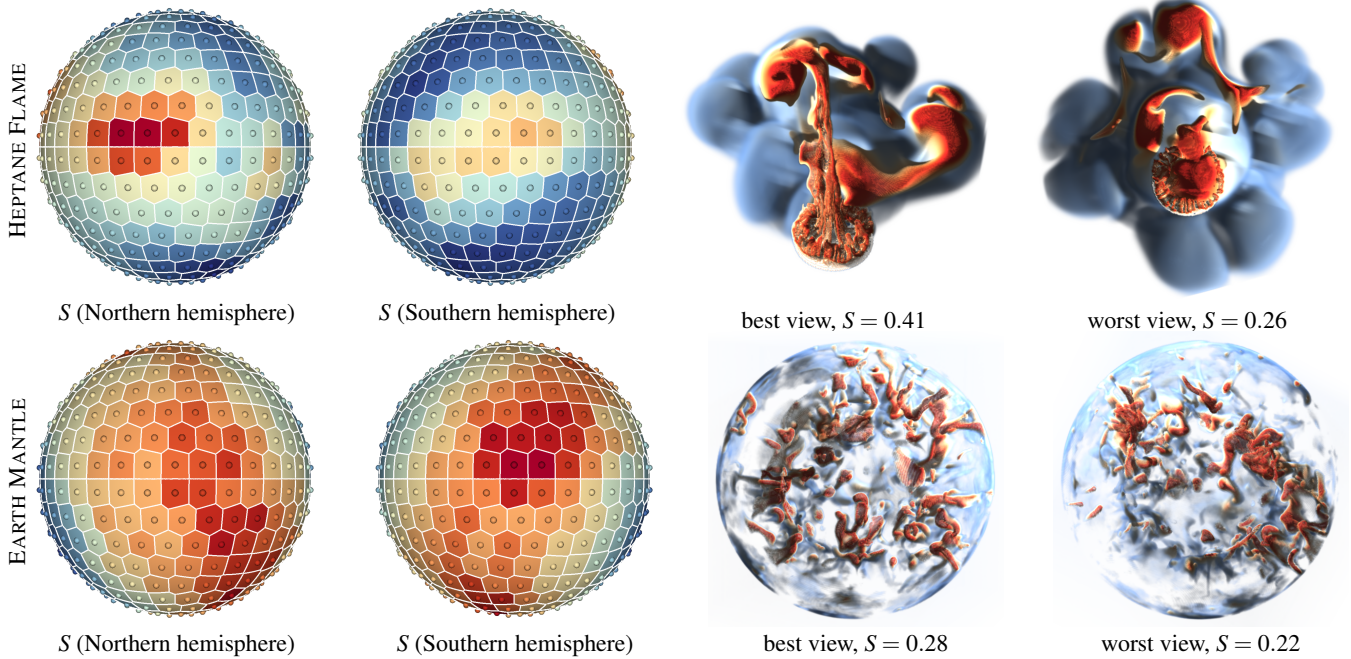
Figure 9: *Visualization of the visibility score S for different viewing directions in the* HEPTANE FLAME *and* EARTH MANTLE *data set, sampled on the Northern and Southern hemisphere (blue is best, red is worst). In addition, the best and worst camera perspective are shown.*

MIXER. We reported the visibility scores S that were obtained by the various methods by sampling them to world space. Since we optimize towards this metric, our approach consistently reached the highest score and showed details not only in high importance ranges.

Viewpoint Optimization. In our search for the optimal viewpoint, the candidates are restricted to a sphere. In Fig. 9, we sampled the sphere of possible viewing directions to convey the smoothness of the objective function. We visualize the visibility score S of Eq. (32) on the Northern and Southern hemisphere, showing that there are distinct viewing directions that are significantly better or worse than others. Further, the best and worst view are shown to convey the expressiveness of the metric. In the best views, the HEPTANE FLAME and the EARTH MANTLE utilize the screen space well, while the worst views are those with the most occlusion.

4.4. Grid Resolution

We discretize the transmittance integrals $T_V(\mathbf{y})$, $T_L(\mathbf{z})$ and the visibility integrals $G_V(\mathbf{y})$, $G_L(\mathbf{z})$ in both view space and light space, and we discretize the unknown extinction field $\mu_t(\mathbf{x})$ in world space. In Fig. 10, we analyze the impact of different grid resolutions on the obtained visual quality. Lowering the resolution of the world space grid results in noticeable voxel artifacts. In comparison, lowering the resolution of the view space grids is less noticeable. A too strong reduction shows grid artifacts in the shadows. As previously discussed in Section 3.5, we use a multi-resolution optimization that progresses along the diagonal of Fig. 10 from low to high resolution.
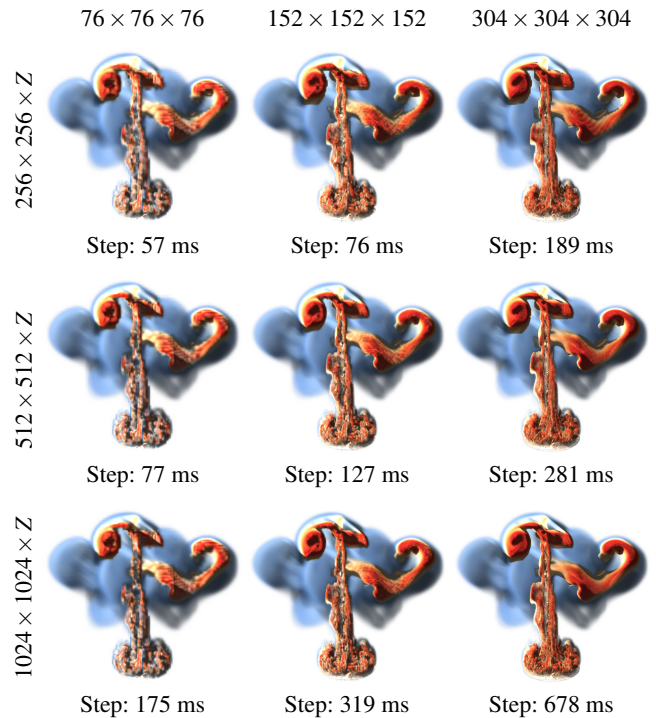

Figure 10: *Here, we systematically vary the resolution of the world space grid for the extinction $\mu(\mathbf{x})$ (in the columns) and the resolution of the view space grid and light space grids for transmittance $T_V(\mathbf{y})$, $T_L(\mathbf{z})$ and visibility integrals $G_V(\mathbf{y})$, $G_L(\mathbf{z})$ (in the rows). Hereby, Z is the value of the world space resolution times two.*

Data Set	World Space	View Space	Grid Updates	Optimization	Rendering	VRAM Usage
Rotating Mixer	$256 \times 256 \times 256$	$1024 \times 1024 \times 512$	447 ms	63 ms	47 ms	9.2GB
Earth Mantle	$256 \times 256 \times 256$	$1024 \times 1024 \times 512$	448 ms	65 ms	43 ms	9.2GB
Visible Human	$472 \times 472 \times 472$	$984 \times 984 \times 984$	793 ms	371 ms	65 ms	21.2GB
Heptane Flame	$304 \times 304 \times 304$	$1024 \times 1024 \times 608$	530 ms	104 ms	44 ms	11.2GB

Table 1: *Performance measurements for our method. All measurements were taken with an image resolution of* 1024×1024.

4.5. Performance

In the following, we report on the performance measurements and the memory consumption of our GPU implementation. All performance measures were taken on a workstation that is equipped with an AMD Ryzen 9 7950X CPU and an NVIDIA RTX 4090 GPU. We list the timings for the update of the view space and light space grids, a single optimization step, and the rendering time per frame in Table 1 for all data sets at a viewport resolution of 1024×1024 pixels. The reported resolution of the world space and view space grids are the ones used for all images throughout the paper, unless mentioned otherwise. While the rendering is possible at interactive rates (about 15 fps), the optimization process takes at full resolution several seconds, since a single frame takes about half a second, including the grid updates. To achieve interactivity, the view is optimized and rendered at lower resolution when moving and continues at full resolution when the camera is standing still, as demonstrated earlier in Fig. 5. The current video memory consumption of up to 21 GB leaves opportunities for further research in the future.

4.6. Discussion

Non-Linearity. In contrast to previous methods [AZD17, GTG17], our non-linear extinction optimization is not guaranteed to have a unique solution. By further adding the viewpoint optimization on top, the problem becomes even harder to solve. Similar to Weiss and Westermann [WW22], the problem can be approached by starting multiple optimizations from different initial conditions, which results in a sampling problem. In Fig. 9, we visualized the visibility score of the visibility optimization on a sphere, showing that the objective function is smooth enough for a gradient-based solver.

Memory Consumption. Currently, we discretize the transmittance fields, the importance integrals, and the unknown extinction field onto dense grids. The maximal resolution is thereby constraint by the available GPU memory. To circumvent this limitation, several approaches are imaginable, including sparse or hierarchical representations, compression algorithms, or approximations by projection into a different function basis [BRGG20].

Light Source Optimization. In this paper, we optimized the position of the camera to find the best view. Likewise, it is imaginable to optimize the light source position. This would, however, result in a trivial but unwanted solution. Since the same energy is minimized from the camera and light direction, cf. Eqs. (23)–(24), while the term in Eq. (22) has comparatively low weight, the light source would likely be placed at the camera location, resulting in a headlight that minimizes shadowing. Headlights, however, are known to have poor spatial perception. Instead, lateral lighting is preferable, which would have to be phrased in an energy term. We consider this an interesting direction for future research.

Perceptual Quality Metric. Our quality metric does not consider the human perception. For example, our metric is invariant to changes in the emitted radiance L_e. From a perceptual point of view, the quality metric should depend on the emitted radiance, since human perception depends on the luminance. In the future, it would be interesting to apply visual quality metrics that include or agree with aspects of the human visual perception [CWM*09, WQC*10, BRB*13].

5. Conclusions

Given a scalar-valued importance field, we proposed a variational formulation for the concurrent optimization of an extinction field and the camera viewpoint for direct volume rendering of three-dimensional scalar fields. For this, we measured the contribution of a voxel to the final pixel color directly by considering extinction and transmittance. We derived the necessary condition for an optimal extinction field and for an optimal camera transformation. The required functional derivatives that need to vanish have been calculated analytically. We employed an iterative multi-resolution gradient descent to find an optimum. The method has been applied to scalar fields from fluid dynamics, geophysics, and medicine.

In the future, it would be interesting to further investigate how the light source could be optimized, such that the visibility and the perception are maximized. At present, our formulation requires the scalar field and the importance field to be differentiable. Supporting non-differentiable scenes would be an interesting avenue for future research. So far, we discretized the unknown extinction field onto a grid, which simplified the optimization. Other function representations could be explored to reduce the required memory footprint. Further, it is imaginable that more specialized numerical solvers might help to accelerate the performance even further. It is imaginable to combine our approach with existing smart visibility techniques [VG05, AZD17] to provide interactive feedback during navigation. Further, varying levels of sparseness [VKG05] could add depth cues that anchor the important objects in the scene. Lastly, finding camera paths that view relevant structures in time-dependent data would be an interesting challenge.

Acknowledgements

References

[APH*03] AGRAWALA M., PHAN D., HEISER J., HAYMAKER J., KLINGNER J., HANRAHAN P., TVERSKY B.: Designing effective step-by-step assembly instructions. *ACM Transactions on Graphics 22*, 3 (jul 2003), 828–837. doi:10.1145/882262.882352. 2

[AZD17] AMENT M., ZIRR T., DACHSBACHER C.: Extinction-optimized volume illumination. *IEEE Transactions on Visualization and Computer Graphics 23*, 7 (2017), 1767–1781. doi:10.1109/TVCG.2016.2569080. 1, 2, 3, 4, 5, 6, 7, 8, 10

[BGKG05] BRUCKNER S., GRIMM S., KANITSAR A., GRÖLLER M. E.: Illustrative context-preserving volume rendering. In *EuroVis* (2005), pp. 69–76. doi:10.2312/VisSym/EuroVis05/069-076. 2

[BLD20] BANGARU S. P., LI T.-M., DURAND F.: Unbiased warped-area sampling for differentiable rendering. *ACM Transactions on Graphics 39*, 6 (nov 2020). doi:10.1145/3414685.3417833. 4

[BLL19] BERGER M., LI J., LEVINE J. A.: A generative model for volume rendering. *IEEE Transactions on Visualization and Computer Graphics 25*, 04 (apr 2019), 1636–1650. doi:10.1109/TVCG.2018.2816059. 3

[BM08] BAVOIL L., MYERS K.: Order independent transparency with dual depth peeling. *NVIDIA OpenGL SDK 1* (2008), 12. 5

[BRB*13] BRAMON R., RUIZ M., BARDERA A., BOADA I., FEIXAS M., SBERT M.: An information-theoretic observation channel for volume visualization. *Computer Graphics Forum 32*, 3pt4 (2013), 411–420. doi:10.1111/cgf.12128. 10

[BRGG20] BAEZA ROJO I., GROSS M., GÜNTHER T.: Fourier opacity optimization for scalable exploration. *IEEE Transactions on Visualization and Computer Graphics 26*, 11 (2020), 3204–3216. doi:10.1109/TVCG.2019.2915222. 2, 10

[BS05] BORDOLOI U., SHEN H.-W.: View selection for volume rendering. In *VIS 05. IEEE Visualization, 2005.* (2005), pp. 487–494. doi:10.1109/VISUAL.2005.1532833. 3

[CJ10] CHEN M., JÄENICKE H.: An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1206–1215. doi:10.1109/TVCG.2010.132. 3

[CM11] CORREA C. D., MA K.-L.: Visibility histograms and visibility-driven transfer functions. *IEEE Transactions on Visualization and Computer Graphics 17*, 2 (2011), 192–204. doi:10.1109/TVCG.2010.35. 2, 4

[CSC07] CORREA C., SILVER D., CHEN M.: Illustrative deformation for data exploration. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1320–1327. doi:10.1109/TVCG.2007.70565. 2

[CWM*09] CHAN M.-Y., WU Y., MAK W.-H., CHEN W., QU H.: Perception-based transparency optimization for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1283–1290. doi:10.1109/TVCG.2009.172. 2, 4, 10

[GBT21] GÜLCHER A. J. P., BALLMER M. D., TACKLEY P. J.: Coupled dynamics and evolution of primordial and recycled heterogeneity in Earth's lower mantle. *Solid Earth 12*, 9 (2021), 2087–2107. doi:10.5194/se-12-2087-2021. 7

[GF63] GELFAND I. M., FOMIN S. V.: *Calculus of variations.* Prentice-Hall Inc., Englewood Cliffs, N. J., 1963. doi:10.2307/3611748. 4

[GRT13] GÜNTHER T., RÖSSL C., THEISEL H.: Opacity optimization for 3D line fields. *ACM Transactions on Graphics (SIGGRAPH) 32*, 4 (2013), 120:1–120:8. doi:10.1145/2461912.2461930. 1, 2, 7

[GRT14] GÜNTHER T., RÖSSL C., THEISEL H.: Hierarchical opacity optimization for sets of 3D line fields. *Computer Graphics Forum (Eurographics) 33*, 2 (2014), 507–516. doi:10.1111/cgf.12336. 2

[GSME*14] GÜNTHER T., SCHULZE M., MARTINEZ ESTURO J., RÖSSL C., THEISEL H.: Opacity optimization for surfaces. *Computer Graphics Forum (Eurographics Conference on Visualization) 33*, 3 (2014), 11–20. doi:10.1111/cgf.12357. 2

[GT18] GÜNTHER T., THEISEL H.: The state of the art in vortex extraction. *Computer Graphics Forum 37*, 6 (2018), 149–173. doi:10.1111/cgf.13319. 7

[GTG17] GÜNTHER T., THEISEL H., GROSS M.: Decoupled opacity optimization for points, lines and surfaces. *Computer Graphics Forum (Eurographics) 36*, 2 (2017), 153–162. doi:10.1111/cgf.13115. 2, 6, 10

[HG41] HENYEY L. G., GREENSTEIN J. L.: Diffuse radiation in the Galaxy. *The Astrophysical Journal 93* (Jan. 1941), 70–83. doi:10.1086/144246. 3

[HHKP96] HE T., HONG L., KAUFMAN A., PFISTER H.: Generation of transfer functions with stochastic search techniques. In *Proceedings of Seventh Annual IEEE Visualization '96* (1996), pp. 227–234. doi:10.1109/VISUAL.1996.568113. 2

[Jar08] JAROSZ W.: *Efficient Monte Carlo methods for light transport in scattering media.* PhD thesis, University of California, San Diego, 2008. 3

[JSYR14] JÖNSSON D., SUNDÉN E., YNNERMAN A., ROPINSKI T.: A survey of volumetric illumination techniques for interactive volume rendering. *Computer Graphics Forum 33*, 1 (2014), 27–51. doi:10.1111/cgf.12252. 2, 3

[KB17] KINGMA D. P., BA J.: Adam: A method for stochastic optimization, 2017. arXiv:1412.6980. 7

[KBM*20] KATO H., BEKER D., MORARIU M., ANDO T., MATSUOKA T., KEHL W., GAIDON A.: Differentiable rendering: A survey, 2020. arXiv:2006.12057. 3

[KD98] KINDLMANN G., DURKIN J.: Semi-automatic generation of transfer functions for direct volume rendering. In *IEEE Symposium on Volume Visualization (Cat. No.989EX300)* (1998), pp. 79–86. doi:10.1109/SVV.1998.729588. 2

[LADL18] LI T.-M., AITTALA M., DURAND F., LEHTINEN J.: Differentiable monte carlo ray tracing through edge sampling. *ACM Transactions on Graphics 37*, 6 (dec 2018). doi:10.1145/3272127.3275109. 4

[LMSC11] LEE T.-Y., MISHCHENKO O., SHEN H.-W., CRAWFIS R.: View point evaluation and streamline filtering for flow visualization. In *2011 IEEE Pacific Visualization Symposium* (2011), pp. 83–90. doi:10.1109/PACIFICVIS.2011.5742376. 3

[MMD10] MARCHESIN S., MONGENET C., DISCHLER J.: Per-pixel opacity modulation for feature enhancement in volume rendering. *IEEE Transactions on Visualization and Computer Graphics 16*, 04 (jul 2010), 560–570. doi:10.1109/TVCG.2010.30. 1, 2, 7, 8

[NDVZJ19] NIMIER-DAVID M., VICINI D., ZELTNER T., JAKOB W.: Mitsuba 2: A retargetable forward and inverse renderer. *ACM Transactions on Graphics 38*, 6 (nov 2019). doi:10.1145/3355089.3356498. 3

[RBB*11] RUIZ M., BARDERA A., BOADA I., VIOLA I., FEIXAS M., SBERT M.: Automatic transfer functions based on informational divergence. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 1932–1941. doi:10.1109/TVCG.2011.173. 2

[RHGJ03] RATIU P., HILLEN B., GLASER J., JENKINS D. P.: Visible human 2.0–the next generation. In *Medicine Meets Virtual Reality 11.* IOS Press, 2003, pp. 275–281. doi:10.3233/978-1-60750-938-7-275. 7

[RRN*20] RAVI N., REIZENSTEIN J., NOVOTNY D., GORDON T., LO W.-Y., JOHNSON J., GKIOXARI G.: Accelerating 3D deep learning with PyTorch3D, 2020. arXiv:2007.08501. 3

[SJ06] SHEN H., JI G.: Dynamic view selection for time-varying volumes. *IEEE Transactions on Visualization and Computer Graphics 12*, 05 (sep 2006), 1109–1116. doi:10.1109/TVCG.2006.137. 2

[TFTN05] TAKAHASHI S., FUJISHIRO I., TAKESHIMA Y., NISHITA T.: A feature-driven approach to locating optimal viewpoints for volume visualization. In *VIS 05. IEEE Visualization, 2005.* (2005), pp. 495–502. doi:10.1109/VISUAL.2005.1532834. 2

[TLB*09] TAO Y., LIN H., BAO H., DONG F., CLAPWORTHY G.: Structure-aware viewpoint selection for volume visualization. In *IEEE Pacific Visualization Symposium* (2009), pp. 193–200. `doi:10.1109/PACIFICVIS.2009.4906856`. 3

[TMWS13] TAO J., MA J., WANG C., SHENE C.-K.: A unified approach to streamline selection and viewpoint selection for 3D flow visualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 3 (2013), 393–406. `doi:10.1109/TVCG.2012.143`. 3

[TWC*16] TAO Y., WANG Q., CHEN W., WU Y., LIN H.: Similarity voting based viewpoint selection for volumes. *Computer Graphics Forum 35*, 3 (2016), 391–400. `doi:10.1111/cgf.12915`. 3

[VG05] VIOLA I., GRÖLLER M. E.: Smart visibility in visualization. In *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging* (Goslar, DEU, 2005), Eurographics Association, p. 209–216. `doi:10.2312/COMPAESTH/COMPAESTH05/209-216`. 1, 2, 4, 10

[VKG04] VIOLA I., KANITSAR A., GROLLER M.: Importance-driven volume rendering. In *IEEE Visualization 2004* (2004), pp. 139–145. `doi:10.1109/VISUAL.2004.48`. 2, 7, 8

[VKG05] VIOLA I., KANITSAR A., GRÖLLER M. E.: Importance-driven feature enhancement in volume visualization. *IEEE Transactions on Visualization and Computer Graphics 11*, 4 (7 2005), 408–418. `doi:10.1109/TVCG.2005.62`. 2, 10

[VMN08] VÁZQUEZ P.-P., MONCLÚS E., NAVAZO I.: Representative views and paths for volume models. In *Smart Graphics* (Berlin, Heidelberg, 2008), Butz A., Fisher B., Krüger A., Olivier P., Christie M., (Eds.), Springer Berlin Heidelberg, pp. 106–117. `doi:10.1007/978-3-540-85412-8_10`. 3

[VSJ21] VICINI D., SPEIERER S., JAKOB W.: Path replay backpropagation: Differentiating light paths using constant memory and linear time. *ACM Transactions on Graphics 40*, 4 (jul 2021). `doi:10.1145/3450626.3459804`. 4

[WQ07] WU Y., QU H.: Interactive transfer function design based on editing direct volume rendered images. *IEEE Transactions on Visualization and Computer Graphics 13*, 5 (2007), 1027–1040. `doi:10.1109/TVCG.2007.1051`. 2

[WQC*10] WU Y., QU H., CHUNG K., CHAN M., ZHOU H.: Quantitative effectiveness measures for direct volume rendered images. In *2010 IEEE Pacific Visualization Symposium (PacificVis 2010)* (Los Alamitos, CA, USA, mar 2010), IEEE Computer Society, pp. 1–8. `doi:10.1109/PACIFICVIS.2010.5429623`. 10

[WW22] WEISS S., WESTERMANN R.: Differentiable direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics 28*, 01 (jan 2022), 562–572. `doi:10.1109/TVCG.2021.3114769`. 3, 10

[XLS10] XU L., LEE T.-Y., SHEN H.-W.: An information-theoretic framework for flow visualization. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1216–1224. `doi:10.1109/TVCG.2010.131`. 3

[YLLY19] YANG C., LI Y., LIU C., YUAN X.: Deep learning-based viewpoint recommendation in volume visualization. *Journal of visualization 22* (2019), 991–1003. `doi:10.1007/s12650-019-00583-4`. 3

[YZM*23] YU Z., ZHANG C., MAURY O., HERY C., DONG Z., ZHAO S.: Efficient path-space differentiable volume rendering with respect to shapes. *Computer Graphics Forum (Proc. Eurographics Symposium on Rendering) 42*, 4 (2023), e14884. `doi:10.1111/cgf.14884`. 4

[ZMY*20] ZHANG C., MILLER B., YAN K., GKIOULEKAS I., ZHAO S.: Path-space differentiable rendering. *ACM Transactions on Graphics 39*, 4 (aug 2020). `doi:10.1145/3386569.3392383`. 4

[ZRPD20] ZEIDAN M., RAPP T., PETERS C., DACHSBACHER C.: Moment-based opacity optimization. In *Eurographics Symposium on Parallel Graphics and Visualization* (2020), Frey S., Huang J., Sadlo F., (Eds.), The Eurographics Association. `doi:10.2312/pgv.20201072`. 2

[ZSGJ21] ZELTNER T., SPEIERER S., GEORGIEV I., JAKOB W.: Monte carlo estimators for differential light transport. *ACM Transactions on Graphics 40*, 4 (jul 2021). `doi:10.1145/3450626.3459807`. 4