



Technische Universiteit
Eindhoven
University of Technology

Department of Mathematics and Computer Science
Visualization Cluster

Uncertainty Visualization for Neural Radiance Field

Master Thesis

Thiam Wai Chua

Supervisors:

Prof. dr. Anna Vilanova
Dr. Nicola Pezzotti
M.Sc. K.W.H. Maas

Eindhoven, January 2024

Abstract

Neural Radiance Field (NeRF) has demonstrated remarkable success in synthesizing realistic scenes. NeRFs generate novel views and detailed reconstructions from two-dimensional (2D) images, utilizing deep learning and camera geometry information. Specifically, NeRF models a scene's volumetric properties by learning the mapping between three-dimensional (3D) spatial coordinates and the radiance of the scene.

Despite the remarkable success of NeRF in scene synthesis, it poses inherent challenges related to uncertainty. For example, uncertainty may arise from model uncertainty, limited training images, or complex scenes. Model uncertainty is intrinsically related to the uncertainty inherent in the model itself, reflecting the model's ability to synthesize scenes accurately within the context of NeRF. The complex scene includes objects with intricate details, such as irregular surfaces or challenging lighting conditions. To address these challenges, it is beneficial to develop a system that enables researchers and practitioners to explore estimated uncertainties in scenes through interactive visualization, fostering a comprehensive understanding of the model uncertainty. This empowers them to make informed and confident decisions and facilitate improvements to the model.

In this research, we develop a system called NeRFDeltaView that facilitates users in exploring the model uncertainty through interactive visualization approaches. Two uncertainty estimation methods are used within the NeRF architecture: the Uncertainty Neural Network and Ensemble model. The Uncertainty Neural Network learns to predict uncertainty values at each 3D spatial coordinate within its neural network, while the Ensemble model estimates uncertainty by comparing predicted radiance from multiple neural networks. We introduce novel visualizations to represent these estimated uncertainties visually, enabling users to interactively explore model uncertainty. To the best of our knowledge, no interactive visualization has been proposed that facilitates such an exploration of the NeRF model uncertainty.

The NeRFDeltaView system has been evaluated through three use cases. The system demonstrates its capabilities in exploring and interpreting model uncertainty. The findings show that the uncertainty estimation method is critical for obtaining reliable results, as both models encounter challenges in estimating uncertainty in a complex scene. Furthermore, our system allows users to identify specific viewing directions characterized by high uncertainties effectively and then enhances model prediction performance by incorporating extra training images in these specific viewing directions. The system offers an alternative perspective to facilitate users in improving and refining the model for higher accuracy scene synthesis through the uncertainty visualization results. It also facilitates making informed and confident decisions and assessing potential risks in various applications, including medical imaging and automotive.

Contents

Contents	iii
List of Figures	v
1 Introduction	1
2 Background	3
2.1 Neural Radiance Field	3
2.1.1 Neural Radiance Field scene representation	4
2.1.2 Volume rendering with radiance field	4
2.1.3 Improvements of the Neural Radiance Field	5
2.2 Volume rendering	5
3 Related work	8
3.1 Uncertainty estimation in deep learning	8
3.2 Uncertainty estimation in Neural Radiance Field	9
3.3 Uncertainty visualization	10
4 NeRFDeltaView	15
5 Uncertainty estimation	18
5.1 Uncertainty Neural Network	18
5.2 Ensemble	20
6 Visualization design	23
6.1 Uncertainty visualization	23
6.1.1 Volume rendering panel	25
6.1.2 Data visualization panel	28
6.1.3 Heatmap	29
6.1.4 Interactions	30
6.2 Implementation details	32
7 Experimental results	34
7.1 Datasets	34
7.2 Training strategy for comparative analysis	37
7.3 Use cases	38
7.3.1 Dataset: chair	38
7.3.2 Dataset: microphone	44
7.3.3 Dataset: drums	48
7.3.4 Summary and discussion of findings	54
8 Conclusion	56

CONTENTS

Bibliography	59
Appendix	64
A Experimental results	65

List of Figures

2.1	Image sourced from Mildenhall et al. [54]. The overview of the NeRF model pipeline. F_Θ represents the multilayer perceptron where Θ is the learning parameters. x , y , and z represent the spatial coordinates of a sample point. θ and ϕ denote the directions of a ray. The output colors $\mathbf{c} = (R, G, B)$, denoted by R , G , and B corresponding to red, green, and blue, respectively. σ represents the output volume density. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value and $\ \cdot\ _2^2$ is squared Euclidean norm.	3
2.2	Images sourced from Preim and Botha [69]. Isosurfaces extracted from computerized tomography (CT) data. (a) A section of the spine in the breast region. (b) Pelvis extracted from CT data.	5
2.3	Direct volume rendering (DVR) of a high potential iron protein with different transfer function settings using Paraview [3]. The histogram is displayed behind the transfer function. In (b), higher scalar values of the high potential iron protein are assigned high opacity, while other scalar values are set to low opacity in the transfer function.	6
2.4	Images sourced from Djurcicov et al. [15]. Direct volume rendering (DVR) of ocean volumetric data with different 2D transfer function settings. Below each representation of the volume-rendered ocean volumetric data, the corresponding transfer function setting is shown.	7
3.1	Images sourced from Lee et al. [42]. (a) A pipeline of the NeRF integrated with the localized uncertainty assessment, guiding additional scene image acquisition for enhanced scene reconstruction. (b) The camera directions and positions on the upper hemisphere of the scene were used for training NeRF initially. (c) The three-dimensional (3D) mesh and the uncertainty map of a specific view of the scene.	9
3.2	Image sourced from NeU-NBV [33]. NeU-NBV employs uncertainty estimation integrated into NeRF to guide measurement acquisition. The NeU-NBV selects the most informative view candidate with the highest uncertainty for acquiring additional scene images, thereby enhancing scene reconstruction.	10
3.3	(a) Images sourced from Lundström et al. [46], displays two possible artery visualizations achieved by adjusting the visualization parameters. (b) Images sourced from Potter [66], displays two possible brain tumor tissue boundary positions resulting from varying visualization parameters.	11
3.4	Image sourced from Weiskopf [84]. Visualization pipeline including uncertainty. The pipeline consists of different stages: data acquisition, filtering and transformations, visual mapping and rendering, perception and cognition of the users, and the interaction between users and previous stages.	11
3.5	A categorization of different visualization of uncertainty methods. The figure is adapted from Raidou [72]. Refer to Raidou [72] for a comprehensive list of visualization of uncertainty methods.	12
3.6	Image sourced Grigoryan and Rheingans [19]. Color indicates uncertain surface.	12

LIST OF FIGURES

3.7	Images sourced Cedelnik et al. [11]. Scalar-valued uncertainty is used to introduce disturbances in grid lines with varying levels of distortion.	13
3.8	Volume-rendered images of (a,b) ocean data and (b) brain data. The normalized standard deviation is linearly mapped to opacity and color in (b) and (c). In (b), red opaque regions represent high standard deviation, while blue transparent regions represent low standard deviation. In (c), voxels with normalized standard deviation less than 0.15 were rendered transparent for a clearer view of regions with higher standard deviation.	14
3.9	(a) Volume visualization of isosurface crossing in uncertain 3D temperature field in the exosphere above Europe and the North Atlantic Ocean [65]. (b) Isosurface of a synthetic dataset and (c) isosurface of an magnetic resonance dataset surrounded by a volume-rendered region representing uncertainty [34].	14
4.1	Overview of the NeRFDeltaView system. The top row is the NeRF models integrated with uncertainty estimation. $Avg[\cdot]$ is an averaging method to estimate the uncertainty δ . The bottom row is the uncertainty visualization technique to visualize the uncertainty estimated by these models.	16
5.1	The overview of the Uncertainty Neural Network model pipeline. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value. Part of the figure is adapted from NeRF [54].	19
5.2	A visualization of the fully-connected multilayer perceptron for the Uncertainty Neural Network. The neural network is similar to NeRF [54] with an additional estimation of uncertainty δ , inspired by Recursive-NeRF [87]. Input vectors are in green, hidden layers in cyan, output vectors in red, and the number in each block represents the vector's dimension, for example, the number in each hidden layer represents the number of neurons. The "+" signifies vector concatenation. The $\gamma(\mathbf{x})$ and $\gamma(\mathbf{d})$ are the positional encoding of the spatial coordinate of the sample point and viewing angle direction of the ray, respectively.	19
5.3	The overview of the Ensemble network architecture pipeline. The $(\mathbf{c}, \sigma)_P$ is the color \mathbf{c} and the volume density σ estimated by the primary MLP. Whereas, the $(\mathbf{c}, \sigma)_{S1}$, $(\mathbf{c}, \sigma)_{S(n-1)}$, and $(\mathbf{c}, \sigma)_{Sn}$ are estimated by the first secondary MLP, (n-1)-th secondary MLP, and n-th secondary MLP, respectively. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value. Part of the figure is adapted from NeRF [54].	21
5.4	Primary and secondary multilayer perceptrons for the Ensemble architecture. Input vectors are in green, hidden layers in cyan, output vectors in red, and the number in each block represents the vector's dimension, for example, the number in each hidden layer represents the number of neurons. The "+" signifies vector concatenation. The $\gamma(\mathbf{x})$ and $\gamma(\mathbf{d})$ are the positional encoding of the spatial coordinate of the sample point and viewing angle direction of the ray, respectively.	21
5.5	The computation of color uncertainty δ_c and density uncertainty δ_σ involves using the method of averaging of the pairwise Euclidean distances for a given point in the three-dimensional (3D) grid. Two separate 3D grids are created, one for storing the δ_c and the other for storing δ_σ , respectively. The notation d_{01}^c represents the pairwise Euclidean distance between $(\mathbf{c})_P$ and $(\mathbf{c})_{S1}$, and $d_{1(n-1)}^c$ represents the pairwise Euclidean distance between $(\mathbf{c})_{S1}$ and $(\mathbf{c})_{S(n-1)}$. Identical notation applies to other pairwise Euclidean distances in the color and volume density predictions.	22

6.1	Screenshot of the interactive tool for uncertainty visualization in the NeRFDeltaView system for Uncertainty NN. Data and control panel (A) contains a control button and meta-information. Volume rendering panel (B) displays the synthetic, isosurface, and uncertainty three-dimensional views. The data visualization panel (C) consists of two tabs. The first tab includes the TF, histograms, and one-dimensional (1D) TF. The second tab shows heatmaps of the mean and standard deviation of uncertainty estimation for selected viewing directions.	24
6.2	Screenshot of the interactive tool for uncertainty visualization in the NeRFDeltaView system for the Ensemble model. The data and control panel (A) contains a control button and meta-information. Volume rendering panel (B) displays the synthetic, isosurface, and uncertainty three-dimensional views. The data visualization panel (C) consists of two tabs. The first tab includes the TFs, histograms, and a two-dimensional (2D) TF of color and density uncertainties. The second tab shows heatmaps of the mean and standard deviation of the color and density uncertainty estimation for selected viewing directions.	25
6.3	Examples of the synthesized scene and uncertainty are presented (a) in separate views for a side-by-side comparison and (b) in the same view using an overloading approach.	26
6.4	The uncertainty is overloaded on the isosurface (a) without using a Z-buffer and (b) with a Z-buffer.	26
6.5	Illustration of the uncertainty 3D view with (a) cutting plane intersects with the chair isosurface and uncertainty, (b) visible chair isosurface, and (c) hidden chair isosurface. Users can translate and rotate the cutting plane as needed. The isosurface object in the uncertainty 3D view can be toggled on or off.	28
6.6	The (a) 2D TF, (b) 2D density plot of the 2D TF, and (c) the color and density uncertainty 3D views are shown after selecting points in the 2D TF on (b). The density plot is estimated using KDE and becomes visible after some points are hidden.	29
6.7	The chair uncertainty scene is surrounded by the sphere on both the (a) upper hemisphere and (b) lower hemisphere. The pyramids represent the cameras. The inserted images display the uncertainty 3D views for the respective camera. The red point in (a) represents 0° for both azimuth and elevation angles, i.e., it is located on the top view of the chair scene. In (c), the clusters of cells are presented in a heatmap. The colors of each cluster of cells represent the viewing directions from the upper or lower hemispheres. The blue-bordered rectangle encompasses cells with viewing directions from the upper hemisphere, while the green-bordered rectangles encompass cells with viewing directions from the lower hemisphere. Lastly, (d) shows an example of a heatmap. The blue points in the heatmap represent the viewing directions of the training images.	31
6.8	The synthetic view, isosurface view, and the uncertainty 3D view of the chair scene at an elevation of -90° and 90° . The mean values are constant at these elevations because the camera is located at the poles of the hemisphere.	32
6.9	The uncertainty 3D views corresponding to the selected cells in the heatmap of the mean density uncertainty when the cursor hovers over these cells.	32
7.1	Example of the (a) chair, (b) microphone, and (c) drums training images. In (d), the drums are surrounded by pyramids in the upper hemisphere [54]. The pyramids represent the camera.	35
7.2	An example of the training images from the limited-angle dataset for the chair, microphone, and drums. The pyramids represent the camera, and the insert images display the training images captured by the respective cameras.	36

LIST OF FIGURES

7.3	An example of a solution space or parameter space with two local minima. The model has initial parameters Θ . $\Theta_{2_{\text{full}}}^*$ represents the optimized parameters of the model trained with the full-angle dataset. $\Theta_{1_{\text{lim}}}^*$ and $\Theta_{1_{\text{full}}}^*$ represent the optimized parameters of the model first trained with limited-angle dataset and then with the full-angle dataset, respectively. The dotted arrows represent possible paths during optimization, each with 100000 iterations.	38
7.4	The synthetic view, uncertainty 3D view, transfer function, and the 1D transfer function for the chair scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.	39
7.5	The mean and standard deviation uncertainty heatmaps in selected viewing directions for the chair scene trained with the Uncertainty NN model using (a) limited-angle and (b) full-angle datasets. The inserted images are the synthetic view and uncertainty 3D view for the corresponding cells.	40
7.6	The synthetic view, color and density uncertainty 3D views, transfer function, and the 2D transfer function for the chair scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.	41
7.7	The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the chair scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets. The inserted images are the synthetic view and uncertainty 3D view for the corresponding cells.	43
7.8	The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the chair scene. In the uncertainty 3D views, a cutting plane intersects with the volumetric data. The mean uncertainty values in the heatmaps have been normalized.	44
7.9	The synthetic view, uncertainty 3D view, transfer functions, 1D transfer functions, and the heatmaps for the microphone scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.	45
7.10	The synthetic view, color and density uncertainty 3D views, transfer functions and the 2D transfer function for the microphone scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.	46
7.11	The (a) mean color uncertainty heatmaps and (b) mean density uncertainty heatmaps in selected viewing directions for the microphone scene trained with the Ensemble model using the limited-angle dataset (left heatmap) and full-angle dataset (right heatmap). The pattern in the standard deviation uncertainty heatmaps is similar to their respective mean uncertainty heatmaps. The standard deviation uncertainty heatmaps can be found in Figure A.1 in Appendix A.	47
7.12	The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the microphone scene. The mean uncertainty values in the heatmaps have been normalized.	48
7.13	The synthetic view, uncertainty 3D view, transfer function, and 1D transfer function for the drums scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.	49
7.14	The mean and standard deviation uncertainty heatmaps in selected viewing directions for the drums scene trained with the Uncertainty NN model using (a) limited-angle and (b) full-angle datasets.	50
7.15	The synthetic view, color uncertainty 3D views, transfer functions, and the 2D transfer functions for the drums scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.	51
7.16	The synthetic view, density uncertainty 3D views, transfer functions, and the 2D transfer functions for the drums scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.	52

7.17	The (a) mean color uncertainty heatmaps and (b) mean density uncertainty heatmaps in selected viewing directions for the drums scene trained with the Ensemble model using limited-angle dataset (left heatmap) and full-angle dataset (right heatmap). The pattern in the standard deviation uncertainty heatmaps is similar to their respective mean uncertainty heatmaps. The standard deviation uncertainty heatmaps can be found in Figure A.2 in Appendix A.	53
7.18	The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the drums scene. The mean uncertainty values in the heatmaps have been normalized.	54
A.1	The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the microphone scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets.	66
A.2	The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the drums scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets.	67

Chapter 1

Introduction

Neural Radiance Field (NeRF) proposed by Mildenhall et al. [54] have emerged as a powerful approach for synthesizing realistic three-dimensional (3D) scenes from two-dimensional (2D) images using deep neural networks, finding applications in robotics and medical imaging. This revolutionary technique has gained significant attention in computer graphics and computer vision communities, due to its ability to generate high-fidelity, novel views of scenes. NeRF leverages a volumetric representation to model the appearance and geometry of a scene by implicitly encoding a mapping from 3D spatial coordinates to scene radiance values.

NeRF has demonstrated its effectiveness in addressing challenges, such as complex lighting conditions and geometric variations in scenes. Despite its impressive capabilities in scene synthesis, NeRF also presents inherent challenges related to uncertainty [61]. According to Hunter and Goodchild [30], uncertainty could be characterized as the lack of knowledge regarding error, resulting in a reluctance to accept the outcomes and observations without applying caution. The uncertainty in the context of NeRF could arise from various sources, such as model uncertainty, a limited amount of training images, noisy image data, and complex scenes. NeRF heavily relies on 2D training images to learn the mapping between 3D spatial coordinates and scene radiance values. If the available training data is limited in terms of quantity, the model may have difficulty accurately synthesizing the scenes. Uncertainty arises when the model is trained on noisy images, as this can introduce inconsistencies in the training process, resulting in a less reliable synthesis of scenes by the model. Scenes with complex objects, such as intricate details, might present challenges for the model, leading to uncertainty in scene synthesis. Model uncertainty is intrinsically linked to the uncertainty in the model itself, also called epistemic uncertainty. In the context of NeRF, model uncertainty presents the capability and confidence of the model to synthesize the scene accurately. All in all, the exploration of estimated uncertainty is important for researchers and practitioners to understand the behavior of NeRF model uncertainty in a scene. This understanding is particularly important in applications such as autonomous driving and medical diagnosis, where the decisions influenced by deep learning directly impact individuals' lives.

Several NeRF variants that incorporate uncertainty estimation to enhance their scene synthesis have emerged. For instance, Recursive Neural Radiance Field (Recursive-NeRF) [87] optimizes the synthesis process by adaptively learning uncertainties for query coordinates. NeRF in the Wild (NeRF-W) [51] introduces pixel-level uncertainty estimation based on the variance of the predicted color. Lee et al. [42] propose a ray-based volumetric uncertainty estimator for uncertainty-guided policy. By providing insights into the uncertainty associated with the scenes, the researchers gain a better understanding of the model uncertainty. However, these existing NeRFs that incorporate uncertainty estimation often lack effective interactive visualization methods to communicate the model uncertainty to the researchers. Relating the spatial distribution of uncertainty within a synthesized scene from multiple uncertainty images presents a challenge. To address this challenge, developing an effective interactive visualization could assist researchers in analyzing the

uncertainty, leading to a better understanding of model uncertainty.

The main goal of this thesis is to develop a visualization approach that facilitates researchers and practitioners in exploring uncertainty in a scene and understanding NeRF model uncertainty. A series of interactive visualizations will be implemented to support the exploration of uncertainty and understanding of the behavior of the model uncertainty within a scene. We display the effectiveness of these visualizations across three distinct scenes with different levels of intricacy. The interactive visualization comprises volume-rendered uncertainty views and overview views. The volume-rendered uncertainty views show the detailed spatial distribution of uncertainty within the synthesized scene. Whereas the overview views provide a comprehensive, high-level perspective on the overall uncertainty patterns. The interactions between these views enable users to focus on specific scene features in the synthesized scene, facilitating exploration and analysis of uncertainty, ultimately leading to a deeper understanding of model uncertainty.

To enable this exploration and analysis, we developed NeRFDeltaView (Neural Radiance Field Delta View) system. This system comprises two methods for estimating the uncertainty using the NeRF model and then visualizing these estimated uncertainties through an interactive visualization tool. The first uncertainty estimation method involves directly predicting uncertainty using a neural network (NN) and a loss function computation adapted from Recursive-NeRF [87]. The second method revolves around an ensemble NeRF model inspired by Deep Ensembles [41], where multiple NNs are employed to estimate scene properties. These two uncertainty estimation methods capture NeRF model uncertainty, with the former estimating uncertainty from a single deterministic NN, and the latter estimating uncertainty from multiple NNs with varying parameter initialization. Through the interactive visualization of these estimated uncertainties, researchers and practitioners can gain insights into the reliability of synthesized scenes, enabling them to make informed and confident decisions, for example, additional training images from specific viewing directions for further optimization. Researchers and practitioners can leverage this system to identify limitations in the model uncertainty, aiding in the improvement of the model architecture and exploring the correlation between viewing direction and NeRF model uncertainty.

The remaining part of the thesis is structured as follows. In Chapter 2, we present the background of NeRF model architecture originally proposed by Mildenhall et al. [54] and volume rendering. Chapter 3 provides an overview of relevant works in the field, including various NeRF models that incorporate uncertainty estimation and uncertainty visualization. Chapter 4 provides an overview of the NeRFDeltaView system. Next, in Chapter 5, we delve into two methods for uncertainty estimation with NeRF within the NeRFDeltaView system. The visualization design in the system is described in Chapter 6. Lastly, Chapter 7 focuses on the discussion of the results and findings, while Chapter 8 presents the conclusion, limitations, and potential future works.

Chapter 2

Background

In this chapter, we first provide a concise introduction to the background of Neural Radiance Field (NeRF) [54], establishing a foundation for comprehending the research presented in this thesis. Subsequently, we provide a brief overview of the topics in volume rendering.

2.1 Neural Radiance Field

NeRF [54] is a novel approach for three-dimensional (3D) scene reconstruction from a set of two-dimensional (2D) training images. It represents a powerful technique in the field of computer graphics and computer vision for synthesizing scenes in a photorealistic manner. NeRF operates by modeling a scene's properties (color and density), through a neural network. Figure 2.1 visually represents the NeRF model pipeline. This network takes spatial coordinates as inputs and predicts the corresponding scene properties. The pipeline of NeRF includes the following steps: (Figure 2.1a) selecting sampling points along rays corresponding to pixels in a synthesized image, (Figure 2.1b) using a multilayer perceptron (MLP) to predict colors and densities at these sampling points, (Figure 2.1c) generating pixel color by considering the colors and densities along the camera rays through volume rendering, and (Figure 2.1d) comparing the generated pixel color via volume rendering with the ground truth pixel color.

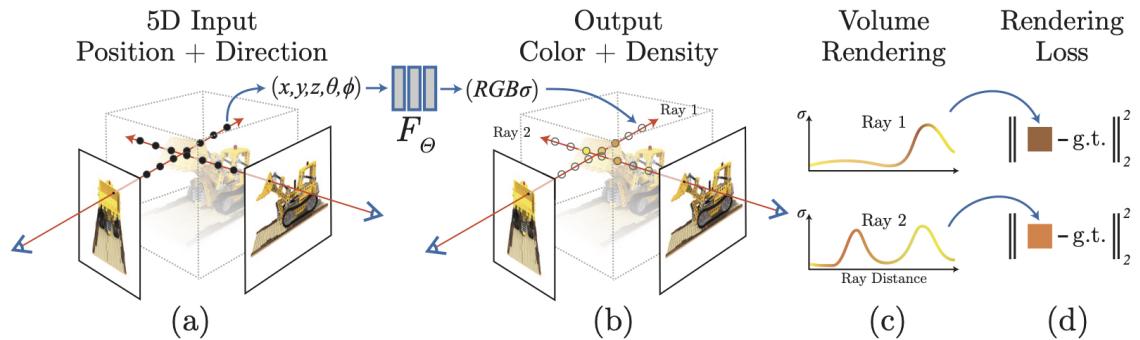


Figure 2.1: Image sourced from Mildenhall et al. [54]. The overview of the NeRF model pipeline. F_Θ represents the multilayer perceptron where Θ is the learning parameters. x, y , and z represent the spatial coordinates of a sample point. θ and ϕ denote the directions of a ray. The output colors $c = (R, G, B)$, denoted by R, G , and B corresponding to red, green, and blue, respectively. σ represents the output volume density. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value and $\|\cdot\|_2^2$ is squared Euclidean norm.

2.1.1 Neural Radiance Field scene representation

The NeRF model represents 3D scenes as a radiance field with an MLP F_Θ as shown in Figure 2.1, where Θ represents the learning parameters. The radiance field describes color \mathbf{c} and volume density σ for each sample point within the scene. The color information is crucial for determining the visual appearance of objects in the scene and is direction-dependent, while density is vital for understanding the occupancy of the scene. The scene representation with an MLP F_Θ is expressed as

$$F_\Theta(\mathbf{x}, \mathbf{d}) \rightarrow (\mathbf{c}, \sigma), \quad (2.1)$$

where Θ represents the parameters of the MLP, $\mathbf{x} = (x, y, z)$ is the spatial coordinate of a sample point, and $\mathbf{d} = (\theta, \phi)$ represents the viewing direction, where θ and ϕ are the azimuthal and polar viewing angles, respectively.

2.1.2 Volume rendering with radiance field

The rendering of color $C(\mathbf{r})$ of any camera ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$ passing through the scene is performed by using classical volume rendering [35] with camera position \mathbf{o} , viewing direction \mathbf{d} and near and far bounds t_n and t_f , respectively, using

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t) \cdot \sigma(\mathbf{r}(t)) \cdot \mathbf{c}(\mathbf{r}(t), \mathbf{d}) \cdot dt, \quad (2.2)$$

where $T(t)$ denotes the accumulated transmittance along the camera ray $\mathbf{r}(t)$ from t_n to t , i.e., the probability that the camera ray $\mathbf{r}(t)$ travels from t_n to t without being hitting any other particle, given by

$$T(t) = e^{- \int_{t_n}^t \sigma(\mathbf{r}(u)) \cdot du}. \quad (2.3)$$

These integrals, equation 2.2 and equation 2.3, can be computed discretely. NeRF used a stratified sampling approach, where sample points are drawn from a uniformly binned ray. Then, equation 2.2 and equation 2.3 can be approximated through

$$\hat{C}(\mathbf{r}) = \sum_{i=1}^m \alpha_i T_i \mathbf{c}_i, \quad (2.4)$$

where $\hat{C}(\mathbf{r})$ is the approximate volume-rendered color, α_i is the the opacity value at sample point i , and $T_i = e^{- \sum_{j=1}^{i-1} \sigma_j \Delta_j}$. The opacity value α_i is computed using

$$\alpha_i = 1 - e^{\sigma_i \Delta_i}, \quad (2.5)$$

where $\Delta_i = t_{i+1} - t_i$ is the distance between adjacent samples, i.e. the distance from sample point i to sample point $i + 1$.

For the rendering loss function \mathcal{L}_R to optimize the MLP parameters Θ , the squared error approach is used,

$$\mathcal{L}_R = \sum_{r \in \mathcal{R}} \|\hat{C}(\mathbf{r}) - C_{g.t.}(\mathbf{r})\|_2^2, \quad (2.6)$$

where $\|\cdot\|_2^2$ is squared Euclidean norm. The $C_{g.t.}(\mathbf{r})$ is the ground truth color of the training image pixel for camera ray \mathbf{r} and \mathcal{R} is the set of rays in each batch.

2.1.3 Improvements of the Neural Radiance Field

To improve the NeRF model ability to capture high-frequency detailed features. Mildenhall et al. [54] introduced two key enhancements: positional encoding and hierarchical volume sampling.

Positional encoding maps the \mathbf{x} and \mathbf{d} to high-frequency variables before input into the MLP, enabling NeRF to capture high-frequency detail in synthesized views. Hierarchical volume sampling improves MLP efficiency by using two models, a coarse and a fine one. The coarse model guides the refined sampling of the fine model, strategically selecting points for higher density regions. This enhances color accuracy in rendered images, with a loss function minimizing discrepancies between the models' predictions and ground truth colors. For further details, we refer to Mildenhall et al. [54].

2.2 Volume rendering

The goal of this section is to give a brief introduction to volume rendering concepts, such as surface rendering, isosurface, direct volume rendering, transfer function, and histogram. We utilize volume rendering to visualize the uncertainty for a scene. We refer to works by Preim and Botha [69] and Preim and Bartz [68] for further details.

Volumetric data are utilized in many application fields, for example medical imaging and physical simulation. Volumetric datasets consist of points in 3D space, each storing a set of values. For example, a set of points could be arranged on a 3D regular grid, which is an isotropically spaced rectilinear grid, with data values represented as scalars. Rendering volumetric data can be achieved through either direct volume rendering (DVR) or an indirect volume rendering (IVR) approach.

In IVR, an intermediate representation, such as surface mesh, is extracted from the dataset and then rendered. Surface rendering (SR) is one of the IVR methods that creates an isosurface. The marching cube algorithm is designed to construct the isosurface. An isosurface is a 3D surface representation based on an isovalue or constant scalar value selected by the user in a 3D grid,

$$f(x, y, z) = k, \quad (2.7)$$

where $f(\cdot)$ is a function in a 3D space, (x, y, z) are the coordinates of a point in 3D space, and k is a constant or isovalue. Figure 2.2 shows examples of an isosurface of the spine and pelvis extracted from computerized tomography (CT) data. In our research, we employed IVR to render the isosurface of objects within a scene.

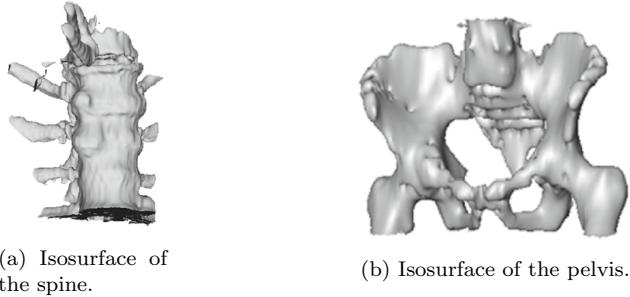


Figure 2.2: Images sourced from Preim and Botha [69]. Isosurfaces extracted from computerized tomography (CT) data. (a) A section of the spine in the breast region. (b) Pelvis extracted from CT data.

DVR is a rendering method used in computer graphics and medical imaging to generate images directly from volumetric data v [43], such as 3D medical data or scientific simulations. The DVR requires every point value of the volumetric data to be mapped to optical properties like color c and opacity α [52]. This mapping is accomplished using a transfer function,

$$TF(v) = (c, \alpha) \quad (2.8)$$

where $TF(v)$ is the transfer function for a volumetric data v . Figure 2.3 displays the volume-rendered high potential iron protein with two different transfer function settings. In Figure 2.3b, the iron protein with higher scalar values is assigned high opacity, while the remaining iron protein is set to low opacity. As a result, the volume rendering view clearly highlights the regions of the iron protein with high scalar values. The histogram, in the context of volume rendering, is the distribution of sample point values within the volumetric dataset. Understanding the histogram helps in setting up an effective transfer function. Figure 2.3b shows the histogram behind the transfer function, serving as a guide for determining settings such as assigning high opacity to regions of the iron protein with higher scalar values. In our research, we utilize DVR to visualize volumetric data filled with uncertainty scalar value. Unlike transfer functions where we mapped scalar values such as uncertainty to opacity α and color c , 2D transfer functions were designed to use color to show regions with varying uncertainty. Figure 2.4 shows the volume rendering of the ocean volumetric data with two different 2D transfer functions. Two and five colors are used in the 2D transfer functions to map corresponding sample points in the ocean volumetric data in Figure 2.4a and Figure 2.4b, respectively. In our research, we employ transfer functions along with histograms, and 2D transfer functions.

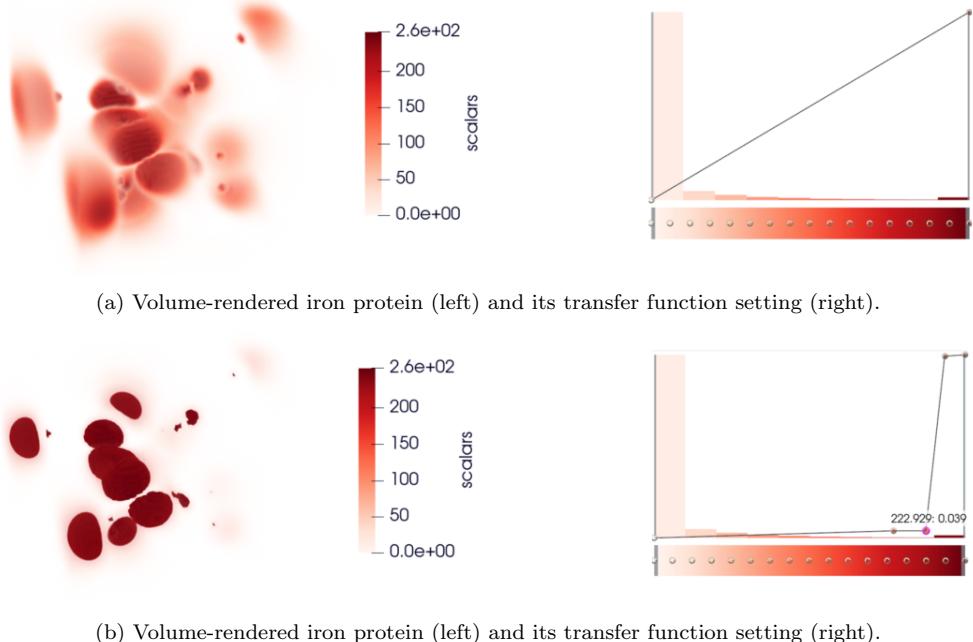


Figure 2.3: Direct volume rendering (DVR) of a high potential iron protein with different transfer function settings using Paraview [3]. The histogram is displayed behind the transfer function. In (b), higher scalar values of the high potential iron protein are assigned high opacity, while other scalar values are set to low opacity in the transfer function.

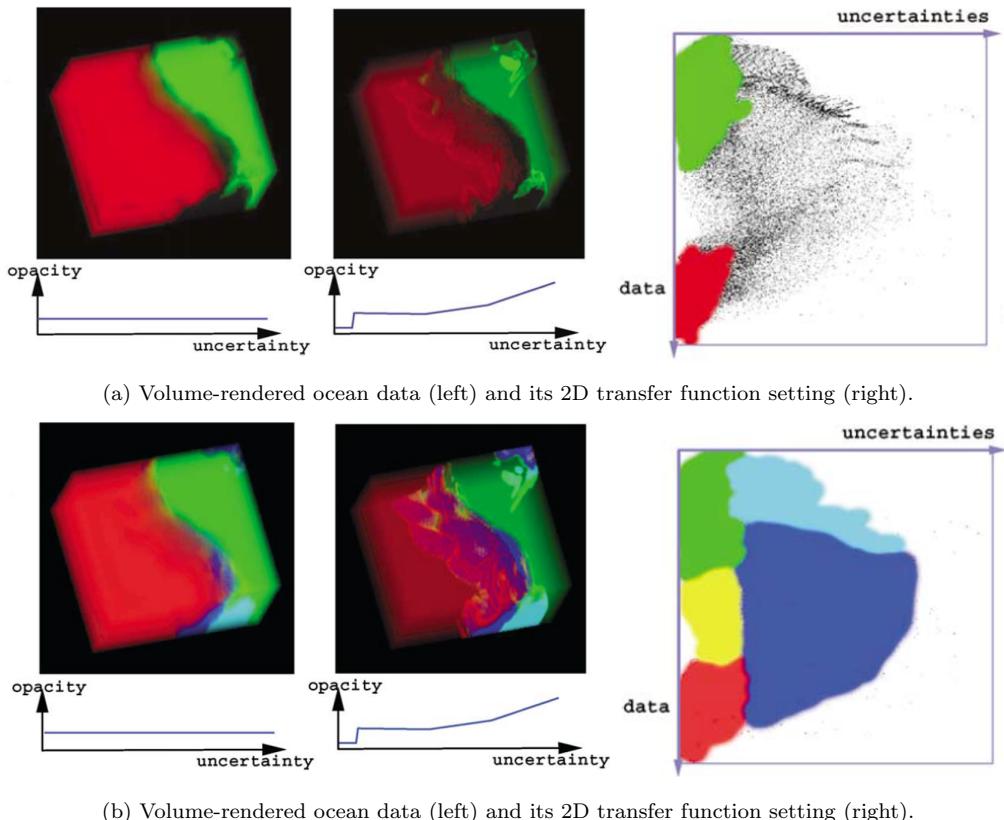


Figure 2.4: Images sourced from Djurcicov et al. [15]. Direct volume rendering (DVR) of ocean volumetric data with different 2D transfer function settings. Below each representation of the volume-rendered ocean volumetric data, the corresponding transfer function setting is shown.

Chapter 3

Related work

In this chapter, we review the relevant literature related to our work. We first explore uncertainty estimation in deep learning. Then, we delve into various techniques within Neural Radiance Field (NeRF) developed in previous studies that incorporate uncertainty. Finally, we discuss the related work in the field of uncertainty visualization.

3.1 Uncertainty estimation in deep learning

In machine learning, uncertainty plays a crucial role in quantifying the confidence or reliability of model predictions. Two main types of uncertainty are often considered: aleatoric and epistemic uncertainties [32, 27, 39]. Aleatoric uncertainty refers to uncertainty in the input data, for example, introduced by noise or random processes [36]. In contrast, epistemic uncertainty relates to the uncertainty associated with the model’s learnable parameters [36].

Uncertainty estimation has been extensively studied in the field of deep learning, with numerous studies exploring this area [23, 10, 48, 81, 18]. Furthermore, uncertainty estimation techniques have been applied to various computer vision tasks [71, 6, 85]. In earlier works, Bayesian Neural Networks (BNNs) [56, 57] were proposed to estimate uncertainty in both network weights and outputs by approximating their marginal distributions. BNNs have shown promise in addressing the problem of overfitting in neural networks by modeling uncertainty in weights. However, integrating BNNs into the NeRF architecture conflicts with the objectives of NeRF models, which are designed to learn a scene by overfitting training images.

Besides BNNs, researchers have proposed efficient strategies to incorporate uncertainty estimation into deep neural networks [75, 9] to capture the epistemic uncertainty. Among these approaches, Deep Ensembles [41] and Monte Carlo Dropout [16] are two popular methods known for their flexibility with various network architectures [4, 25, 5, 44]. Deep Ensembles generate a set of outputs from multiple neural networks trained independently using different parameter initializations. This ensemble of networks provides a means to estimate the output uncertainty. On the other hand, in Monte Carlo Dropout, stochastic dropout is introduced during inference. By performing multiple forward passes over the same input with different dropout configurations, the variance over the resulting set of outputs is used as the output uncertainty. Gustafsson et al. [20] introduced a framework for comparing uncertainty quantification methods between ensembles and Monte Carlo dropouts using real-life applications. According to this framework, they found that ensemble methods consistently provide more reliable and practically useful uncertainty estimates compared to Monte Carlo dropouts.

3.2 Uncertainty estimation in Neural Radiance Field

The Neural Radiance Field (NeRF) [54] method has achieved impressive results in synthesizing photo-realistic views of scenes. However, it lacks the ability to quantify the uncertainty associated with the model estimates. Understanding and quantifying uncertainties is relevant for making informed decisions and ensuring the reliability of prediction in various applications. Therefore, various uncertainty-aware NeRF methods have been proposed in the literature to address this limitation, such as Stochastic Neural Radiance Field (S-NeRF) [78], Conditional-Flow NeRF (CF-NeRF) [77], Probabilistic NeRF (ProbNeRF) [26], NeRF in the Wild (NeRF-W) [51], Self-NeRF [7], and Uncertainty-aware MedNeRF (UMedNeRF) [28]. These proposed approaches present challenges for users attempting to utilize their uncertainty estimations methodology because implementing them requires substantial modifications to the original NeRF architecture, and they lack direct human interaction. Recursive Neural Radiance Field (Recursive-NeRF) [87] is an efficient NeRF approach to synthetic scenes that leverages a level-of-detail strategy, and it is akin to the original NeRF model. It learns uncertainties that indicate the quality of predicted color and volumetric intensity at each level. Predictions with higher uncertainty progress to larger neural networks. This splits the workload adaptively, avoiding unnecessary parameter increases. Density-aware NeRF Ensembles [80] train an ensemble of NeRF neural networks with different parameter initialization and combine their predictions through averaging.

Apart from general uncertainty-aware methods, several methods focusing on localized uncertainty assessment have also been proposed, such as Lee et al. [42], ActiveRMAP [88], and NeUNBV [33] and NeurAR [73]. Localized uncertainty assessment involves evaluating uncertainty within specific or limited angles in scene reconstruction, focusing on understanding uncertainty in particular viewpoints during the scene reconstruction process. Figure 3.1 shows the pipeline of integration of NeRF incorporates with the localized uncertainty assessment in the scene reconstruction, the initial camera directions and positions, and the example of the uncertainty map of a specific view of the scene. Figure 3.2 shows the localized uncertainty assessment visualizations of a scene to guide measurement acquisition scene reconstruction, the view with the highest uncertainty is used as a candidate for acquiring additional training images. These methods lack direct human interaction, emphasizing the necessity for interactive visualization to enhance users in exploring and analyzing uncertainty, facilitating a better understanding of NeRF model uncertainty. Our approach focuses on user understanding of the model uncertainty through human interaction with uncertainty visualization.

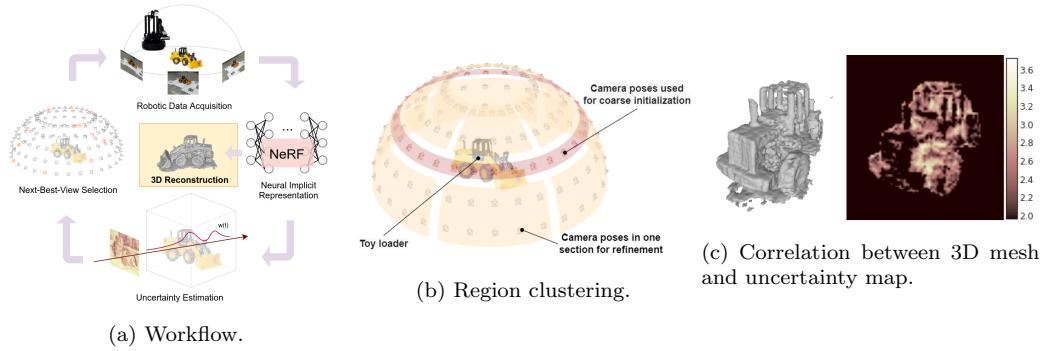


Figure 3.1: Images sourced from Lee et al. [42]. (a) A pipeline of the NeRF integrated with the localized uncertainty assessment, guiding additional scene image acquisition for enhanced scene reconstruction. (b) The camera directions and positions on the upper hemisphere of the scene were used for training NeRF initially. (c) The three-dimensional (3D) mesh and the uncertainty map of a specific view of the scene.

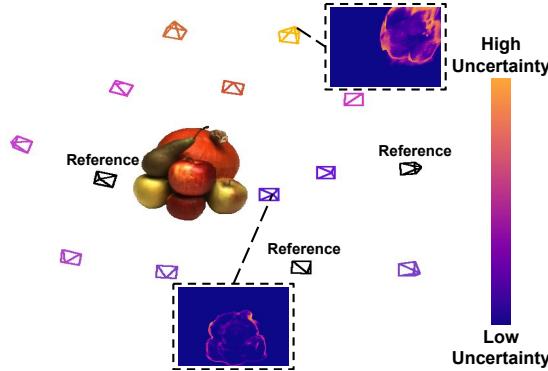


Figure 3.2: Image sourced from NeU-NBV [33]. NeU-NBV employs uncertainty estimation integrated into NeRF to guide measurement acquisition. The NeU-NBV selects the most informative view candidate with the highest uncertainty for acquiring additional scene images, thereby enhancing scene reconstruction.

Our work uses two uncertainty estimation methods that capture NeRF model uncertainty. The first method is adopted from Recursive-NeRF [87], whereas the second method is adopted from Density-aware NeRF Ensemble [80] and Deep Ensembles [41]. The former estimates uncertainty from a single deterministic neural network (NN). While, the latter estimates uncertainty from multiple NNs with varying parameter initialization that captures epistemic uncertainty. These methods are selected to estimate uncertainty because they can be easily implemented and integrated into the original NeRF architecture [54]. The integration of Recursive-NeRF’s uncertainty estimation approach into the original NeRF architecture takes place without introducing substantial model complexity, preventing unnecessary increases in network parameters and maintaining a manageable model size. On the other hand, Deep Ensembles has demonstrated superior performance in various domains [13, 21, 60, 49, 86].

3.3 Uncertainty visualization

Uncertainty visualization provides users with a more complete and accurate understanding of the data for analysis and decision-making [62]. It is widely applied in the field of medical imaging. Figure 3.3 shows an example of subtle differences in the artery and brain tumor representations by changing the visualization parameters in the transfer function settings. In this uncertainty visualization alongside the confidence level could guide doctors and surgeons to improve surgical results. This section begins with an overview of uncertainty’s role in the visualization pipeline, followed by a discussion of various visualization approaches by Pang et al. [62] and Raidou [72]. Finally, we discuss several uncertainty visualization techniques.

Weiskopf [84] provides an overview of the role of uncertainty in the visualization pipeline, as shown in Figure 3.4. The figure shows that uncertainty exists at different stages in the visualization pipeline, including data acquisition, filtering and transformations, visual mapping and rendering, and the perception and cognition of users, along with the interaction between users and these stages. Weiskopf [84] emphasizes the need for different stages to be made uncertainty-aware and the uncertainty can propagate through the pipeline. Our uncertainty visualization system fits this visualization pipeline and facilitates users’ exploration of the scenes, aiding in understanding NeRF model uncertainty. For example, NeRF learns parameters that transform the data in training images to represent the three-dimensional (3D) scene, introducing uncertainty in this transformation. Then, we visualize the synthesized scene through visual mapping and rendering, allowing user interaction to obtain uncertainty information. Visual perception and cognition play

crucial roles in visualization. However, assessing uncertainty in perception and cognition is challenging [84]. Assessing uncertainty in the perception and cognition of users is out of the scope of our research.

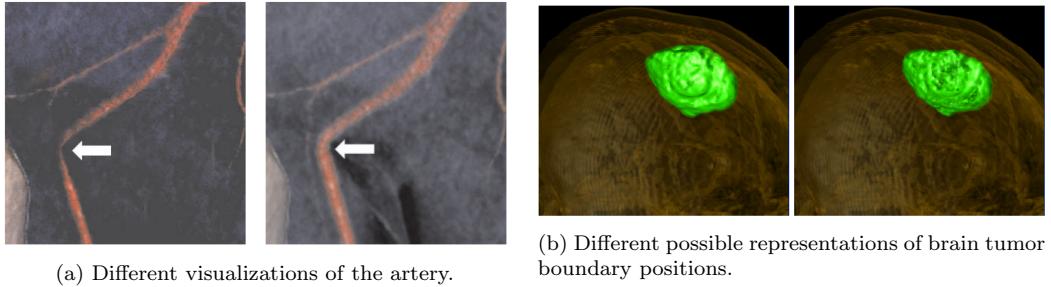


Figure 3.3: (a) Images sourced from Lundström et al. [46], displays two possible artery visualizations achieved by adjusting the visualization parameters. (b) Images sourced from Potter [66], displays two possible brain tumor tissue boundary positions resulting from varying visualization parameters.

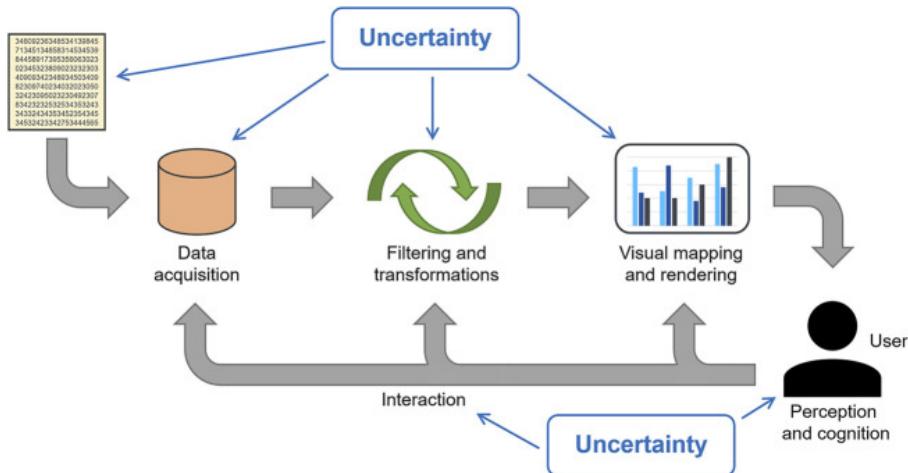


Figure 3.4: Image sourced from Weiskopf [84]. Visualization pipeline including uncertainty. The pipeline consists of different stages: data acquisition, filtering and transformations, visual mapping and rendering, perception and cognition of the users, and the interaction between users and previous stages.

Pang et al. [62] classified the uncertainty visualization approaches into three classes at a high level: overloading, side-by-side comparison, and seamless integration. Overloading involves graphically displaying the uncertainty on top of the conventional visualization within the same image. In a side-by-side comparison, two or more similar uncertainty visualizations of a dataset are rendered side-by-side; they may differ due to the use of different rendering algorithms or different filtering functions. Unlike the overloading approach, the seamless integration approach directly integrates the uncertainty with the data in the unified visualization rendering. At a lower level, Pang et al. categorized several uncertainty visualization techniques into several categories, including glyphs, animation, and opacity. Here, we briefly discuss opacity, which is relevant to our research. Opacity is used to indicate uncertainty in a volumetric dataset, where the less uncertain regions could be rendered as an opaque area, while the high uncertainty regions could be rendered as a transparent area. This opacity technique is applied in the drug molecule volumetric dataset to indicate intra-molecule positional uncertainty [74]. In our research, we implement the overloading

and side-by-side comparison approaches.

In the research work of Raidou [72], Raidou categorizes the visualization of uncertainty methods into several categories, as shown in Figure 3.5. In the following, we discuss the free graphical variables category because it is related to our research. Color is commonly used in uncertainty visualization because it is not only easy to perceive but also enables quick and intuitive interpretation, making it a highly effective means for conveying uncertainty information in visualizations. Figure 3.6 illustrates the representation of the uncertain surface using color. Hengl et al. [24] utilized Hue-Saturation-Intensity color model for visualization of uncertainty. Coninx et al. [12] presented a method to visualize uncertain scalar data fields by combining color scale visualization techniques with animated. Apart from color, Djurcicov et al. [15] incorporated the uncertainty information directly into the volume rendering. Specifically, two-dimensional transfer functions were used to show regions with varying uncertainty with different colors. Adding holes, noise, texture, and speckles to the locations of uncertainty in the post-processing process was proposed. However, adding these free graphical variables, including color and transparency, could be difficult to interpret [72]. In our research, we employ color and opacity as free graphical variables in our uncertainty visualization.

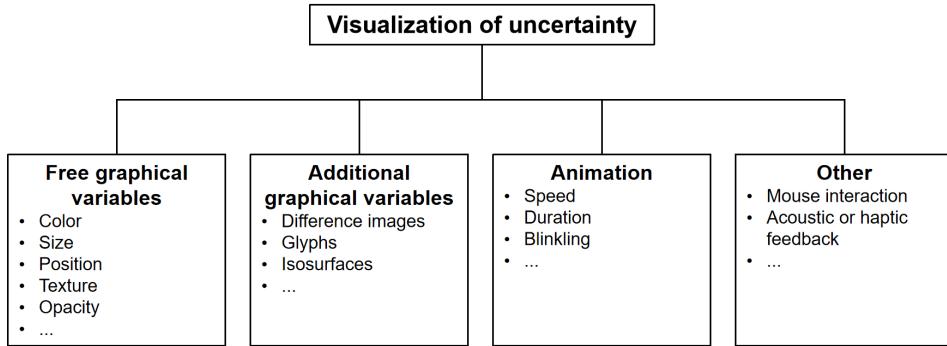


Figure 3.5: A categorization of different visualization of uncertainty methods. The figure is adapted from Raidou [72]. Refer to Raidou [72] for a comprehensive list of visualization of uncertainty methods.

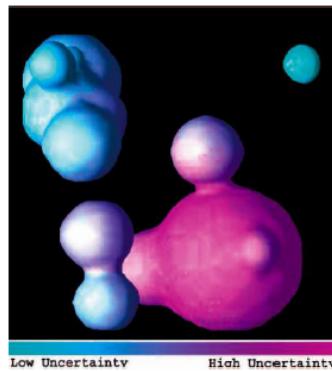


Figure 3.6: Image sourced Grigoryan and Rheingans [19]. Color indicates uncertain surface.

Visualizing uncertainty data can be accomplished with different types of graphs and plots. Potter et al. [67] introduced summary plots to visualize summary statistics and uncertainty. Cedilnik et al. [11] proposed a method for visualizing uncertainty in 2D scalar maps using grid lines. The levels of distortion to the grid lines are determined by the uncertainty of the underlying

data values (Figure 3.7). In the case of 3D volumetric data, the statistics values, such as mean, standard deviation, or variance, can be computed on a voxel level, given 3D multi-valued scalar data, and they can then be mapped to color and opacity using a transfer function. For example, Djurcicov et al. [15] demonstrates the use of volume rendering to visualize 3D measurements of water temperature and salinity in ocean data (Figure 3.8a). Uncertainty is quantified through Monte Carlo simulations based on ocean models, and it is represented as a scalar-valued variance. Figure 3.8b and Figure 3.8c display the uncertainty visualizations of ocean and brain datasets using the direct volume rendering method with a transfer function. This figure shows the visualizations of normalized standard deviation, which is quantified as uncertainty, the red regions represent high uncertainty, while the blue regions represent low uncertainty. The drawback of these uncertainty visualizations is they produce confusing results since depth information is lost. This occurs because the uncertainty values throughout the rays are projected onto the projection plane. Another uncertainty visualization technique involves combining the isosurface methods with volume rendering methods, as shown in Figure 3.9. This is considered a seamless integration uncertainty visualization approach. Figure 3.9a and Figure 3.9b show an isosurface dataset bracketed with volume-rendered regions indicating uncertainty around the isosurface. Figure 3.9c shows the isosurface of a magnetic resonance imaging dataset surrounded by a volume-rendered region of low opacity to indicate uncertainty in surface position. In our uncertainty visualization design, we employ a fundamental volume rendering technique to visualize the volumetric uncertainty data, similar to Figure 3.8.

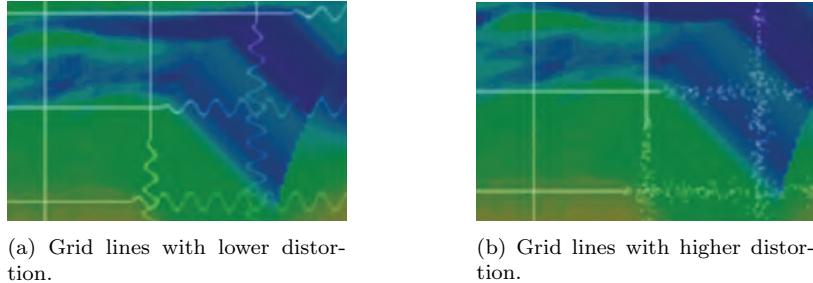


Figure 3.7: Images sourced Cedelnik et al. [11]. Scalar-valued uncertainty is used to introduce disturbances in grid lines with varying levels of distortion.

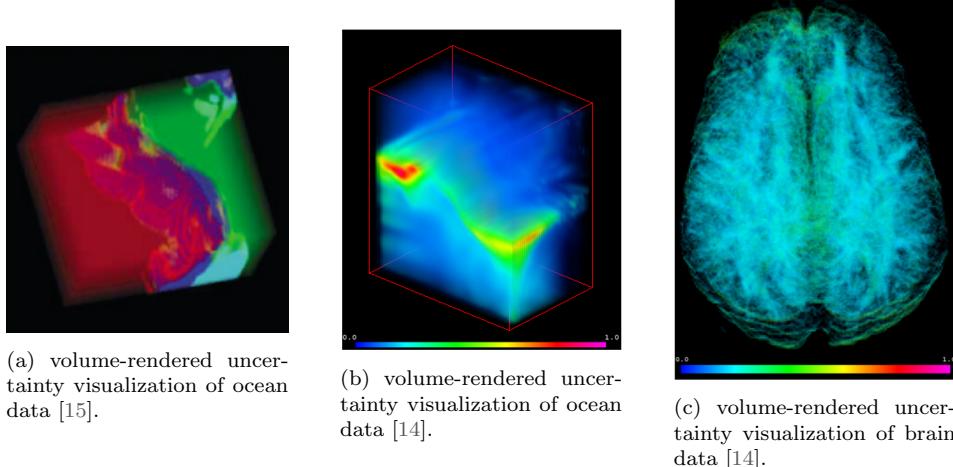


Figure 3.8: Volume-rendered images of (a,b) ocean data and (b) brain data. The normalized standard deviation is linearly mapped to opacity and color in (b) and (c). In (b), red opaque regions represent high standard deviation, while blue transparent regions represent low standard deviation. In (c), voxels with normalized standard deviation less than 0.15 were rendered transparent for a clearer view of regions with higher standard deviation.

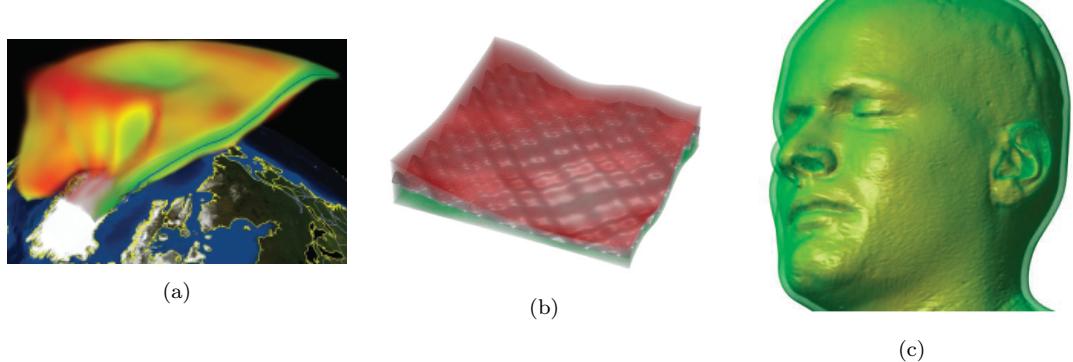


Figure 3.9: (a) Volume visualization of isosurface crossing in uncertain 3D temperature field in the exosphere above Europe and the North Atlantic Ocean [65]. (b) Isosurface of a synthetic dataset and (c) isosurface of an magnetic resonance dataset surrounded by a volume-rendered region representing uncertainty [34].

Chapter 4

NeRFDeltaView

This chapter introduces the NeRFDeltaView (Neural Radiance Field Delta View) system, designed to explore uncertainty in Neural Radiance Field (NeRF) models. The system is composed of uncertainty estimation methods integrated with the NeRF model and interactive uncertainty visualizations. This facilitates users to make informed decisions based on the visualizations with uncertainty predicted by the models. Users can leverage uncertainty information estimated by NeRF to identify limitations in the model uncertainty, aiding in the improvement of the model architecture, as well as in exploring the correlation between viewing direction and NeRF model uncertainty.

Several potential user questions may be posed by the users when analyzing uncertainty:

- **Q1:** How do the scene features relate to NeRF model uncertainty?
- **Q2:** Can we identify the limitations of NeRF model based on the model uncertainty analysis?
- **Q3:** Is there a correlation between the amount and viewing direction of training images and NeRF model uncertainty?

The first question (**Q1**) is centered around interpreting the relationship between scene features and NeRF model uncertainty. The scene features could be referred to as elements or characteristics of objects within the synthesized scenes. For instance, these scene features could be intricate details such as irregular surfaces or obstructed objects. The objective is to identify the scene features where the model struggles, guiding to understand the model uncertainty. This question triggers an exploration into why these particular features pose challenges for the model predictions. Exploring and analyzing which scene features contribute to uncertainty is crucial for guiding decisions on additional training images, particularly for specific viewing directions, improving model architecture design, or refining hyperparameters, such as the learning rate.

The second question (**Q2**) is focused on examining the model's limitations by analyzing its behavior in relation to uncertainty. By examining the uncertainty distribution across synthesized scenes through visualizations, we gain insights into regions in the synthesized scene where these methods may exhibit shortcomings. For instance, a crucial aspect of this investigation involves understanding why a particular region did not exhibit high uncertainty when it was expected to. This analysis gains insights into the model's behavior and limitations concerning uncertainty.

The third question (**Q3**) investigates whether a correlation exists between the amount and viewing direction of training images and the model uncertainty. It aims to explore how variations in viewing direction and the quantity of training images influence the pattern of uncertainty estimation within a scene. This analysis contributes to a deeper understanding of the relationship between these factors and the reliability of the models. Hence, it enables users to make informed

decisions, such as adding training images in specific viewing directions for further optimization of model parameters and enhance overall performance.

The overview of the NeRFDeltaView system is depicted in Figure 4.1. This system comprises two main components: uncertainty estimation methods and uncertainty visualization. Two NeRF model variants are employed for the uncertainty estimation methods. Questions **Q1**, **Q2**, and **Q3** emphasize the necessity for different visualizations to visualize the uncertainties estimated by these methods. Question **Q1** highlights that volume rendering visualization is required to allow users to explore the scene features in the synthesized scene where the model encounters challenges, facilitating our understanding of the model uncertainty. An overview visualization that displays a comprehensive perspective on the overall uncertainty patterns around a scene integrated with the amount and viewing direction of training images serves the objective of Question **Q3**. Exploration of the uncertainty could be performed using both the volume rendering and overview visualizations to identify the limitations of the model, thereby addressing Question **Q2**.

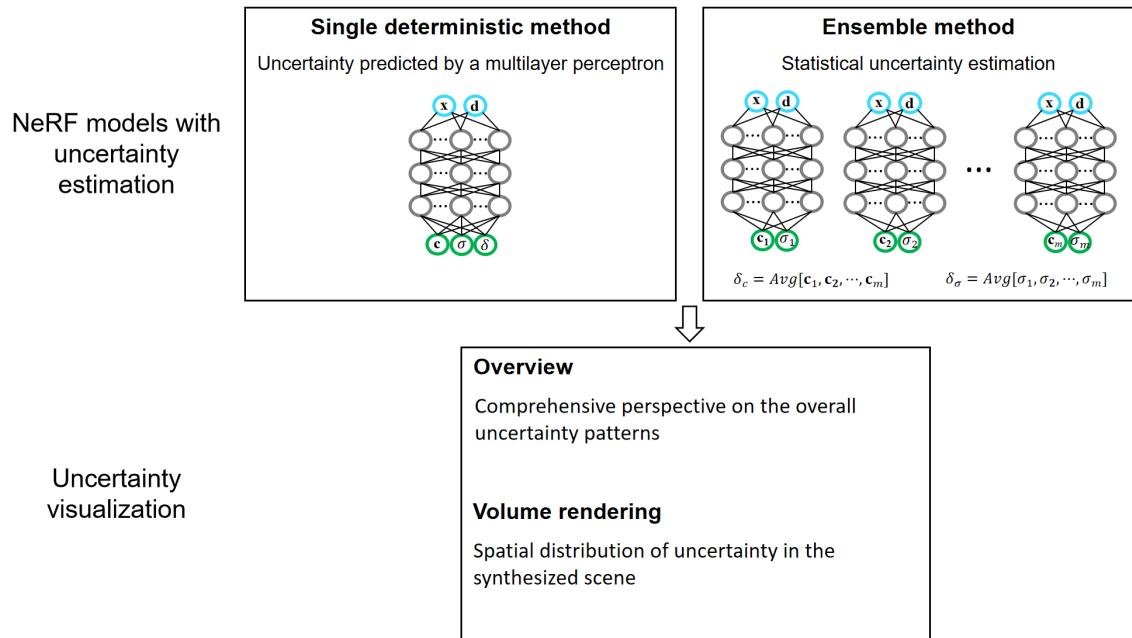


Figure 4.1: Overview of the NeRFDeltaView system. The top row is the NeRF models integrated with uncertainty estimation. $\text{Avg}[\cdot]$ is an averaging method to estimate the uncertainty δ . The bottom row is the uncertainty visualization technique to visualize the uncertainty estimated by these models.

In the following, we describe briefly both NeRF models with uncertainty estimation and the uncertainty visualization components. The first uncertainty estimation method is the single deterministic approach. We adopt the uncertainty estimation method from Recursive Neural Radiance Field (Recursive-NeRF) [87]. This approach estimates the uncertainty of the model's prediction, i.e., the output estimated uncertainty by the model represents the capability and confidence of the model to synthesize the scene accurately. It outputs an uncertainty prediction based on a single forward pass within a deterministic network [17]. The deterministic network, represented as a multilayer perceptron (MLP) estimating color \mathbf{c} , volume density σ , and an uncertainty δ from the input spatial coordinate of a sample point \mathbf{x} and the viewing angle direction \mathbf{d} , is illustrated in Figure 4.1. An uncertainty δ is output by the MLP F_Θ for a sample point,

$$F_\Theta(\mathbf{x}, \mathbf{d}) \rightarrow (\mathbf{c}, \sigma, \delta), \quad (4.1)$$

where Θ represents the parameters of the MLP. This differs from the original NeRF formulation (Equation 2.1), which outputs only the color \mathbf{c} and volume density σ . Details of this uncertainty estimation method are elaborated in Chapter 5.

The second method is the ensemble approach, which combines the predictions of several different deterministic networks during inference [17], inspired by Density-aware NeRF Ensemble [80] and Deep Ensembles [41]. This approach estimates the uncertainty of the model based on the estimated color \mathbf{c} and volume density σ from several different deterministic networks, represented as MLPs, with varying parameter initialization that captures epistemic uncertainty. Figure 4.1 illustrates each deterministic network estimating \mathbf{c} and σ from the input \mathbf{x} and \mathbf{d} ,

$$\begin{aligned} F_{\Theta_1}(\mathbf{x}, \mathbf{d}) &\rightarrow (\mathbf{c}_1, \sigma_1) \\ F_{\Theta_2}(\mathbf{x}, \mathbf{d}) &\rightarrow (\mathbf{c}_2, \sigma_2) \\ &\vdots \\ F_{\Theta_{m-1}}(\mathbf{x}, \mathbf{d}) &\rightarrow (\mathbf{c}_{m-1}, \sigma_{m-1}) \\ F_{\Theta_m}(\mathbf{x}, \mathbf{d}) &\rightarrow (\mathbf{c}_m, \sigma_m) \end{aligned} \tag{4.2}$$

where m is the number of deterministic networks. The uncertainty δ is then estimated from these multiple sets of \mathbf{c} and σ using an averaging approach. Notably, two uncertainties are estimated for a sample point: color uncertainty δ_c and density uncertainty δ_σ ,

$$\begin{aligned} \delta_c &= \text{Avg}[\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{m-1}, \mathbf{c}_m] \\ \delta_\sigma &= \text{Avg}[\sigma_1, \sigma_2, \dots, \sigma_{m-1}, \sigma_m] \end{aligned} \tag{4.3}$$

where $\text{Avg}[\cdot]$ is an averaging method used to estimate the uncertainty δ , such as mean. This uncertainty estimation method is detailed in Chapter 5.

The uncertainty visualization component comprises two subcomponents: overview and volume rendering. In the overview subcomponent, we generate a visualization that can be used to comprehensively visualize the uncertainty patterns in a scene. This approach allows users to grasp the overall uncertainty pattern and then delve into specific viewing directions through volume rendering visualization. Furthermore, the uncertainty pattern in the overview visualization can be analyzed to investigate the correlation between the amount and viewing of training images and NeRF model uncertainty (**Q3**). In the volume rendering subcomponent, the volume-rendered uncertainty visualization is generated for users to explore specific scene features in the synthesized scene (**Q1**). Users can identify limitations in the model based on the model uncertainty analysis through the exploration of the uncertainty distribution in the synthesized scene (**Q2**). For instance, high uncertainties may not appear in obstructed regions, suggesting that the model faces challenges in estimating uncertainty correctly for such scenes. Chapter 6 provides a detailed explanation of the uncertainty visualization component within the system. Chapter 7 presents the results of our uncertainty visualization tool in the NeRFDeltaView system using three distinct use cases.

Chapter 5

Uncertainty estimation

In this chapter, we will explore two uncertainty estimation methods, both of which are integral components of NeRFDeltaView system, as explained in Chapter 4. The first method is the Uncertainty Neural Network, where the uncertainty estimation method is adopted from Recursive Neural Radiance Field (Recursive-NeRF) [87]. The second approach involves using an ensemble architecture for uncertainty estimation, inspired by Density-aware NeRF Ensemble [80] and Deep Ensembles [41].

5.1 Uncertainty Neural Network

This section introduces our first uncertainty estimation method for the Neural Radiance Field (NeRF), which is the single deterministic method, as explained in Chapter 4. We call this single deterministic method as Uncertainty Neural Network (Uncertainty NN). This integration of NeRF with uncertainty estimation is based on the uncertainty prediction approach introduced in Recursive Neural Radiance Field (Recursive-NeRF) [87]. We chose this approach because it can be easily implemented and integrated into the original NeRF architecture [54]. Additionally, implementing Recursive-NeRF’s uncertainty estimation method in Uncertainty NN does not introduce substantial model complexity, thus avoiding unnecessary parameter increments and maintaining a manageable model size. This approach estimates uncertainty value by the Uncertainty NN, which represents the confidence of the model in accurately synthesizing the scene.

Figure 5.1 illustrates the pipeline of the Uncertainty NN. We first discuss the fully-connected multilayer perceptron (MLP) for the Uncertainty NN, followed by an explanation of the training process for the MLP. The architecture of the Uncertainty NN’s MLP is depicted in Figure 5.2. In addition to estimating the color \mathbf{c} and volume density σ as in NeRF [54], the MLP also outputs the uncertainty δ , as formulated in Equation 4.1. The MLP comprises 8 hidden layers activated with Rectified Linear Unit (ReLU) activation function [1] with 128 neurons in the first 7 hidden layers and 64 neurons in the last hidden layer. The architecture is similar to the original NeRF MLP [54], except it has 128 neurons per layer instead of 256 neurons per layer. This improvement from 128 to 256 neurons per layer is insignificant based on observations from Instant Neural Graphics Primitives (Instant-NGP) [55]. Like in the original NeRF MLP architecture, the skip connection approach [63] is employed to enhance the information flow throughout the network, improving the training results.

To train the MLP of Uncertainty NN, the classical volume rendering technique [35] is used to composite the colors and volume densities into an image. Subsequently, the rendering loss \mathcal{L}_R is computed using Equation 2.6, as explained in Chapter 2. In addition, following the uncertainty loss formulations from the Recursive-NeRF [87]. We compute the error loss

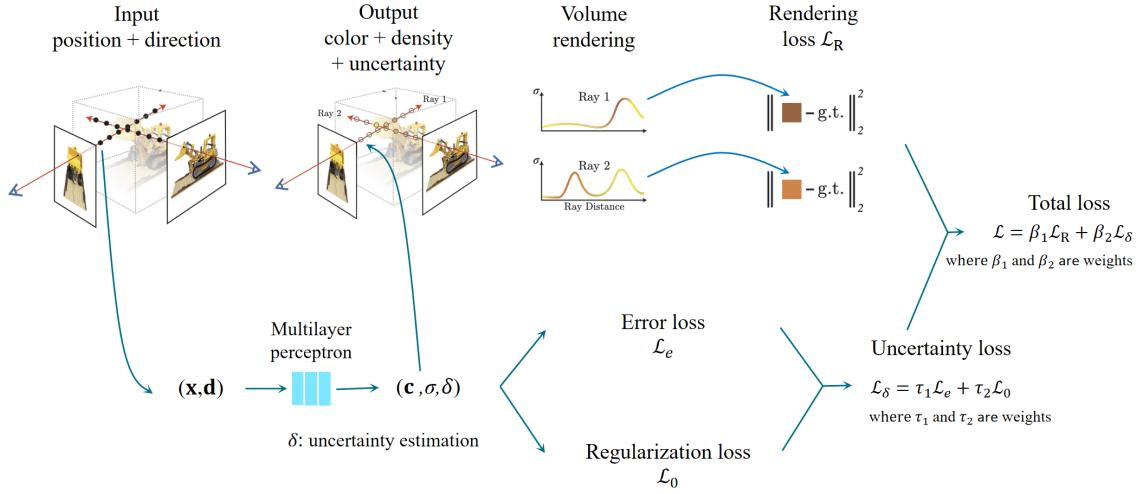


Figure 5.1: The overview of the Uncertainty Neural Network model pipeline. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value. Part of the figure is adapted from NeRF [54].

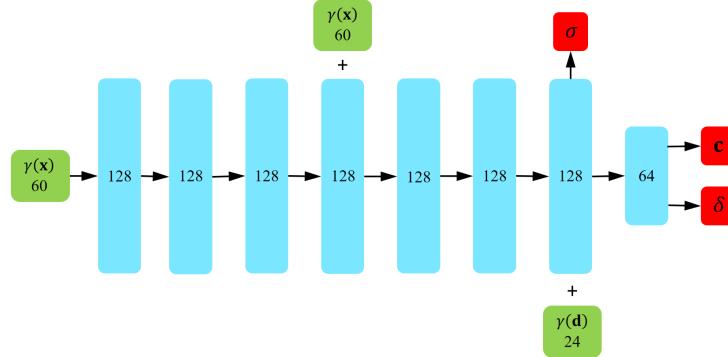


Figure 5.2: A visualization of the fully-connected multilayer perceptron for the Uncertainty Neural Network. The neural network is similar to NeRF [54] with an additional estimation of uncertainty δ , inspired by Recursive-NeRF [87]. Input vectors are in green, hidden layers in cyan, output vectors in red, and the number in each block represents the vector's dimension, for example, the number in each hidden layer represents the number of neurons. The "+" signifies vector concatenation. The $\gamma(\mathbf{x})$ and $\gamma(\mathbf{d})$ are the positional encoding of the spatial coordinate of the sample point and viewing angle direction of the ray, respectively.

$$\mathcal{L}_e = \sum_{r \in \mathcal{R}} \sum_{i=1}^N \max(E(\mathbf{r}) - \delta_{\mathbf{r},i}, 0), \quad (5.1)$$

$$E(\mathbf{r}) = \left\| \hat{C}(\mathbf{r}) - C_{\text{g.t.}}(\mathbf{r}) \right\|_2^2, \quad (5.2)$$

where $E(\mathbf{r})$ is the squared error of ray \mathbf{r} , N is the number of sampling points for ray \mathbf{r} , and $\delta_{\mathbf{r},i}$ is the predicted uncertainty at sample point i of ray \mathbf{r} . $\hat{C}(\mathbf{r})$ is the volume-rendered color of ray \mathbf{r} . This loss involves the computation of the squared error of the pixel to provide supervision for δ . The objective is to ensure that if a pixel shows a high error, the corresponding δ of the sample points should also be high. Therefore, the δ is penalized in a manner similar to the rendering loss \mathcal{L}_R of the pixel.

To prevent δ from exploding, i.e., δ is encouraged to be as close to zero as possible. The regularization loss \mathcal{L}_0 is introduced,

$$\mathcal{L}_0 = \sum_{\mathbf{r} \in \mathcal{R}} \sum_{i=1}^N \max(\delta_{\mathbf{r},i}, 0). \quad (5.3)$$

The uncertainty loss \mathcal{L}_δ is determined by taking a weighted sum of \mathcal{L}_e and \mathcal{L}_0 ,

$$\mathcal{L}_\delta = \tau_1 \mathcal{L}_e + \tau_2 \mathcal{L}_0 \quad (5.4)$$

where τ_1 and τ_2 are weights. In Recursive-NeRF [87], these weights are set to $\tau_1 = 1.0$ and $\tau_2 = 0.01$. Recursive-NeRF used \mathcal{L}_δ with unbalanced weights τ_1 and τ_2 so that the network applies a larger penalty, implying that the model is more sensitive to deviations in uncertainty predictions when they are lower than expected. Finally, we obtain the total loss \mathcal{L} by taking the weighted sum of \mathcal{L}_R (Equation 2.6) and \mathcal{L}_δ ,

$$\mathcal{L} = \beta_1 \mathcal{L}_R + \beta_2 \mathcal{L}_\delta \quad (5.5)$$

where β_1 and β_2 are weights. These weights are set to $\beta_1 = 1.0$ and $\beta_2 = 0.1$ in the Recursive-NeRF [87] based on the analysis of the model performance and sensitivity to uncertainty measures. This total loss \mathcal{L} is utilized for learning the parameters in the MLP. After training the Uncertainty NN model, we utilize the trained model’s MLP to estimate the uncertainty δ for a given point in a three-dimensional (3D) grid.

5.2 Ensemble

In this section, we explain our second uncertainty estimation method for NeRF, which we denote as an Ensemble model, inspired by Density-aware NeRF Ensemble [80] and Deep Ensemble [41]. We select this architecture because it can be seamlessly implemented and integrated into the original NeRF architecture [54] and ensemble architecture has demonstrated superior performance in various domains [13, 21, 60, 49, 86]. This Ensemble model estimates the NeRF model uncertainty based on estimated color \mathbf{c} and volume density σ from several neural networks with varying parameter initialization that captures epistemic uncertainty.

Figure 5.3 illustrates the Ensemble model pipeline. We first discuss the architecture of the fully-connected MLPs for model, followed by an explanation of the training process for the MLPs. Instead of training a single MLP, we employ multiple MLPs, including a primary MLP and several secondary MLPs. The parameter initialization differs for these MLPs. We design the primary MLP to have a larger number of parameters compared to the secondary MLPs (Figure 5.4). This approach allows us to generate a rendered image using the primary MLP with better fidelity and quality, which will be used as a synthesized scene view. The primary MLP is similar to the original NeRF MLP [54], but it employs 128 neurons per layer instead of 256 neurons per layer, as performance observations from Instant-NGP [55] showed comparable results for both architectures. The secondary MLPs are designed with fewer parameters to introduce diversity into the Ensemble model by varying the architecture of MLPs. Additionally, reducing the number of parameters in the secondary MLPs ensures efficient utilization of available hardware resources during both training and inference phases, striking a balance between model complexity, computational feasibility, and training time.

To train the MLPs in the Ensemble model, each MLP estimates its color \mathbf{c} and volume density σ as shown in Figure 5.3. For example, $(\mathbf{c}, \sigma)_P$ is the color and volume density estimated by the primary MLP, where subscript P means primary MLP. $(\mathbf{c}, \sigma)_{S1}$, $(\mathbf{c}, \sigma)_{S(n-1)}$, and $(\mathbf{c}, \sigma)_{Sn}$ are the color and volume density estimated by the first, $(n-1)$ -th, and n -th secondary MLPs, respectively,

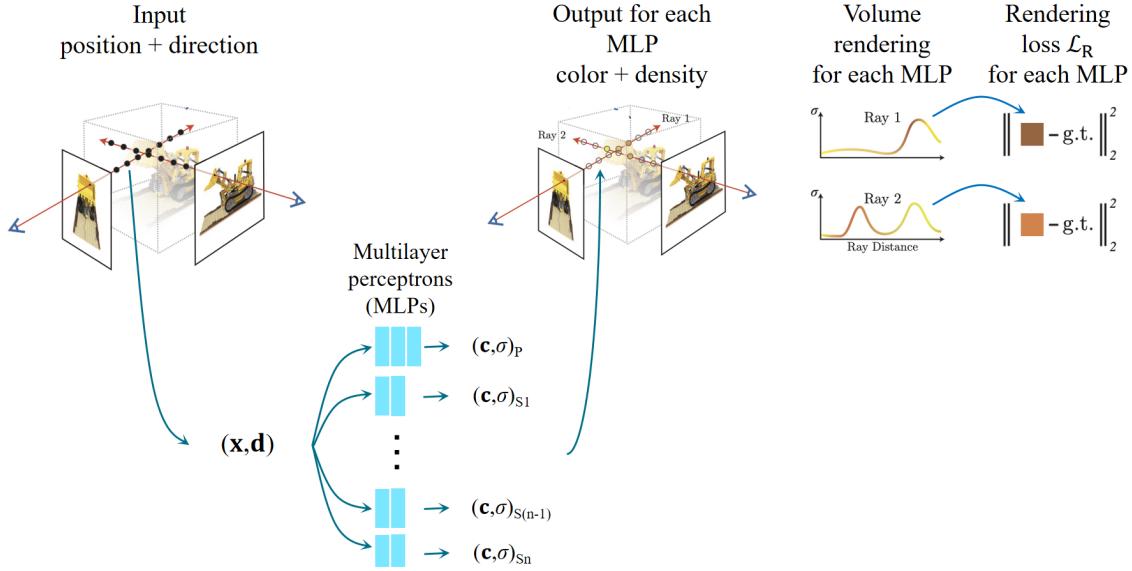


Figure 5.3: The overview of the Ensemble network architecture pipeline. The $(\mathbf{c}, \sigma)_P$ is the color \mathbf{c} and the volume density σ estimated by the primary MLP. Whereas, the $(\mathbf{c}, \sigma)_{S1}, (\mathbf{c}, \sigma)_{S(n-1)}$, and $(\mathbf{c}, \sigma)_{Sn}$ are estimated by the first secondary MLP, $(n-1)$ -th secondary MLP, and n -th secondary MLP, respectively. The abbreviation "g.t." stands for "ground truth," which refers to a pixel's color's actual value. Part of the figure is adapted from NeRF [54].

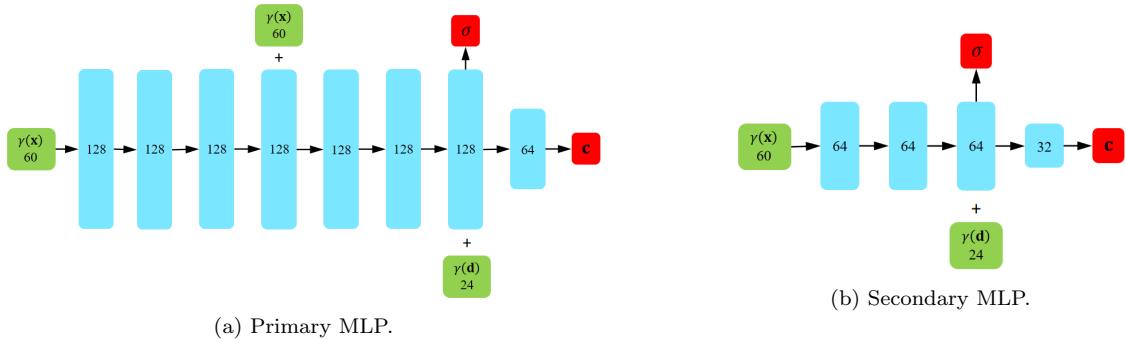


Figure 5.4: Primary and secondary multilayer perceptrons for the Ensemble architecture. Input vectors are in green, hidden layers in cyan, output vectors in red, and the number in each block represents the vector's dimension, for example, the number in each hidden layer represents the number of neurons. The "+" signifies vector concatenation. The $\gamma(\mathbf{x})$ and $\gamma(\mathbf{d})$ are the positional encoding of the spatial coordinate of the sample point and viewing angle direction of the ray, respectively.

where the subscript S means secondary MLP. Afterward, classical volume rendering techniques [35] are employed to composite the colors and volume densities into an image for each MLP. The rendering losses \mathcal{L}_R are computed using Equation 2.6 for each MLP to update the parameters within each MLP, respectively.

In the following, we discuss our approach for estimating the color and density uncertainties. For Density-aware NeRF Ensemble [80] and Deep Ensemble [41], the predictions from the ensemble models are combined through averaging, such as mean, and uncertainty is expressed as the variance over the ensemble individual MLP predictions. However, computing variance might not provide a representative uncertainty in our case because we are dealing with color space, which is a

high-dimensional space, and it is not well defined by an isotropic Gaussian. An isotropic Gaussian refers to a Gaussian distribution where the variance is the same in all directions. Therefore, we take a heuristic approach in which the uncertainty is represented by the mean of the pairwise Euclidean distances because it considers a measure of variability between sample points within the color space. Firstly, the pairwise Euclidean distances are computed between each MLP in the Ensemble for both color and volume density. After that, the means of color and density pairwise Euclidean distances are used to represent the color and volume density uncertainties, respectively. Figure 5.5 illustrates the estimation of color uncertainty δ_c and volume density uncertainty δ_σ by averaging the pairwise Euclidean distances for both color and volume density predictions made by the MLPs for a given point in a 3D grid. Two separate 3D grids are created to store δ_c and δ_σ separately. It is worth noting that while the model is initially trained with positional encoding of the input viewing direction, this is set to zero during 3D grid generation because we sample the points in a grid that is independent of the viewing direction. However, this approach disregards the influence of viewing direction on color estimation. Consequently, the estimated color uncertainty may not fully represent the true uncertainty of color in the scene, especially considering the scene with complex lighting conditions.

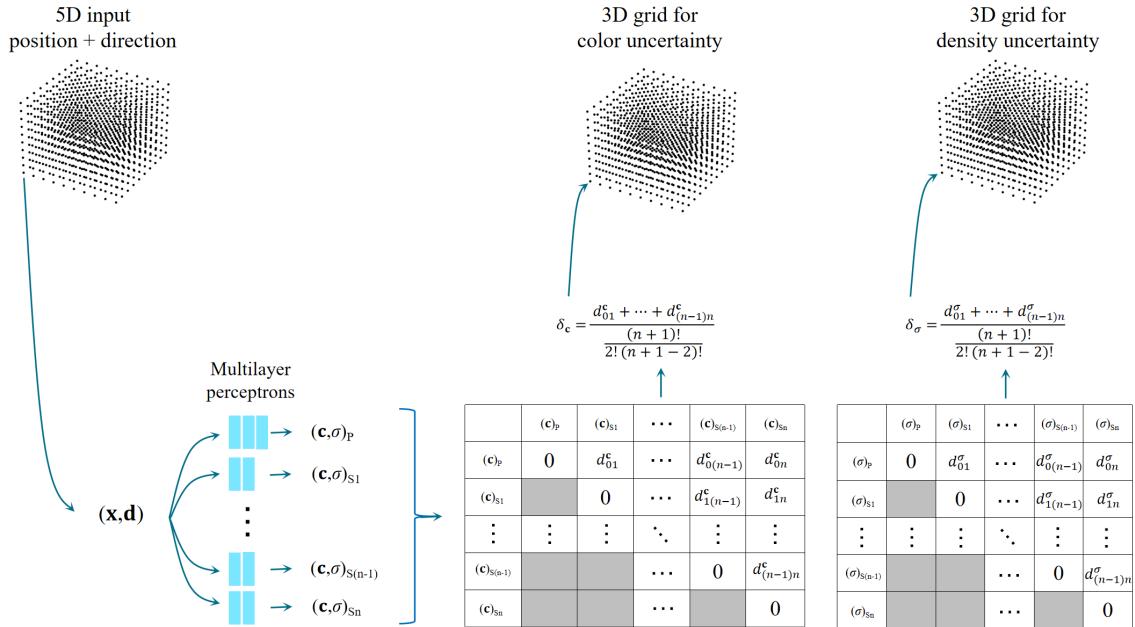


Figure 5.5: The computation of color uncertainty δ_c and density uncertainty δ_σ involves using the method of averaging of the pairwise Euclidean distances for a given point in the three-dimensional (3D) grid. Two separate 3D grids are created, one for storing the δ_c and the other for storing δ_σ , respectively. The notation d_{01}^c represents the pairwise Euclidean distance between $(\mathbf{c})_P$ and $(\mathbf{c})_{S1}$, and $d_{1(n-1)}^c$ represents the pairwise Euclidean distance between $(\mathbf{c})_{S1}$ and $(\mathbf{c})_{S(n-1)}$. Identical notation applies to other pairwise Euclidean distances in the color and volume density predictions.

Chapter 6

Visualization design

In this chapter, we will delve into a detailed description of the visualization, an integral component of the NeRFDeltaView system, as explained in Chapter 4. This uncertainty visualization comprises two subcomponents: volume rendering and overview. We also highlight the interactions between various visualization views. Finally, we briefly mention the implementation of the system.

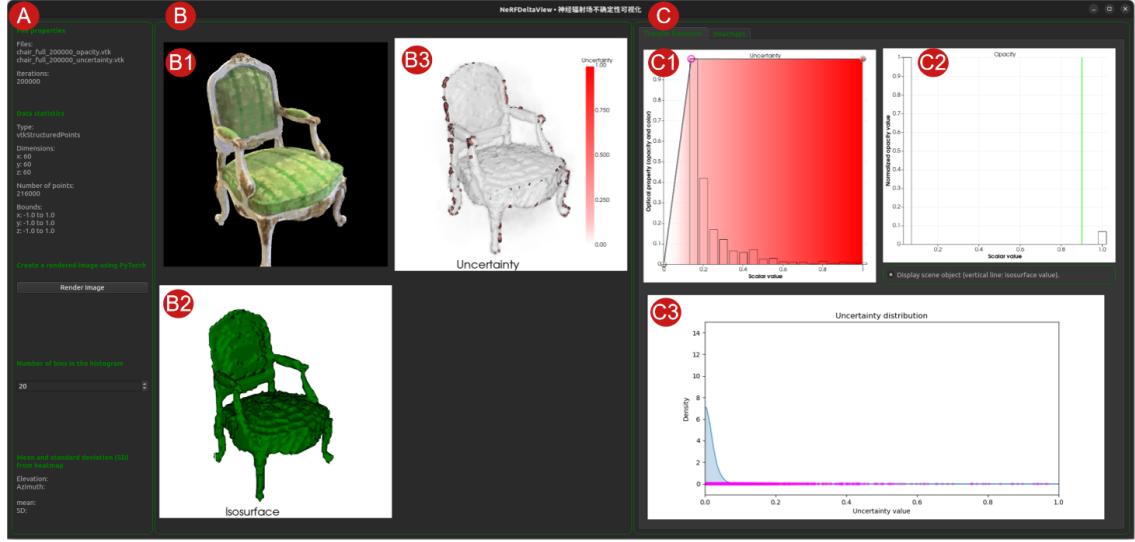
6.1 Uncertainty visualization

This section provides details of the visualization tool for uncertainty visualization in the NeRF-DeltaView system. Uncertainty visualization provides users with an in-depth understanding of the data for analysis and to perform confident and informed decision-making. The main purpose of our tool is to highlight the significance of interactive visualization in enabling users to explore the estimated uncertainty in synthesized scenes, identify model limitations, and aid in understanding NeRF model uncertainty. This is particularly relevant to address the lack of direct human interaction with proposed approaches, such as Lee et al. [42], ActiveRMAP [88], and NeU-NBV [33] and NeurAR [73]. We correlate the visualization views in the tool to the three user questions (**Q1**, **Q2**, and **Q3**) outlined in Chapter 4.

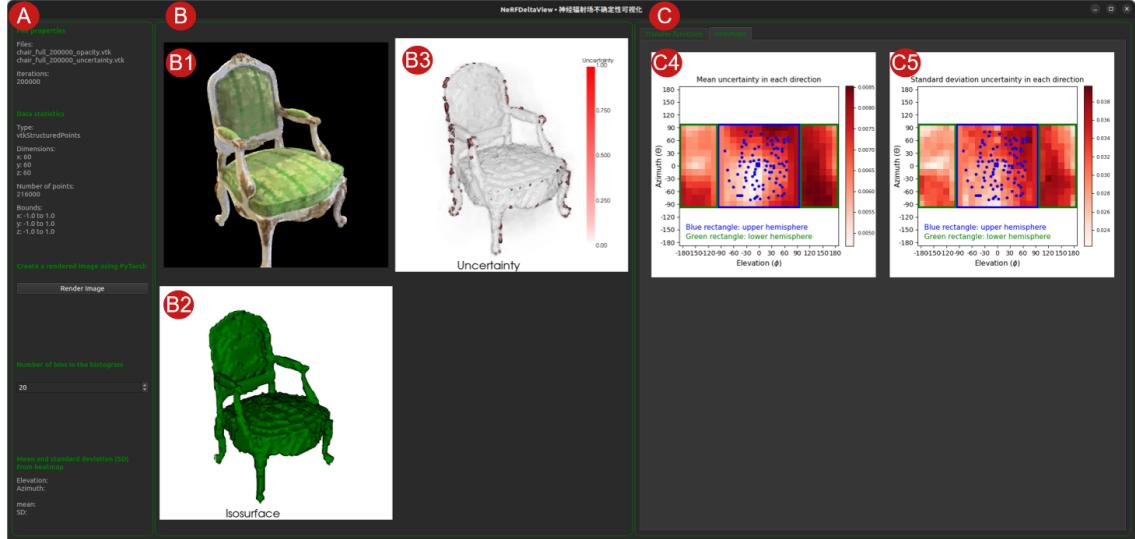
Figure 6.1 and Figure 6.2 show the screenshot of the tool designed for the Uncertainty Neural Network (Uncertainty NN) and Ensemble models, respectively. The tool composed of three panels: the data and control panel, the volume rendering panel, and the data visualization panel.

- **Data and control panel:** As shown in Figure 6.1A and Figure 6.2A, this panel displays the general information about the imported volumetric data and widgets to control the views.
- **Volume rendering panel:** This panel is shown in Figure 6.1B and Figure 6.2B, users can interact with the views in this panel to explore specific scene features in the synthesized scene, supporting **Q1** and **Q2**. The volume-rendered uncertainty views show the detailed spatial distribution of uncertainty within the synthesized scene.
- **Data visualization panel:** This panel is shown in Figure 6.1C and Figure 6.2C. It contains two tabs. The first tab contains the transfer functions (TFs) and histograms, which map uncertainty values with color and opacity. The heatmaps in the second tab provide a comprehensive, high-level perspective on the overall uncertainty patterns; they can be used to investigate the correlation between the amount and viewing direction of training images and the model uncertainty, addressing **Q3**.

The controls and views in these panels are linked. Further details on the views in these panels and their interactions will be discussed in the following sections. We will explain each view, primarily



(a) Interface displaying the TF, histogram, and 1D TF of uncertainty in panel C.



(b) Interface showing heatmaps for the mean and standard deviation of uncertainty estimation for selected viewing angles in panel C.

Figure 6.1: Screenshot of the interactive tool for uncertainty visualization in the NeRFDeltaView system for Uncertainty NN. Data and control panel (A) contains a control button and meta-information. Volume rendering panel (B) displays the synthetic, isosurface, and uncertainty three-dimensional views. The data visualization panel (C) consists of two tabs. The first tab includes the TF, histograms, and one-dimensional (1D) TF. The second tab shows heatmaps of the mean and standard deviation of uncertainty estimation for selected viewing directions.

using the visualization tool for Ensemble (Figure 6.2) as an example. The explanations in the following sections are also applicable to the visualization tool for Uncertainty NN (Figure 6.1). We have incorporated uncertainty visualization approaches, specifically overloading and side-by-side comparison, described by Pang et al. [62], as detailed in Chapter 3.

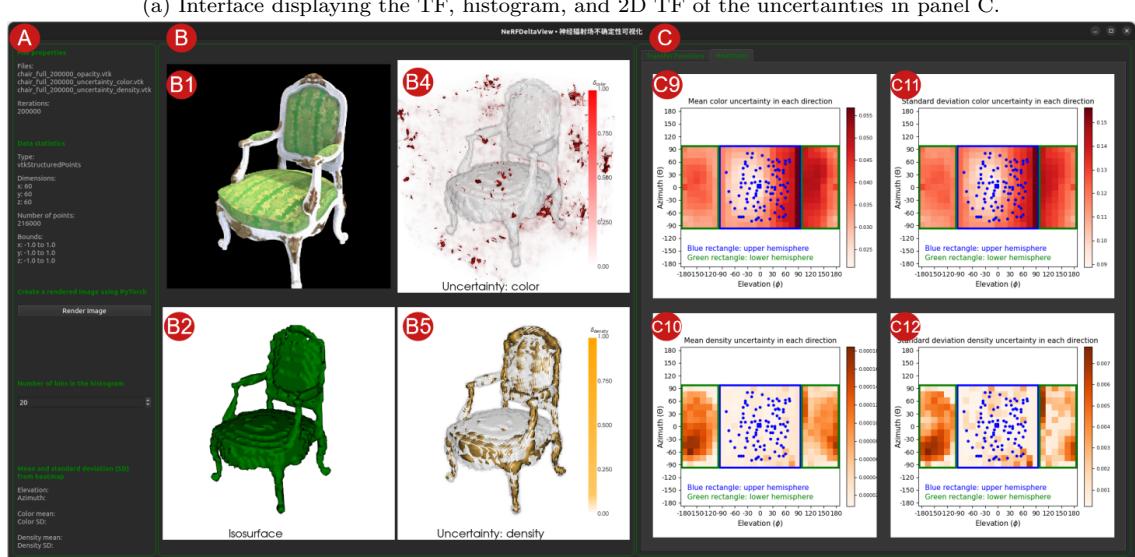
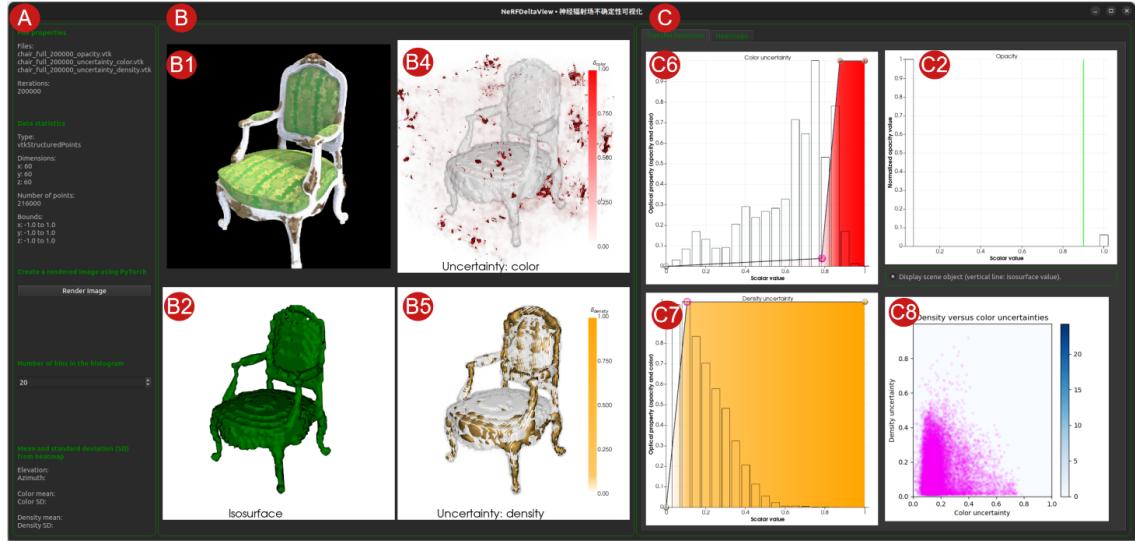
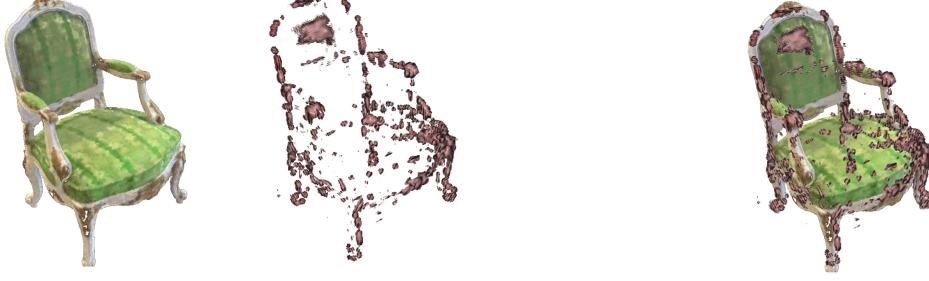


Figure 6.2: Screenshot of the interactive tool for uncertainty visualization in the NeRFDeltaView system for the Ensemble model. The data and control panel (A) contains a control button and meta-information. Volume rendering panel (B) displays the synthetic, isosurface, and uncertainty three-dimensional views. The data visualization panel (C) consists of two tabs. The first tab includes the TFs, histograms, and a two-dimensional (2D) TF of color and density uncertainties. The second tab shows heatmaps of the mean and standard deviation of the color and density uncertainty estimation for selected viewing directions.

6.1.1 Volume rendering panel

The goal of the volume rendering panel is to enable users to understand the spatial relationship of uncertainty within the synthesized scene. One of the approaches is to render the synthesized scene and uncertainty in two different views, but this approach makes it challenging to observe the relation, as an example shown in Figure 6.3a. However, overloading the uncertainty onto the synthesized scene in the same view (Figure 6.3b) encounters performance issues because rendering

synthesized scenes is considerably time-consuming.



(a) Synthesized scene and uncertainty are presented in separate views.

(b) Synthesized scene and uncertainty are presented in the same view.

Figure 6.3: Examples of the synthesized scene and uncertainty are presented (a) in separate views for a side-by-side comparison and (b) in the same view using an overloading approach.

Therefore, we opt to render the uncertainty and isosurface in the same view because the spatial relation of the uncertainty in the scene can be easily discerned, and rendering both uncertainty and isosurface volumetric data can be achieved considerably rapidly. The isosurface represents the three-dimensional (3D) surface representation of objects in the synthesized scene based on an isovalue. However, overloading of the uncertainty onto the isosurface in the same view results in a misleading perspective, where the uncertainty behind the object is visible, as can be seen in Figure 6.4a. To address this, we use the Z-buffer to provide users with accurate depth perception of uncertainties in the uncertainty views. The Z-buffer obtained from the isosurface is utilized to determine the visibility of uncertainty for a particular viewing direction, ensuring the uncertainty occluded by the object remains invisible. This is illustrated in Figure 6.4b, where the uncertainty in the rear leg is not visible due to occlusion. We refer to works by Hearn et al. [22], Hughes et al. [29], and Marschner and Shirley [50] for a comprehensive overview of Z-buffers. We incorporate uncertainty visualization with a Z-buffer, as illustrated in Figure 6.4b, into our visualization tool.

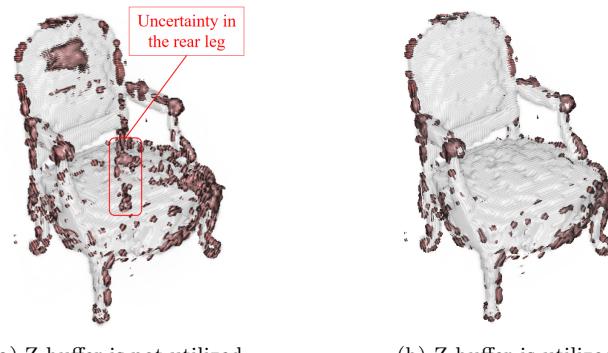


Figure 6.4: The uncertainty is overloaded on the isosurface (a) without using a Z-buffer and (b) with a Z-buffer.

We use the side-by-side comparison approach to visualize the rendered synthesized scene and the uncertainty view combined with the isosurface. The uncertainty 3D, isosurface, and synthetic views in this panel are explained in detail in the following sections.

Uncertainty 3D view

Uncertainty 3D views are important for users to explore the spatial uncertainty within the synthesized scene and understand the limitations of the model. This visualization approach serves as the volume rendering subcomponent in the NeRFDeltaView system, as discussed in Chapter 4. The color uncertainty 3D view and density uncertainty 3D view, along with the isosurface object, are depicted in Figure 6.2B4 and Figure 6.2B5, respectively. We use color as a free graphical variable to visualize uncertainty, as discussed in Section 3.3 and in the work of Raidou [72] because color is commonly used in uncertainty visualization and it effectively conveys the uncertain scalar value in our case. We employ the Z-buffer obtained from the isosurface to determine the visibility of uncertainty for a particular viewing direction, making the uncertainty behind the object invisible, as explained in the previous section.

These views are developed to provide a representation of estimated uncertainties made by the model. They can be used to identify the scene features where the model struggles and identify potential limitations in the model, guiding to understanding the model uncertainty, addressing **Q1** and **Q2**. The generation of color and density uncertainty 3D grids is described in detail in Section 5.2. These color and density uncertainty 3D grids are then overloaded onto the isosurface object, which is generated by the opacity 3D grid; the generation of the opacity 3D grid will be discussed in the following section. Subsequently, direct volume rendering (DVR) is performed for these 3D grids, and TFs are utilized to map the uncertainty value of each sample point in the 3D grids to optical properties such as color c and opacity α , as discussed in Section 2.2. The TFs are shown in Figure 6.2C6 and Figure 6.2C7. The uncertainty values range from 0 to 1, as indicated in the color bar in each uncertainty 3D view, where 0 represents the lowest uncertainty and 1 represents the highest uncertainty.

Users may want to quickly visualize and identify regions with high uncertainty while hiding the objects. Therefore, we have added a radio button (Figure 6.2C2) that allows users to toggle the display of the isosurface object in the uncertainty 3D views on or off. An example of displaying and hiding the chair isosurface in the uncertainty 3D views can be found in Figure 6.5b and Figure 6.5c, respectively. Hiding the isosurface objects in the uncertainty 3D views provides an additional perspective for users to identify other regions with high uncertainty and focus specifically on these regions of interest quickly, for instance, to identify the uncertainties behind or within the objects. It is worth noting that the Z-buffer is deactivated while the object is hidden. However, hiding the isosurface object in the uncertainty 3D view loses the context of the uncertainty spatial distribution relationship with the isosurface object. Therefore, the cutting plane widget is integrated into uncertainty 3D views. Users can interact with the cutting plane by translating and rotating it. This allows users to explore uncertainty within the object by intersecting the cutting plane with the isosurface and visualizing uncertainty while maintaining spatial context. Figure 6.5a illustrates that the cutting plane intersects with the chair, revealing high uncertainties inside the chair while the chair's isosurface remains visible.

Isosurface and synthetic views

Figure 6.2B2 displays the isosurface view of an object. The isosurface object is important for obtaining the Z-buffer to determine the visibility of uncertainty, ensuring that the uncertainty occluded by the isosurface object remains invisible. The isosurface object is generated from the opacity 3D regular grid based on the isovalue pre-defined by the user. The opacity 3D regular grid stores individual opacity values α for each sample point. The opacity α is calculated using the following equation from the optical model for direct volume rendering [53],

$$\alpha_i = 1 - e^{-\sigma_i}, \quad (6.1)$$

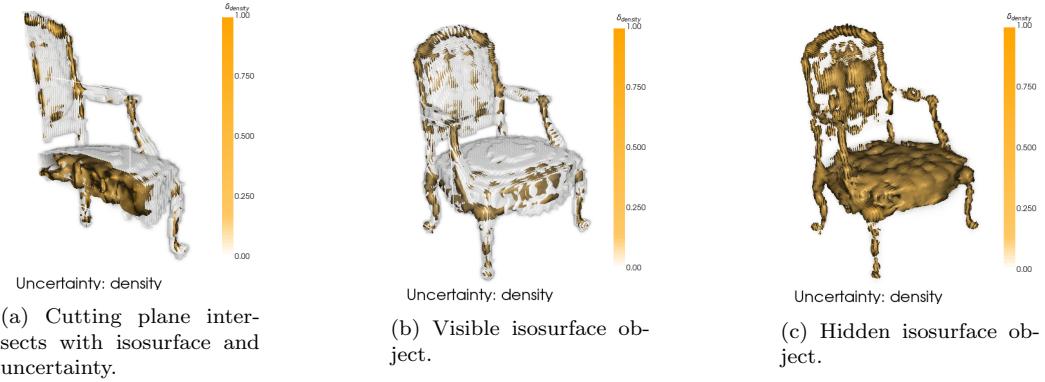


Figure 6.5: Illustration of the uncertainty 3D view with (a) cutting plane intersects with the chair isosurface and uncertainty, (b) visible chair isosurface, and (c) hidden chair isosurface. Users can translate and rotate the cutting plane as needed. The isosurface object in the uncertainty 3D view can be toggled on or off.

where i is the sample point, and the volume density σ_i is estimated by the model’s multilayer perceptron (MLP) for sample point i . Additionally, the isosurface view offers an alternative perspective on the quality of the synthesized scene, complementing the synthetic view (Figure 6.2B1).

The synthetic view (Figure 6.2B1) is rendered by the trained model based on the current viewing direction in the isosurface view and uncertainty 3D view. This visualization enables users to assess the quality of synthetic scenes from different viewing directions, providing insights into the limitations of NeRF model (**Q2**) and discovering correlations between the amount and viewing direction of training images and NeRF model uncertainty based on the rendering quality observed in the synthetic view (**Q3**).

6.1.2 Data visualization panel

The data visualization panel contains two tabs. The first tab includes the TFs, histograms, and a two-dimensional (2D) TF. The second tab consists of heatmaps displaying the mean and standard deviation of uncertainty for selected viewing directions. The TFs, histograms, and a 2D TF will be discussed in this section. The heatmaps will be discussed in the next section.

Transfer function and histogram

The TFs for color uncertainty and density uncertainty are displayed in Figure 6.2C6 and Figure 6.2C7, respectively. These TFs allow users to control the opacity of the points in their respective views to highlight the scene features that are interesting. For example, users can identify the scene features that exhibit high uncertainty by assigning high opacity through the TF.

Behind each TF, there are normalized histograms of the color and density uncertainties. These histograms depict the distribution of uncertainty values in the 3D grids, offering insights into the uncertainty estimation distribution in a 3D grid. These histograms are used as a guide for TF specification, i.e., increase opacity at one peak in the histogram to the next in order to emphasize these regions, as shown in Figure 6.2C7.

2D transfer function

The color and density uncertainties are estimated for the Ensemble model, and the relationship between these two uncertainties is represented through the 2D TF (Figure 6.2C8). In contrast, the TF for Uncertainty NN is one-dimensional (1D), as depicted in Figure 6.1C3, because Uncertainty NN outputs only one uncertainty value for a given sample point in the 3D grid. The distribution of points in the 2D TF allows users to effectively compare the estimated color and density uncertainty values in a visualization. For example, in Figure 6.6a, the color uncertainty distribution appears to be wider than that of density uncertainty. Most points are estimated to have relatively low color and density uncertainties; this can be seen in the density of the 2D TF estimated using Kernel Density Estimation (KDE), displayed behind the scatter points on a blue color scale (Figure 6.6b). The density plot provides users with an alternative perspective on the relationship between the color and density uncertainties, guiding users in the selection of scatter points. Additionally, the 2D TF provides additional information for comparing different models; for example, it reveals how scatter points in the 2D TF change after a model is further optimized with additional training images.

When specific points are selected in the 2D TF, only the sample points in the uncertainty 3D grids that correspond to the selected points become visible. To illustrate, Figure 6.6c illustrates that points beneath the chair seat in uncertainty 3D views are visible, corresponding to the selected points in the 2D TF in Figure 6.6b. This interaction allows users to effectively identify the scene features in the uncertainty 3D view corresponding to the selected points in the 2D TF.

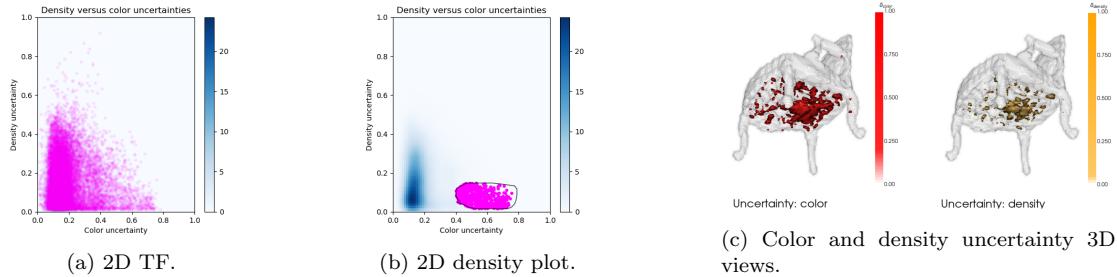


Figure 6.6: The (a) 2D TF, (b) 2D density plot of the 2D TF, and (c) the color and density uncertainty 3D views are shown after selecting points in the 2D TF on (b). The density plot is estimated using KDE and becomes visible after some points are hidden.

6.1.3 Heatmap

We have designed a heatmap that enables users to grasp the comprehensive perspective of the overview uncertainty pattern and then delve into specific viewing directions through an uncertainty 3D view. Additionally, heatmap allows users to visualize the correlation between the amount and viewing direction of training images and NeRF model uncertainty, addressing **Q3**. This visualization approach serves as the overview subcomponent in the NeRFDeltaView system, as discussed in Chapter 4.

Figure 6.2C9 to Figure 6.2C12 show the heatmaps designed to provide an overview of uncertainty statistics for selected viewing directions around the synthesized scene. These heatmaps represent the 2D mapping of a sphere surrounding a synthesized scene (Figure 6.7a and Figure 6.7b). In these heatmaps, each cell corresponds to an uncertainty statistic, such as the mean or

standard deviation, of the projected uncertainties in an uncertainty 3D view for a specific viewing direction. Examples of uncertainty 3D views for some selected viewing directions are presented as inserted images in Figure 6.7a and Figure 6.7b.

Here, we briefly explain the process of generating the heatmaps. Figure 6.7a and Figure 6.7b depict an uncertainty scene surrounded by an upper and lower hemisphere, respectively. These figures show the intersections of latitude and longitude lines on the sphere, representing camera locations, while the inserted images are the uncertainty 3D view for selected viewing directions of the camera. The angle between latitude lines and between longitudinal lines could be 1° , 10° , 15° , or other angles in elevation and azimuth directions. It is worth noting that the Z-buffer is activated for every uncertainty 3D view, ensuring that the uncertainty occluded by the object remains invisible, as explained in Section 6.1.1. We then calculate the mean and standard deviation of the projected uncertainties in an uncertainty 3D view for these selected viewing directions. These uncertainty statistics are subsequently displayed in each cell in the heatmaps, with each cell corresponding to specific viewing directions in the hemisphere. Figure 6.7d shows an example of a heatmap of mean density uncertainty. The blue points in the heatmap represent the viewing directions of the training images.

Figure 6.7c displays clusters of cells within a heatmap with each cluster representing viewing directions from the upper or lower hemispheres. The red cells within the blue-bordered rectangle, with azimuths and elevations between -90° and 90° , indicate viewing directions from the upper hemisphere (Figure 6.7a). The orange cells surrounded by the green-bordered rectangles form a cluster representing viewing directions from the lower hemisphere (Figure 6.7b). These blue-bordered and green-bordered rectangles are used in actual heatmaps, as shown in Figure 6.7d.

Despite its advantages, such as providing a perspective overview and highlighting patterns, utilizing a heatmap as a 2D mapping of a sphere has a downside. This downside is that it may flatten the representation of 3D data, potentially leading to distortions. For instance, the poles of the sphere may appear disproportionately large. To illustrate, Figure 6.8 shows the uncertainty statistic values remaining constant at elevations of -90° and 90° throughout azimuths between -90° and 90° , as the camera is located at the poles of the hemisphere. These uncertainty values at elevations of -90° and 90° in the heatmap at the poles may give a false impression that they occupy a relatively large region in the sphere. Additionally, 2D mapping of a sphere may result in a loss of spatial context, posing a challenge for users to visualize the distribution of uncertainties across various spatial directions in the perspective overview. We choose a heatmap as a 2D mapping of a sphere for the perspective overview of the uncertainty pattern because it is easy to implement and adequately serves our overview subcomponent.

6.1.4 Interactions

The views are linked and interact with each other, enabling users to explore the uncertainty of a scene estimated by the models. Some links between views have been mentioned and discussed in Section 6.1.2. In this section, we will explain the remaining links between views, primarily focusing on the visualization tool in the NeRFDeltaView system for Ensemble (Figure 6.2). The explanations in this section are also applicable to the visualization tool in the NeRFDeltaView system for Uncertainty NN (Figure 6.1).

Link between views in volume rendering panel

Users can interact with the isosurface view and uncertainty 3D views by performing zoom, rotate, and pan operations. The goal of these operations is to empower users to focus on and explore the specific scene features of interest in the views (**Q1**). All these views share the same

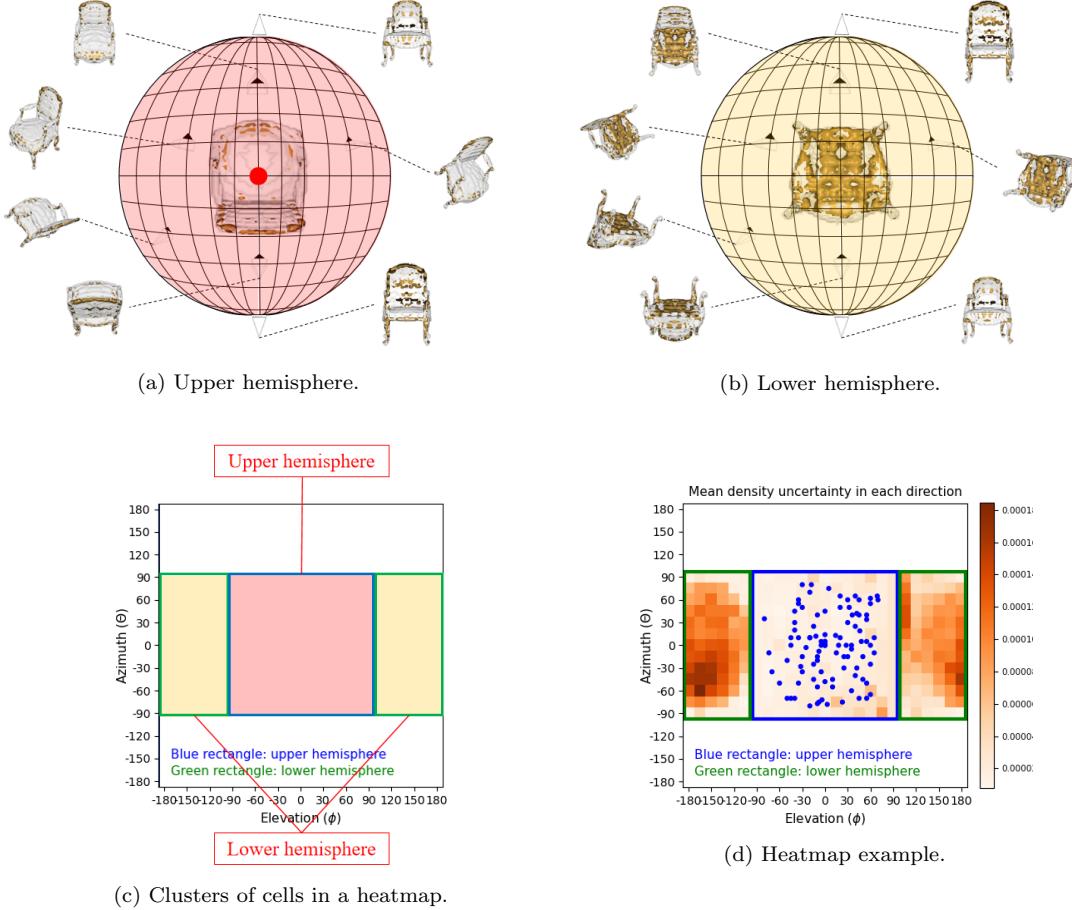


Figure 6.7: The chair uncertainty scene is surrounded by the sphere on both the (a) upper hemisphere and (b) lower hemisphere. The pyramids represent the cameras. The inserted images display the uncertainty 3D views for the respective camera. The red point in (a) represents 0° for both azimuth and elevation angles, i.e., it is located on the top view of the chair scene. In (c), the clusters of cells are presented in a heatmap. The colors of each cluster of cells represent the viewing directions from the upper or lower hemispheres. The blue-bordered rectangle encompasses cells with viewing directions from the upper hemisphere, while the green-bordered rectangles encompass cells with viewing directions from the lower hemisphere. Lastly, (d) shows an example of a heatmap. The blue points in the heatmap represent the viewing directions of the training images.

camera, meaning that the viewing directions are updated simultaneously. The shared camera provides a consistent perspective across different views, fostering a seamless exploration process and side-by-side comparison.

The synthetic view is rendered by the trained model based on the current viewing direction in the isosurface view and uncertainty 3D views after pressing the button on the data and control panel (Figure 6.2A). This allows users to assess the rendering quality of synthesized scenes from different viewing directions, providing insights into the limitations of the model (**Q2**) and the relationship between the model uncertainty and the quantity and viewing direction of training images (**Q3**).

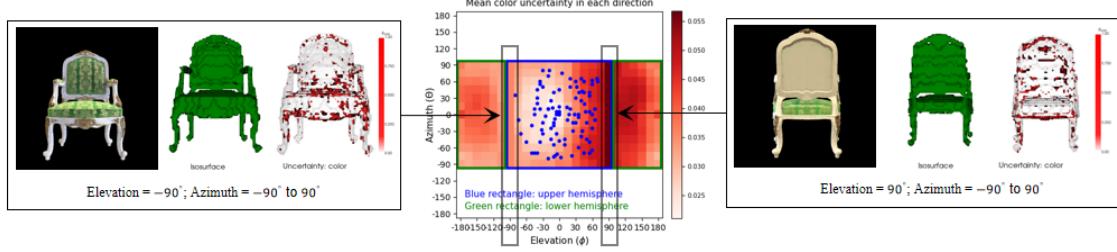


Figure 6.8: The synthetic view, isosurface view, and the uncertainty 3D view of the chair scene at an elevation of -90° and 90° . The mean values are constant at these elevations because the camera is located at the poles of the hemisphere.

Link between views in volume rendering panel and data visualization panel

Based on the "Overview first, then details on demand" principle [79], users can swiftly identify patterns of interest from an uncertainty overview, and then delve into further exploration in the volume rendering views. For example, identifying the viewing direction gives a high mean uncertainty in an uncertainty overview, and then users can focus on the scene features of interest in this specific viewing direction (**Q1**). To facilitate this process, we establish a link between the heatmaps and uncertainty 3D views. Users can hover over the heatmaps, which will simultaneously update the viewing direction for these views depending on the cursor's position (elevation and azimuth angles) on the heatmaps. This interactive feature streamlines identifying the uncertainty 3D view in a specific direction. To illustrate, Figure 6.9 displays the uncertainty 3D view corresponding to the cell where the cursor has hovered in the heatmap.

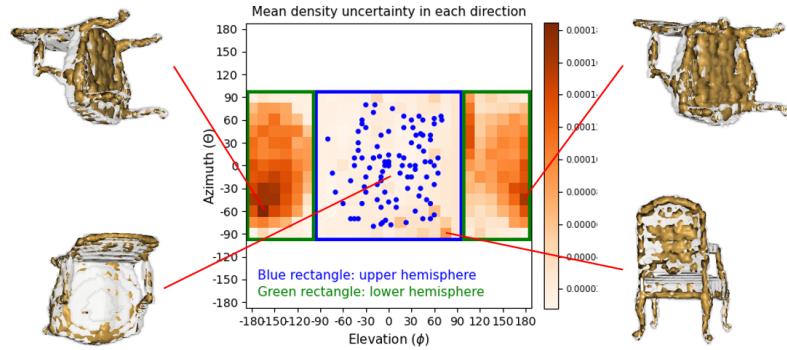


Figure 6.9: The uncertainty 3D views corresponding to the selected cells in the heatmap of the mean density uncertainty when the cursor hovers over these cells.

6.2 Implementation details

We trained the Uncertainty NN and the Ensemble models using PyTorch [64] with CUDA [59] on an NVIDIA RTX 4090 GPU [8]. The training duration for a model for a single scene with 100000 iterations is approximately 12 hours. The generation of opacity and uncertainty 3D grids, as well as the computation of mean and standard deviation uncertainties, are conducted on an AMD Ryzen 9 5900HX CPU. We implemented the visualization tool using Python [82], VTK from Python wrappers [76] and PyQt6 [70] on an AMD Ryzen 9 5900HX CPU. Different rendering packages and techniques were employed for various views. For example, the isosurface view, uncertainty

3D views, transfer functions, and histogram are rendered using packages from VTK from Python wrappers, while the 2D transfer function and heatmaps are generated using Matplotlib [31].

For the Uncertainty NN model, the following weights for the uncertainty loss \mathcal{L}_δ (Equation 5.4) and total loss \mathcal{L} (Equation 5.5) are used: $\tau_1 = 0.1$ and $\tau_2 = 0.01$, $\beta_1 = 1.0$ and $\beta_2 = 0.1$. These weights are set based on the analysis of the model performance and sensitivity to uncertainty measures. For both the Uncertainty NN and Ensemble models, we use the ADAM optimizer [38] with a learning rate set to 5.0×10^{-3} in the beginning and decays exponentially to 5.0×10^{-5} over the course of optimization. We employ a batch size of 2048 rays, and configure the sampling point per ray as 64 and 128 for the coarse and fine MLPs, respectively. These optimizer, learning rate, batch size, and sampling point per ray settings are as original NeRF [54].

In the context of 3D grid generation, we explored several numbers of sample points along each of the X , Y , and Z axes, including 30, 50, 60, 80, and 100, to strike a balance between generation duration and rendering quality in isosurface view and uncertainty 3D views. We opted to generate the opacity and uncertainty 3D regular grids with 60 sample points along each axis using the trained model because it provided a balanced compromise between generation duration and rendering quality in those views. The duration to generate the 3D grid and launch the tool increases significantly as the number of sample points increases. For example, the 3D grid generation duration increases from 3 minutes to 14 minutes when the number of sample points increases from $60 \times 60 \times 60$ to $100 \times 100 \times 100$.

Chapter 7

Experimental results

In this chapter, we begin by providing details of the datasets and the training strategy for comparative analysis. Subsequently, we present the results of our NeRFDeltaView system, focusing on three use cases.

7.1 Datasets

We evaluate our visualization tool based on the three synthesis scene datasets. These datasets are also utilized in the original Neural Radiance Field (NeRF) [54]. These three synthesis scenes are the chair, microphone, and drums, as shown in Figure 7.1. These scenes are selected because they cover a wide range of geometrical detail and lighting conditions. The chair has a relatively simple three-dimensional (3D) structure, and the drums scene is a good mix of complex lighting conditions and complicated geometric details, for example, the cymbal stands and hoops in the tom-toms and drums (Figure 7.1c). Whereas the microphone scene falls between the chair and drums scenes in terms of complex lighting conditions and geometric details, particularly the woven wire grill in the microphone capsule (Figure 7.1b). Each scene contains 100 training images and 200 testing images, all with dimensions of 800×800 pixels, along with corresponding transformation matrices, which are used to determine the camera position \mathbf{o} and viewing direction \mathbf{d} , as described in Section 2.1.2. These scenes were rendered from the viewpoints sampled on the upper hemisphere where the object is placed in the center, as depicted in Figure 7.1d.

We create two sets of datasets for each scene: full-angle and limited-angle datasets. This approach allows us to investigate the correlation between the quantity and viewing direction of the training images and the model uncertainty, addressing **Q3** as described in Chapter 4. In Section 7.2, we will discuss the training strategy employed for the models using both full-angle and limited-angle datasets. The full-angle dataset consists of 100 training images obtained from the original NeRF dataset [54]. In contrast, the limited-angle dataset is created from a subset of the full-angle dataset, where only a limited view of an object in the scene is provided. The choice for the limited-angle datasets of the chair, microphone, and drums are described below.

- Chair: The training images are selected in a way that the front part of the backrest is hardly visible (Figure 7.2a). There are 39 training images in the limited-angle dataset.
- Microphone: The training images are chosen so that the right metal strip of the microphone capsule is hardly visible (Figure 7.2b). There are 36 training images in the limited-angle dataset.
- Drums: The training images are selected so that the pedal of the bass drum is hardly visible (Figure 7.2c). There are 22 training images in the limited-angle dataset.



(a) Chair.



(b) Microphone.



(c) Drums.



(d) Drums surrounded by the pyramids on the upper hemisphere [54].

Figure 7.1: Example of the (a) chair, (b) microphone, and (c) drums training images. In (d), the drums are surrounded by pyramids in the upper hemisphere [54]. The pyramids represent the camera.

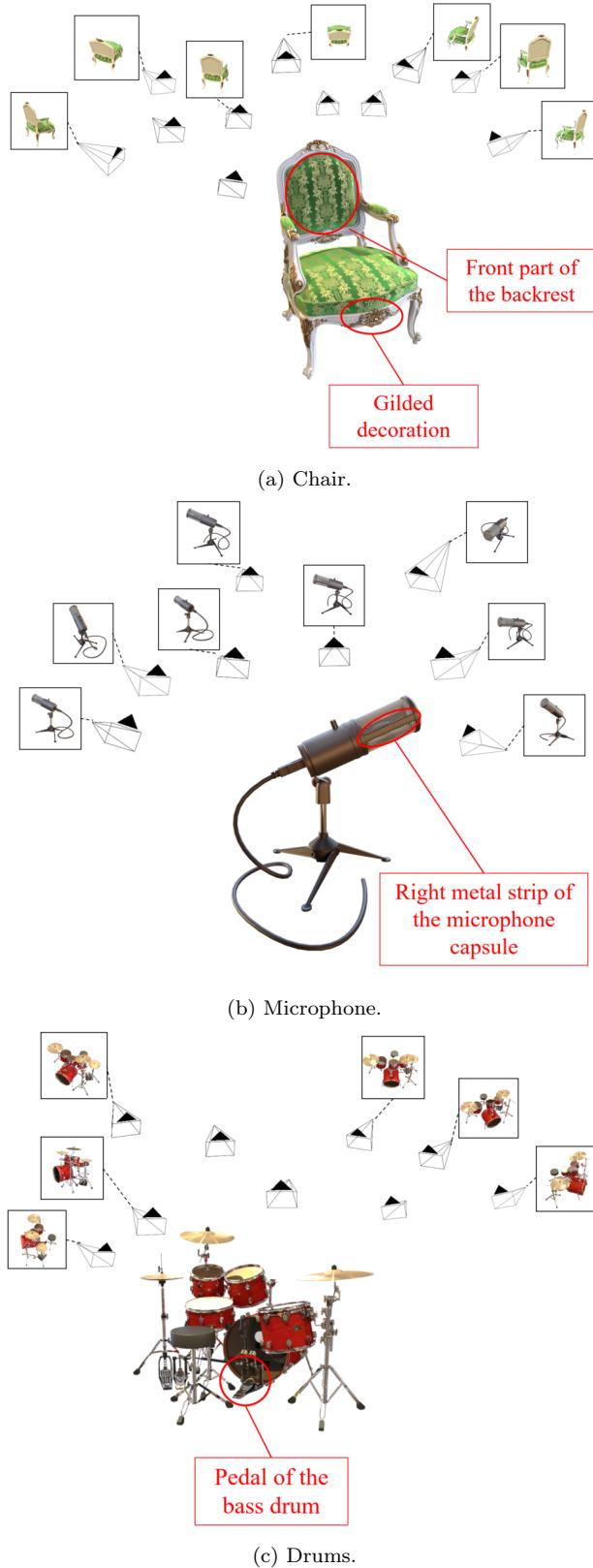


Figure 7.2: An example of the training images from the limited-angle dataset for the chair, microphone, and drums. The pyramids represent the camera, and the insert images display the training images captured by the respective cameras.

7.2 Training strategy for comparative analysis

In this section, we describe the training strategy employed for the model, leveraging both limited-angle and full-angle datasets. The fundamental goal is to conduct a meaningful comparative analysis between the models initially trained with limited-angle and subsequently refined with full-angle datasets. This comparison allows us to evaluate how the model adapts and performs in scene synthesis and uncertainty estimation when exposed to additional training images with a broader range of viewing directions around the scene, providing valuable insights into understanding the model uncertainty, addressing **Q3** as described in Chapter 4.

Figure 7.3 displays the training strategy schematically. An approach is to train the model with full-angle and limited-angle datasets separately from the same initial parameters Θ . Figure 7.3 shows the models with optimized parameters $\Theta_{1_{\text{lim}}}^*$ and optimized parameters $\Theta_{2_{\text{full}}}^*$ are trained with limited-angle dataset and full-angle dataset, respectively. Both models are trained from the same initial parameters Θ . However, there is no guarantee that the models trained from the same initial parameters will converge to the same parameter subspace. This is because the optimization process could be influenced by the nature of the data, i.e., training dataset size and variety could lead to variations in the optimization process [2]. Figure 7.3 demonstrates that comparing the models with optimized parameters $\Theta_{1_{\text{lim}}}^*$ and $\Theta_{2_{\text{full}}}^*$ may not be meaningful, as both might converge to distinct parameter subspace even though they are initiated with identical initial parameters Θ . Therefore, for a meaningful comparison of the model results obtained from both the limited-angle and full-angle datasets, we applied an approach where we compared the model initially trained with the limited-angle dataset and the model subsequently refined with full-angle datasets. This approach will be explained in detail in the following paragraph.

We initiate the training of the model with initial parameters Θ using the limited-angle dataset, obtaining optimized parameters $\Theta_{1_{\text{lim}}}^*$. Subsequently, we continue training the model, this time with the full-angle dataset, and obtain optimized parameters $\Theta_{1_{\text{full}}}^*$. Each model training comprises 100000 iterations, as illustrated in Figure 7.3. This training strategy ensures that the model continues to optimize within the same parameter subspace, allowing for a meaningful comparison of the model results obtained from both the limited-angle and full-angle datasets. In other words, the optimized parameters $\Theta_{1_{\text{lim}}}^*$ and $\Theta_{1_{\text{full}}}^*$ are within the same parameter subspace. This simulates a situation in practice; for example, a model is further optimized with a new set of training images in active robot vision applications [42].

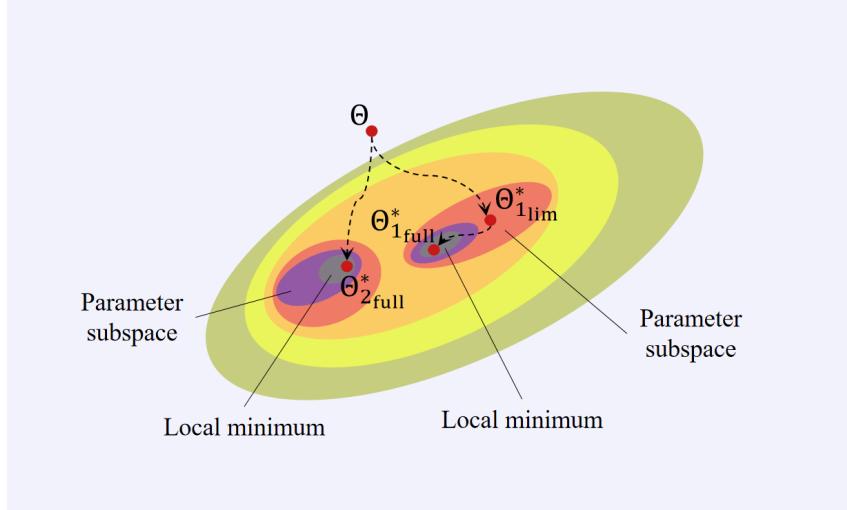


Figure 7.3: An example of a solution space or parameter space with two local minima. The model has initial parameters Θ . $\Theta_{2\text{full}}^*$ represents the optimized parameters of the model trained with the full-angle dataset. $\Theta_{1\text{lim}}^*$ and $\Theta_{1\text{full}}^*$ represent the optimized parameters of the model first trained with limited-angle dataset and then with the full-angle dataset, respectively. The dotted arrows represent possible paths during optimization, each with 100000 iterations.

7.3 Use cases

In this section, we demonstrate the practical usability results of our interactive uncertainty visualization in the NeRFDeltaView system. We focus on three scenes: a chair, a microphone, and drums. For each scene, we perform a qualitative comparison of the results between limited-angle and full-angle datasets for both the Uncertainty Neural Network (Uncertainty NN) model and the Ensemble model. Additionally, we compare the results qualitatively between both models. We correlate the results with the three user questions (**Q1**, **Q2**, and **Q3**) outlined in Chapter 4. In the end, we summarize and discuss our findings.

7.3.1 Dataset: chair

Chair: Qualitative results for Uncertainty NN

This section explores the uncertainty visualization for the chair scene trained by the Uncertainty NN model. Figure 7.4 shows the chair scene's synthetic and uncertainty 3D views trained with the Uncertainty NN, along with transfer functions (TFs) settings, for both the limited-angle and full-angle datasets. In the model trained with the limited-angle dataset, we expect to see relatively high uncertainty in the parts of the chair not visible in the training images. Specifically, these include the front part of the backrest, gilded decoration, and the bottom part of the seat. In Figure 7.4a, these regions are rendered incorrectly in the synthetic view. For example, the green texture structure on the front part of the backrest is not visible. To check whether these regions indeed contain higher uncertainties, we select the points with uncertainty exceeding 0.2 in the one-dimensional (1D) TF, as shown in Figure 7.4a. We define the uncertainty values higher than 0.2 as relatively high uncertainty for conservative purposes. The corresponding points in the uncertainty 3D grid linked to the selected points in the 1D TF become visible. However, we have discovered that points with relatively high uncertainty do not appear in the front part of the backrest, gilded decoration, and the bottom part of the seat; instead, the high uncertainty appears on the edges of the chair in the uncertainty 3D view, as shown in Figure 7.4a. This observation

contradicts our expectations. This result highlights the need for further investigation of why high uncertainty is exhibited on the edges of the chair instead of the front part of the backrest, gilded decoration, and the bottom part of the seat and how these scene features relate to the model uncertainty (**Q1**). It could be explained by the limitation of the uncertainty estimation model being trained to overfit the training images, and the uncertainty value is learned to be as close to zero as possible in the regularization loss (**Q2**).

For the Uncertainty NN model trained with the full-angle dataset, we observed that the front part of the backrest and the gilded decoration are rendered correctly in the synthetic view (Figure 7.4b). However, we are unable to verify the correctness of the rendering for the bottom part of the seat, as this region is not visible in any of the training images. We do not observe relatively high uncertainty exhibited at the bottom part of the seat in the uncertainty 3D view (Figure 7.4b), which further emphasizes the need for further investigation of why relatively high uncertainty is not exhibited in this region (**Q1**). It could be explained by the possible reasons mentioned earlier, i.e., overfitting to the training images, and the uncertainty value is learned to be as close to zero as possible in the regularization loss.

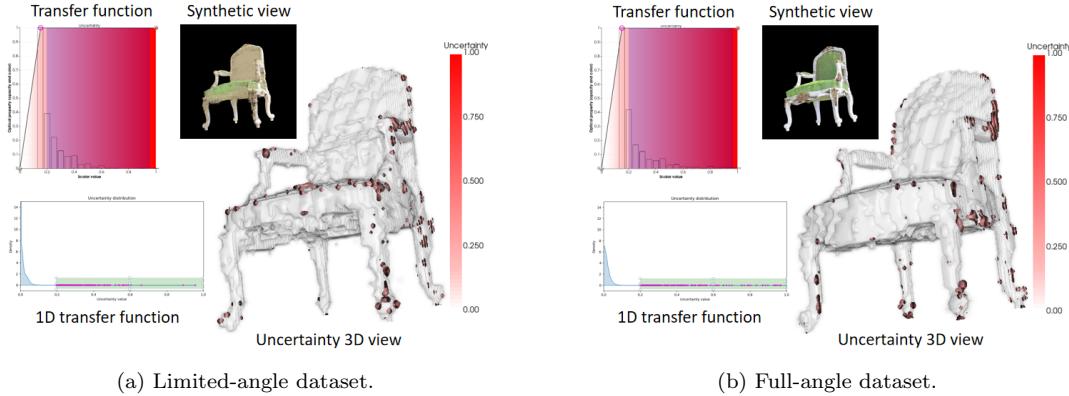


Figure 7.4: The synthetic view, uncertainty 3D view, transfer function, and the 1D transfer function for the chair scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.

Figure 7.5 displays the heatmaps of mean and standard deviation uncertainties in selected viewing directions for the limited-angle and full-angle datasets. The range in the color bar differs for each heatmap, depending on the maximum and minimum values in their respective statistical data. The inserted images in the figure display the synthetic view and uncertainty 3D view for the corresponding cells. In an uncertainty heatmap, we expect the uncertainty value in the cells to be lower around the viewing direction of a training image because the model has learned to estimate color and density optimally in that specific direction.

In Figure 7.5a, we observe that the mean uncertainties are relatively lower for cells in the vicinity of the training images, aligning with expectations. This suggests the factor of viewing direction of various training images influences the uncertainty distribution in the chair scene for the limited-angle dataset (**Q3**). We also observe that the heatmap cells exhibiting relatively high mean uncertainty correspond to views from the upper hemisphere, which is consistent with the invisible front part of the backrest in the limited-angle dataset, as shown in Figure 7.5aB.

In Figure 7.5b, the number of cells with relatively high mean uncertainty within the blue-bordered rectangle does not appear to significantly decrease as the number of training images increases (**Q3**), compared with Figure 7.5a. Additionally, the heatmap cells showing relatively low uncertainty within the green-bordered rectangle are due to that there is no relatively high

uncertainty present at the bottom part of the seat (Figure 7.5bB). These deviations from our expectations are probably because the Uncertainty NN model learns limited information from the additional training images in the full-angle dataset. This result suggests the need for further investigation into the limitations of Uncertainty NN (**Q2**). We notice a comparable pattern in both mean and standard deviation in Figure 7.5, i.e., the cells with relatively high mean values also exhibit relatively high standard deviation values. This observation may lead users to question the reliability of the estimation of high mean values.

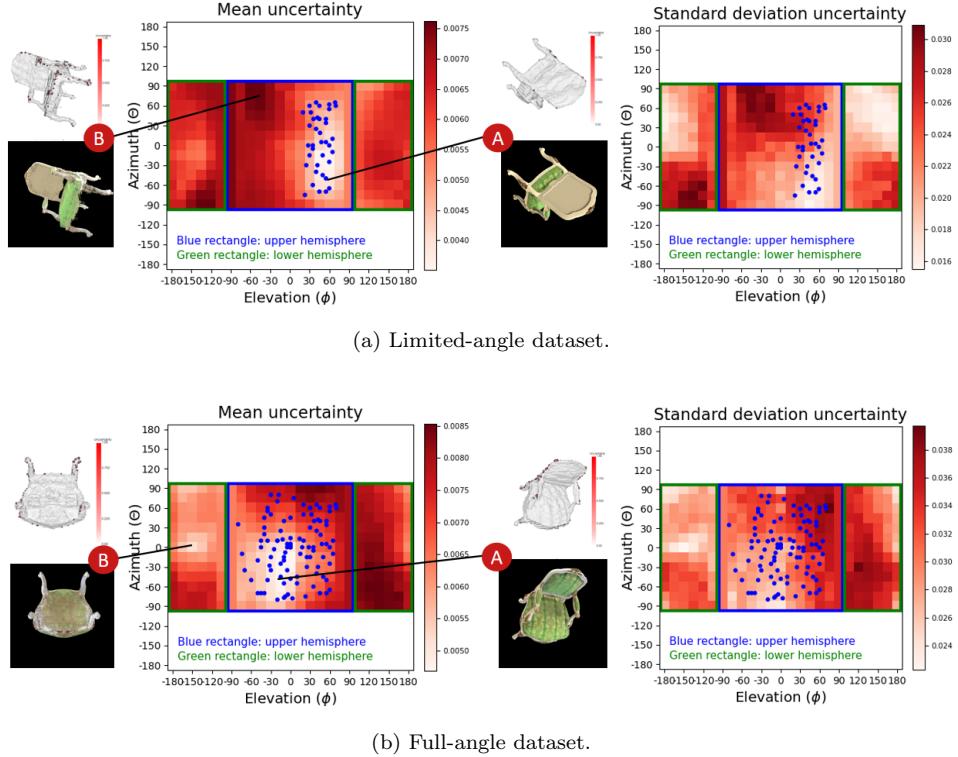


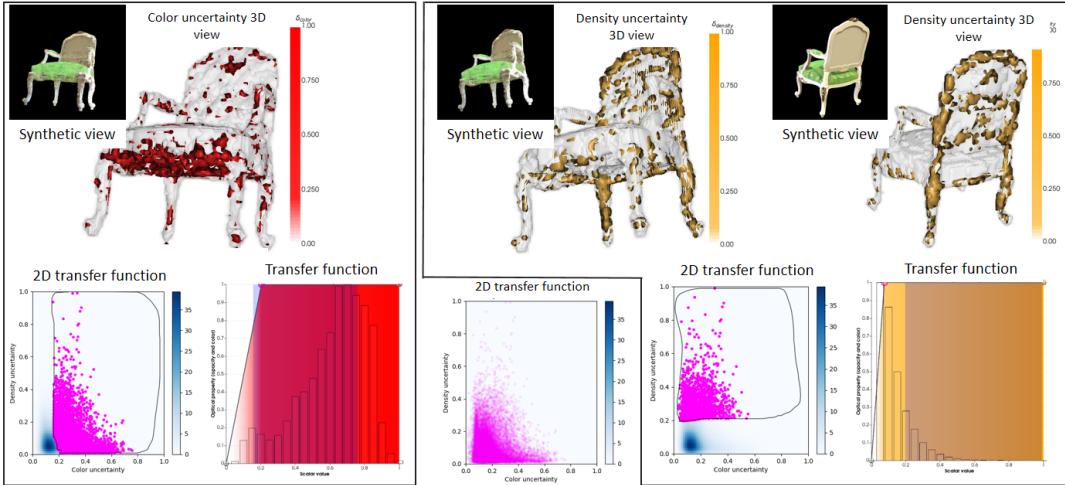
Figure 7.5: The mean and standard deviation uncertainty heatmaps in selected viewing directions for the chair scene trained with the Uncertainty NN model using (a) limited-angle and (b) full-angle datasets. The inserted images are the synthetic view and uncertainty 3D view for the corresponding cells.

Chair: Qualitative results for Ensemble

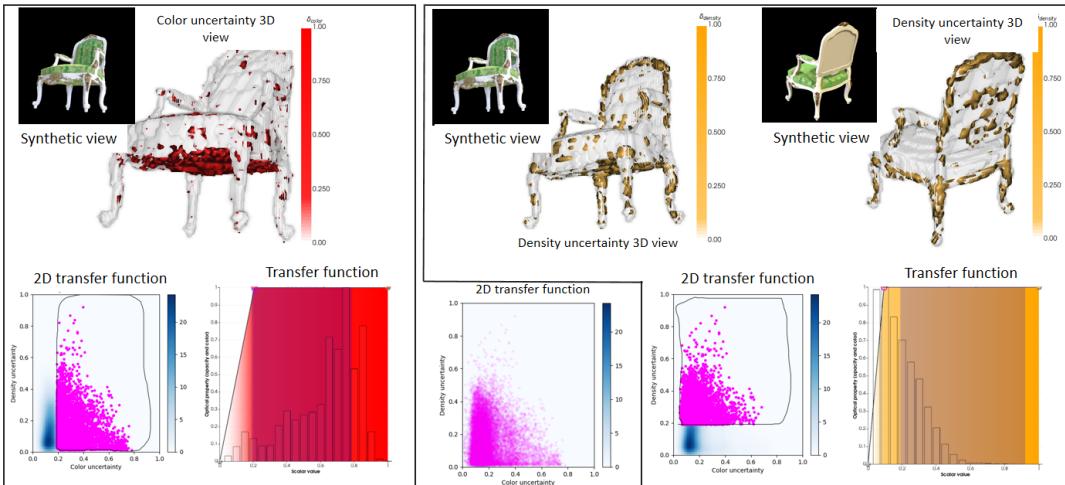
In this section, we explore uncertainty visualization for the chair scene trained by the Ensemble model. Figure 7.6 presents the chair scene's synthetic view, color and density uncertainty 3D views trained with the Ensemble model, along with their respective TFs and two-dimensional (2D) TF settings, for both the limited-angle and full-angle datasets. In the model trained with the limited-angle dataset, the front part of the backrest is rendered incorrectly in the synthetic view, with the green texture structure not being visible (Figure 7.6a), as this region is not visible in the limited-angle training images. In Figure 7.6a, we observe color and density uncertainty values exceeding 0.2 at the bottom of the seat and the front part of the backrest in the color and density uncertainty 3D views (**Q1**) after selecting points with uncertainty higher than 0.2 in the 2D TF, aligning with our expectations. However, we observe relatively high certainty at the rear part of the backrest, even though this region is visible in the limited-angle training images. This observation contradicts our expectations and emphasizes the necessity for further investigation of

why high uncertainty is exhibited in the back part of the backrest in future work (**Q1**).

Figure 7.6b shows the results for the Ensemble model trained with the full-angle dataset. We observe a reduction in the uncertainty area on the front part of the backrest in the uncertainty 3D views compared to those from the limited-angle dataset. This reduction is due to the visibility of this region in the full-angle training images. The bottom of the seat still exhibits relatively high uncertainty values, as expected, because this region is not visible in limited-angle and full-angle training images. Surprisingly, the rear part of the backrest continues to show relatively high uncertainty values, despite the inclusion of more images, including views of the rear part of the backrest, in the full-angle dataset. This suggests future work investigating the limitations of the model might be necessary (**Q2**).



(a) Limited-angle dataset.



(b) Full-angle dataset.

Figure 7.6: The synthetic view, color and density uncertainty 3D views, transfer function, and the 2D transfer function for the chair scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.

Figure 7.7 shows the heatmaps of mean and standard deviation for color and density uncer-

tainties for limited-angle and full-angle datasets. The figure also includes inserted synthetic and uncertainty 3D views corresponding to the respective cells. We observe a significant reduction in both mean and standard deviation uncertainty values to the full-angle dataset when comparing the heatmaps of the limited-angle dataset. This indicates that increasing the training images for this scene significantly reduces the uncertainty (**Q3**). In the mean and standard deviation density uncertainty heatmaps for the limited-angle dataset in Figure 7.7a, the uncertainty values are relatively high in the cells between elevations 75° and 105° due to the relatively high uncertainty exhibited in the rear part of the backrest, as shown in Figure 7.7aD. In Figure 7.7b, the relatively high mean standard deviation color and density uncertainty within the green-bordered rectangles correspond to the uncertainty viewing from the lower hemisphere, which includes the high uncertainty at the bottom of the seat, as shown in Figure 7.7bB and Figure 7.7bC.

We observe similar patterns in the mean and standard deviation uncertainty heatmaps for both limited-angle and full-angle datasets. This implies that the cells with relatively high mean values also tend to exhibit relatively high standard deviation values, raising potential concerns about the reliability of high mean values estimation. Similar patterns are also observed in the mean and standard deviation uncertainty heatmaps for the microphone and drums scenes trained with the Ensemble model.

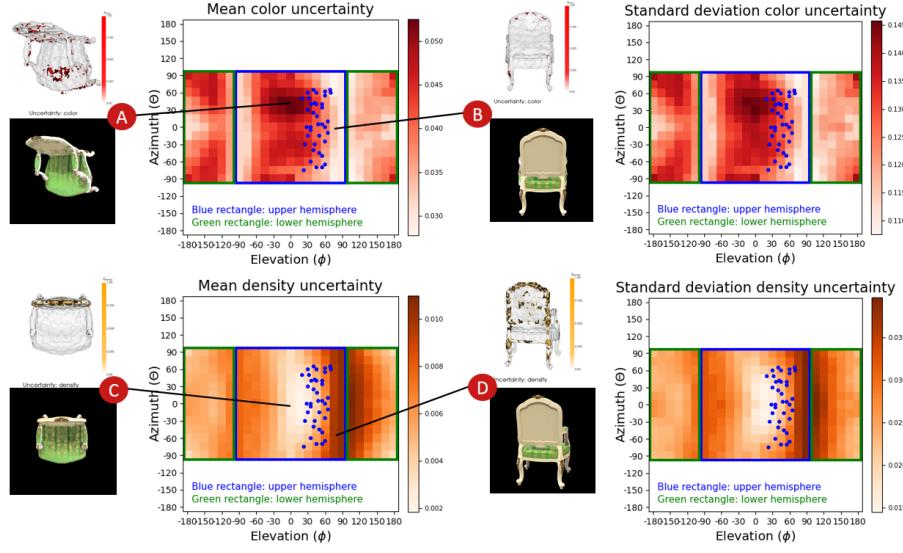
Chair: Qualitative results comparison between Uncertainty NN and Ensemble

Figure 7.8 presents a comparison between the cross-sectional uncertainty 3D views and normalized mean uncertainty heatmaps of the Uncertainty NN and Ensemble models for the chair scene. In the cross-sectional uncertainty 3D view, we expect to observe high uncertainty within the seat because none of the training images show inside the seat. The interactive cutting plane allows for intersection with the volumetric data to identify uncertainty values within the seat and validate this expectation. The ensemble model correctly estimates high uncertainty within the seat, as shown in the uncertainty 3D view in Figure 7.8b, contrasting with the uncertainty 3D view of Uncertainty NN in Figure 7.8a. This might be attributed to the earlier mentioned limitations of the Uncertainty NN model being trained to overfit the training images, and the uncertainty value is learned to be as close to zero as possible in the regularization loss (**Q2**).

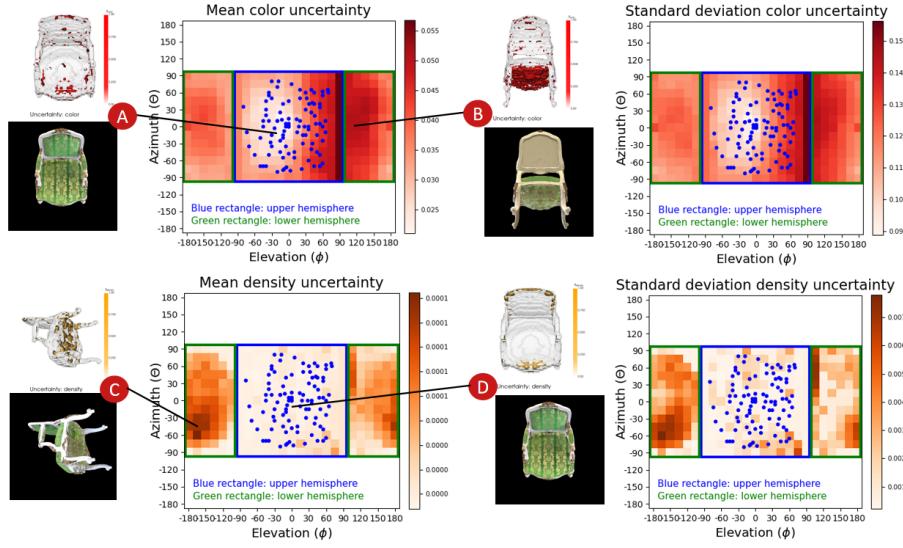
When comparing the heatmaps of both models, the Ensemble's mean density uncertainty heatmap has a more distinct pattern, making it easier for users to understand and interpret the information presented. In both the Uncertainty NN's mean uncertainty heatmap and the Ensemble's mean color uncertainty heatmap, there is higher variability in the vicinity of the training images, i.e., their mean uncertainties are sensitive to slight variations in the viewing direction. This sensitivity is expected, given the estimation of color is sensitive to direction.

In the subsequent analysis of the microphone and drums scenes, we employ a similar procedure to qualitatively analyze the uncertainty visualization for both the Uncertainty NN model and the Ensemble model. To keep the explanation concise for both scenes, the following information and expectations apply to the following analysis and will not be mentioned:

- Uncertainty values higher than 0.2 are defined as relatively high uncertainty.
- The process of selecting of points with uncertainty exceeding 0.2 in the 1D and 2D transfer functions.
- The color bar range in the heatmap is set to the maximum and minimum values in its respective statistical data.



(a) Limited-angle dataset.



(b) Full-angle dataset.

Figure 7.7: The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the chair scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets. The inserted images are the synthetic view and uncertainty 3D view for the corresponding cells.

- The synthetic view and uncertainty 3D view in the heatmap figure correspond to the respective cells.
- We expect the uncertainty value in the heatmap's cells near the training images to be lower.

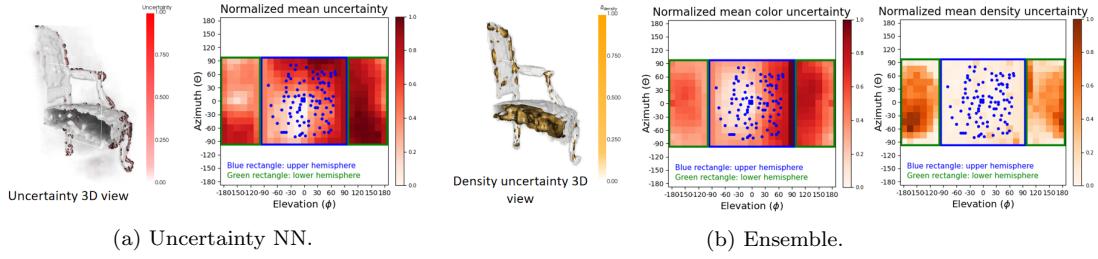


Figure 7.8: The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the chair scene. In the uncertainty 3D views, a cutting plane intersects with the volumetric data. The mean uncertainty values in the heatmaps have been normalized.

7.3.2 Dataset: microphone

Microphone: Qualitative results for Uncertainty NN

The microphone scene poses a higher level of complexity compared to the chair scene, especially the woven wire grill in the microphone capsule and the variations in lighting conditions on the metal surface. In this section, we explore the uncertainty visualization for this scene as trained by the Uncertainty NN. Figure 7.9 shows the microphone scene’s synthetic and uncertainty 3D view trained with the Uncertainty NN, along with the TFs settings, as well as the heatmaps showing mean and standard deviation uncertainties, for both the limited-angle and full-angle datasets.

For the Uncertainty NN trained with the limited-angle dataset, we expect relatively high uncertainty on the right side and the woven wire grill of the microphone capsule because of the limited visibility of the right side of the capsule in the limited-angle training images and the woven wire grill being a high-frequency variation in geometry. This can be seen in the synthetic view in Figure 7.9a, which shows that the scene is incorrectly rendered on the right side and the woven wire grill of the microphone capsule, specifically, the right metal strip on the capsule, is not rendered. From the uncertainty 3D view shown in Figure 7.9a, we observe relatively high uncertainties exhibited on the microphone’s capsule and its surroundings.

For the Uncertainty NN trained with the full-angle dataset, the synthetic view in Figure 7.9b shows the scene is rendered correctly; specifically, the right metal strip on the capsule is visible in the rendered scene. However, the relatively high uncertainty surrounding the microphone capsule is still persisting in the uncertainty 3D view (**Q1**). A more in-depth exploration into the limitations of the model and its behavior in this scene is necessary to explain this phenomenon, addressing **Q2**.

In Figure 7.9a, we observe that the mean and standard deviation uncertainties are relatively lower for the cells near the training images. However, the inclusion of additional training images in the full-angle dataset does not seem to eliminate the high mean and standard deviation uncertainties within the blue-bordered rectangle (**Q3**), as depicted in Figure 7.9b. This is attributed to the persistent uncertainties surrounding the capsule, as shown in the inserted uncertainty 3D view in Figure 7.9bA.

Microphone: Qualitative results for Ensemble

In this section, we explore the qualitative results for the microphone scene trained by the Ensemble model. Figure 7.6 presents the microphone scene’s synthetic view, color and density uncertainty 3D views trained with the Ensemble model, along with their respective TFs and 2D

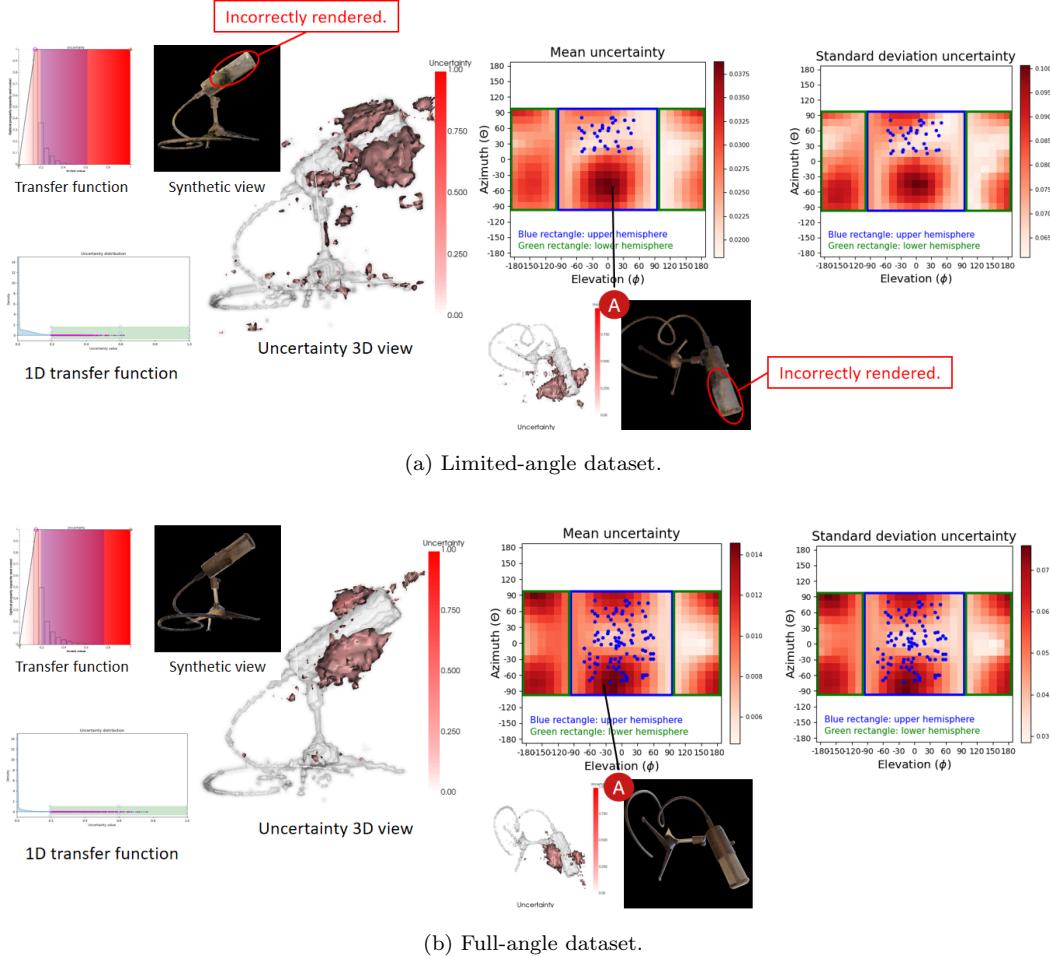
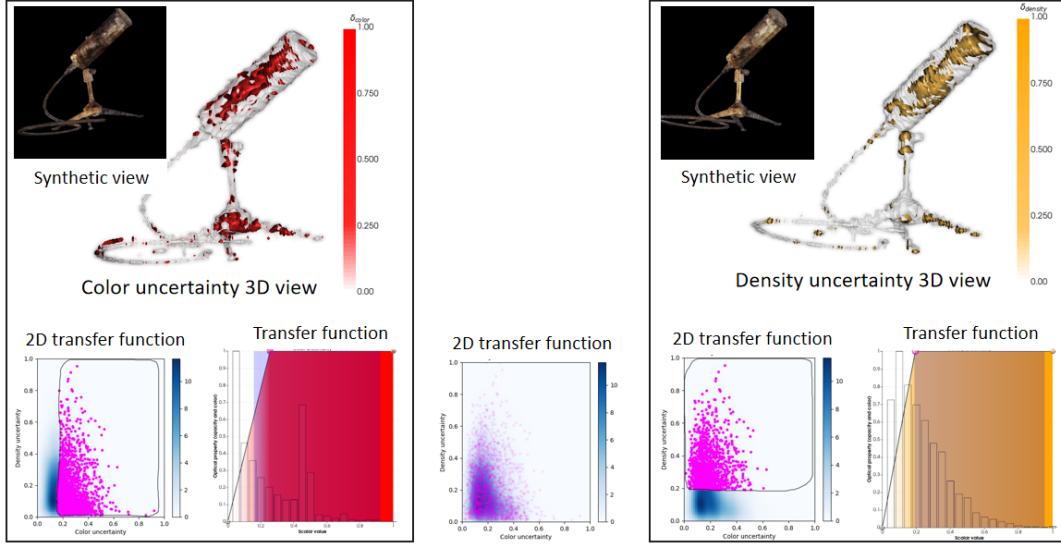


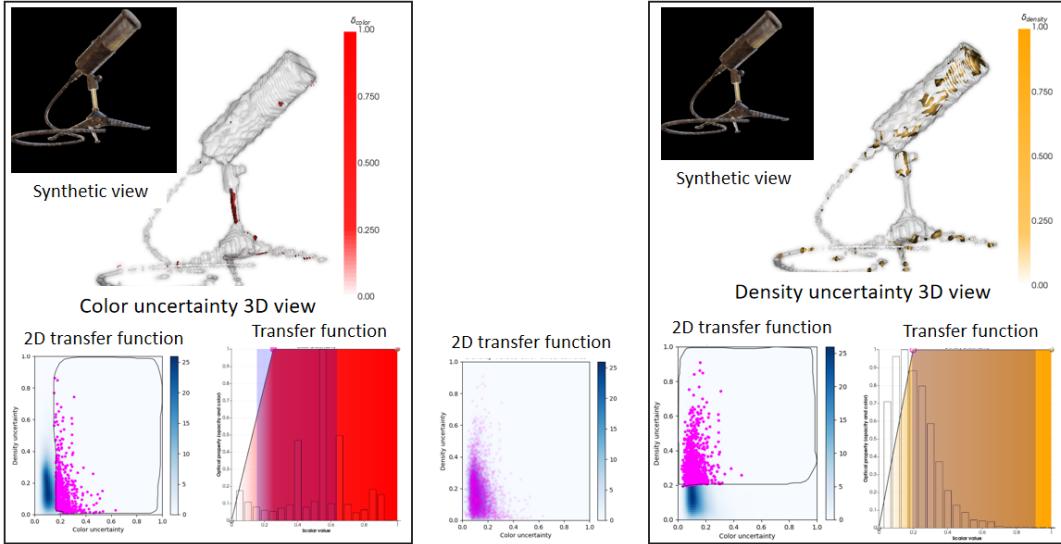
Figure 7.9: The synthetic view, uncertainty 3D view, transfer functions, 1D transfer functions, and the heatmaps for the microphone scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.

TF settings, for both the limited-angle and full-angle datasets. The results align with our expectations. This is indicated by the reduction in sample points with relatively high color and density uncertainties on the right side and the woven wire grill of the microphone capsule from the limited-angle dataset to the full-angle dataset (**Q3**). This trend can be observed in the uncertainty 3D views in Figure 7.10. The rendering quality improves from the limited-angle dataset to the full-angle dataset, as shown in the synthetic views in Figure 7.10. Additionally, in the 2D TF shown in Figure 7.10, we observe a higher concentration of points in the lower left corner for the full-angle dataset compared to the limited-angle dataset. This observation suggests a reduction in color and density uncertainty values resulting from further optimization of the Ensemble model with the full-angle dataset for this scene.

In the mean color uncertainty heatmap for the limited-angle dataset (Figure 7.11a), there is no clear pattern. The patterns in the mean color uncertainty heatmap become clearer for the full-angle dataset, as depicted in Figure 7.11b. However, the inclusion of additional training images in the full-angle dataset does not seem to eliminate the high mean uncertainties within the blue-bordered rectangle. This could be attributed to inaccurate color estimation in complex lighting conditions by the secondary multilayer perceptrons (MLPs), which have a relatively lower number of learning parameters. In short, the color estimations are sensitive to slight variations in the



(a) Limited-angle dataset.

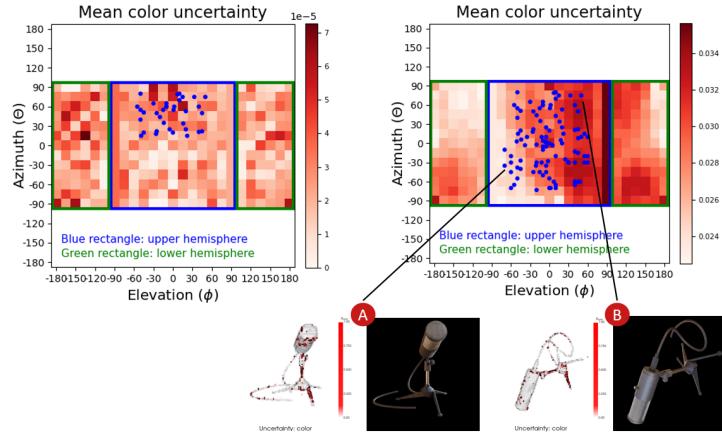


(b) Full-angle dataset.

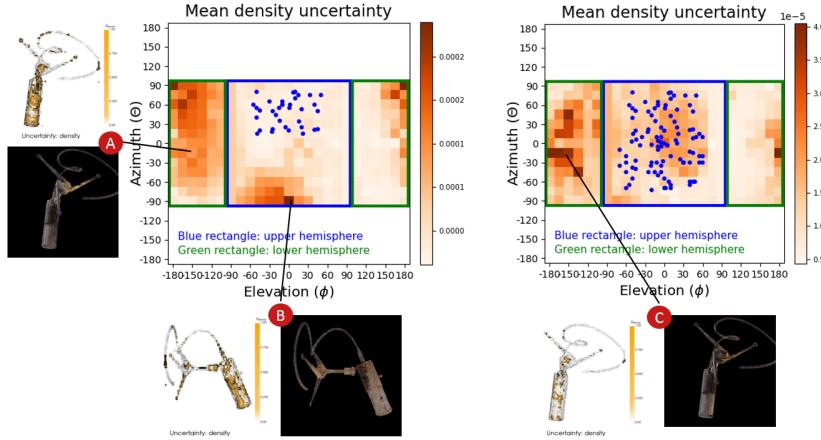
Figure 7.10: The synthetic view, color and density uncertainty 3D views, transfer functions and the 2D transfer function for the microphone scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.

viewing direction in this scene.

For the mean density uncertainty heatmaps in Figure 7.11, the mean density uncertainty values within the blue-bordered rectangle decrease from the limited-angle dataset to the full-angle dataset following the addition of training images (**Q3**). The patterns in these heatmaps are distinct, the clusters of cells with relatively high uncertainty in the green-bordered rectangle. correspond to viewing uncertainty from the lower hemisphere, as shown in Figure 7.11bC. This is because the viewing directions from the lower hemisphere are not included in both the limited-angle and full-angle training images.



(a) Mean color uncertainty heatmaps for the limited-angle dataset (left) and the full-angle dataset (right).



(b) Mean density uncertainty heatmaps for the limited-angle dataset (left) and the full-angle dataset (right).

Figure 7.11: The (a) mean color uncertainty heatmaps and (b) mean density uncertainty heatmaps in selected viewing directions for the microphone scene trained with the Ensemble model using the limited-angle dataset (left heatmap) and full-angle dataset (right heatmap). The pattern in the standard deviation uncertainty heatmaps is similar to their respective mean uncertainty heatmaps. The standard deviation uncertainty heatmaps can be found in Figure A.1 in Appendix A.

Microphone: Qualitative results comparison between Uncertainty NN and Ensemble

Figure 7.12 presents a comparison between the cross-sectional uncertainty 3D views and normalized mean uncertainty heatmaps of the Uncertainty NN and Ensemble models for the microphone scene. The cross-sectional uncertainty 3D view of the microphone from the Ensemble model correctly predicts relatively high uncertainty within the capsule, as shown in Figure 7.12b. In contrast, the Uncertainty NN model shows minimal indication of high uncertainty within the microphone capsule, as illustrated in Figure 7.12a. This observation aligns with the chair scene.

The Ensemble's mean density uncertainty heatmap (Figure 7.12b) exhibits a clearer pattern,

facilitating users' understanding and interpretation of the presented information. In both the Uncertainty NN's mean uncertainty heatmap (Figure 7.12a) and the Ensemble's mean color uncertainty heatmap (Figure 7.12b), higher variability is observed in the vicinity of the training images, as color estimation is sensitive to complex lighting conditions in the microphone scene.

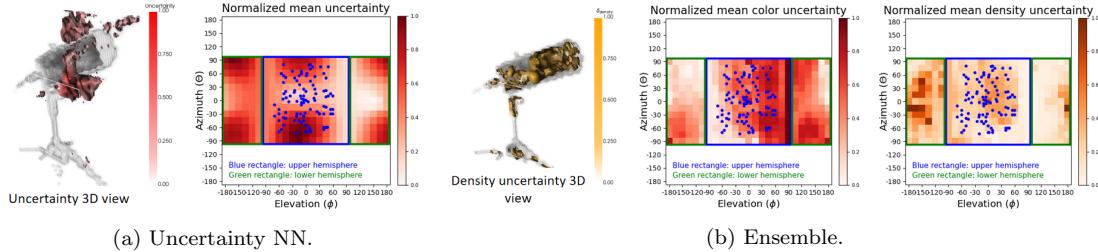


Figure 7.12: The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the microphone scene. The mean uncertainty values in the heatmaps have been normalized.

7.3.3 Dataset: drums

Drums: Qualitative results for Uncertainty NN

The drums scene poses a higher level of complexity compared to the chair and mic scenes, featuring intricate details, particularly in the cymbal stands and hoops in the toms and drums, as well as complex lighting conditions on the drums, toms, and cymbals. In this section, we explore the uncertainty visualization for this scene trained by the Uncertainty NN. Figure 7.13 displays the drums scene's synthetic and uncertainty 3D views trained with the Uncertainty NN, along with the TFs settings, for both the limited-angle and full-angle datasets. We expect relatively high uncertainty in the pedal and bottom of the bass drum, as these regions are not visible in the limited-angle dataset. However, we have discovered that relatively high uncertainty does not appear in these regions, as observed in the uncertainty 3D views in Figure 7.13a. This observation contradicts our expectations and it is similar to our findings in the chair scene trained by the Uncertainty NN.

For the Uncertainty NN model trained with the full-angle dataset, we expect to see relatively high uncertainty at the bottom of the bass drum because this region is not visible in the training images. In the synthetic view shown in Figure 7.13b, we observe that this region is rendered incorrectly, as expected. However, high uncertainty is not visibly exhibited in this region, as shown in the uncertainty 3D view in Figure 7.13b.

Figure 7.14 shows the comparison of the heatmaps between the limited-angle and full-angle datasets for this drums scene. We observe similar findings as in the chair scene trained by the Uncertainty NN. In Figure 7.14a, the mean uncertainties in the vicinity of the training images are relatively lower. This suggests that the viewing directions of the limited-angle dataset training images influence the uncertainty distribution in the scene (**Q3**). In Figure 7.14b, the mean and standard deviation uncertainty values decrease after training with the full-angle dataset, but the cells near the top left corner of the blue-bordered rectangle still exhibit higher uncertainty. This is attributed to the persistence of uncertainty at the stand of the cymbal, as depicted in the Figure 7.14bB, even though additional training images have been given to train the model.

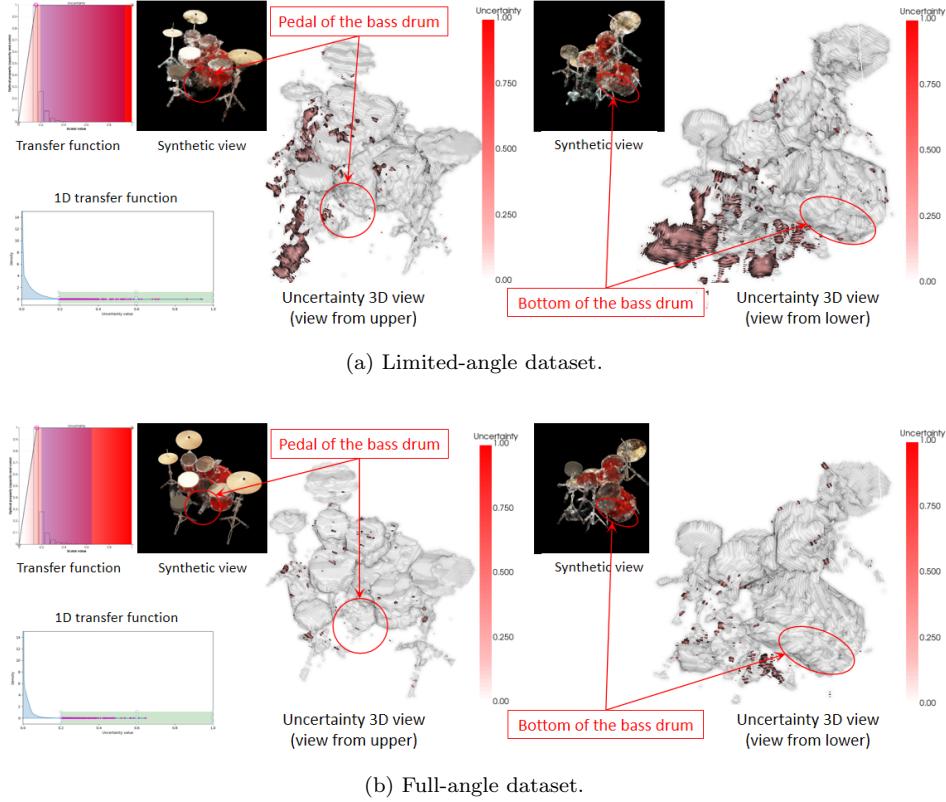


Figure 7.13: The synthetic view, uncertainty 3D view, transfer function, and 1D transfer function for the drums scene trained with the Uncertainty NN model using (a) limited-angle dataset and (b) full-angle dataset.

Drums: Qualitative results for Ensemble

Moving on to the exploration of the drums scene trained by the Ensemble model. Figure 7.15a and Figure 7.15b present two synthetic views and two uncertainty 3D views each, along with the TFs settings. The left view is from the upper hemisphere, while the right view is from the lower hemisphere. The same applies to Figure 7.16a and Figure 7.16b.

We first explore the color uncertainty results in Figure 7.15. In Figure 7.15a, we observe inaccuracies in the rendering of the pedal of the bass drum, toms, and cymbals in the synthetic view, with relatively high uncertainty exhibited in these regions in the uncertainty 3D view for the limited-angle dataset. This might be attributed to the intricate lighting conditions on the reflective surfaces of these drum set components. Following further optimization with the full-angle dataset, the color uncertainties are mitigated, as shown in Figure 7.15b. This observation is evident from the 2D TF in Figure 7.15, where a noticeable number of points along color uncertainty shift to lower color uncertainty values through further optimization of the model with the full-angle dataset. The color uncertainty 3D views, when viewed from the lower hemisphere for both limited-angle and full-angle datasets are comparable, as the training images lack views from the lower hemisphere in both datasets.

Here, we explore the density uncertainty results in Figure 7.16. We observe a moderate reduction in density uncertainty in the density uncertainty 3D view when viewing from the upper hemisphere, from the limited-angle dataset to the full-angle dataset. This observation is supported by the 2D TF in Figure 7.16, where a limited number of points along density uncertainty shift to

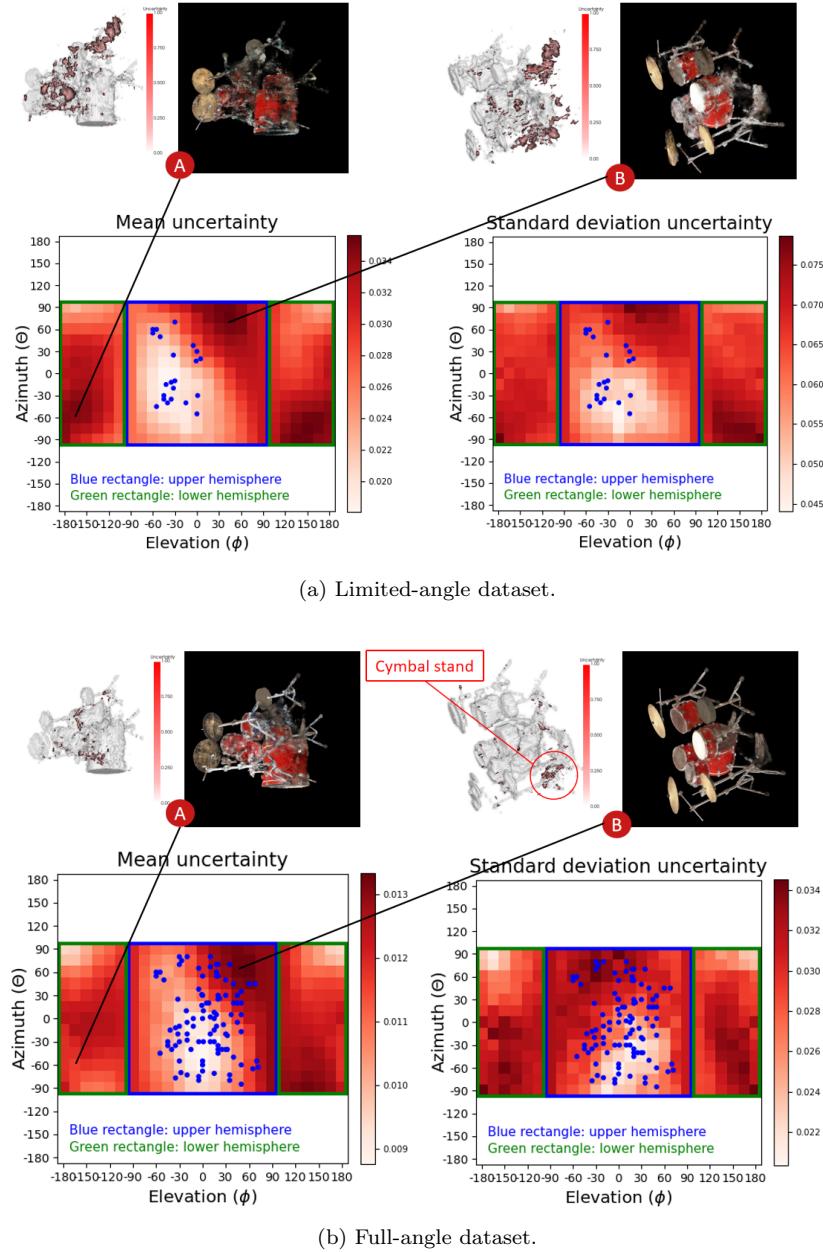
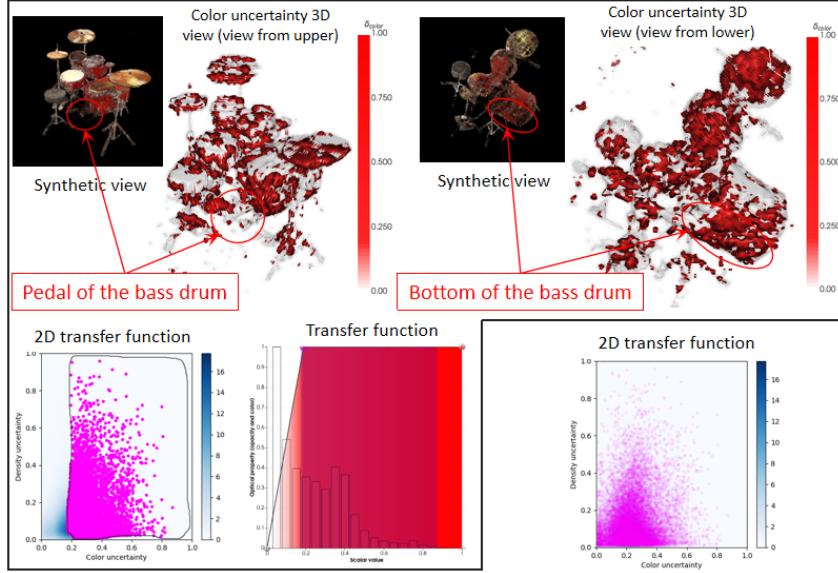


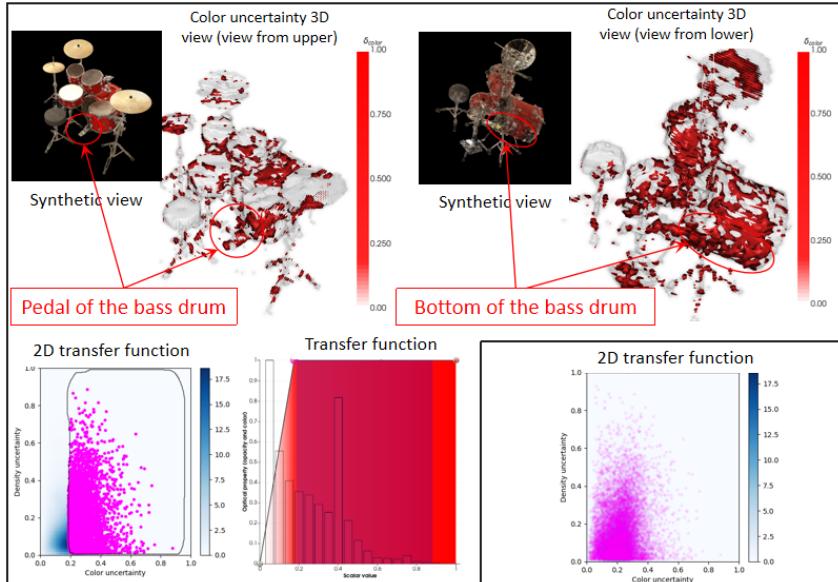
Figure 7.14: The mean and standard deviation uncertainty heatmaps in selected viewing directions for the drums scene trained with the Uncertainty NN model using (a) limited-angle and (b) full-angle datasets.

lower density uncertainty values from the limited-angle to full-angle datasets. This may be due to the intricate geometry in the scene. It is worth noting that relatively high density uncertainty is not exhibited at the bottom of the bass drum for both datasets. This contradicts our expectation, considering the bottom of the bass drum is not visible in any training images.

The comparison of heatmaps between the limited-angle and full-angle dataset for the mean color and mean density uncertainties are shown in Figure 7.17. In the mean color uncertainty heatmaps (Figure 7.17a), surprisingly, the mean color uncertainty value increases in the full-angle dataset compared to the limited-angle dataset. This observation suggests the need for further in-



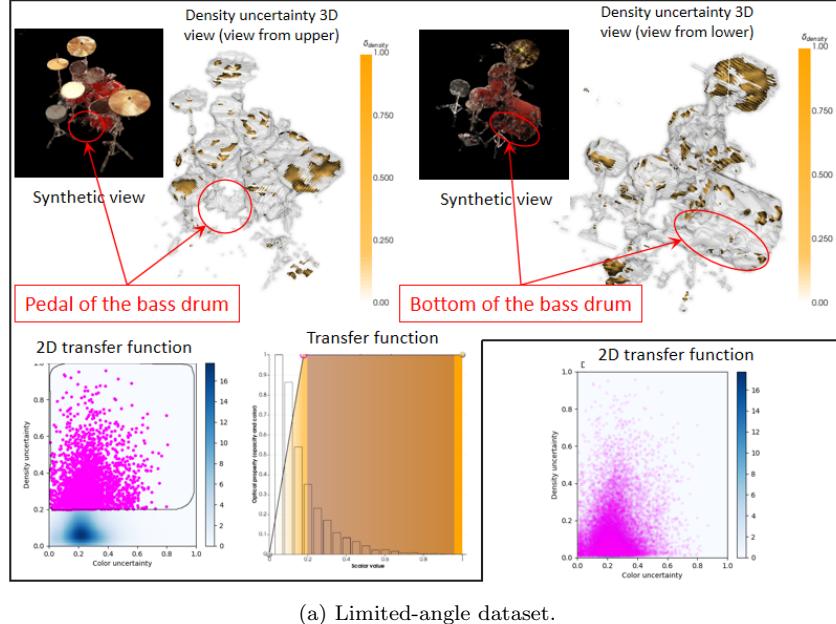
(a) Limited-angle dataset.



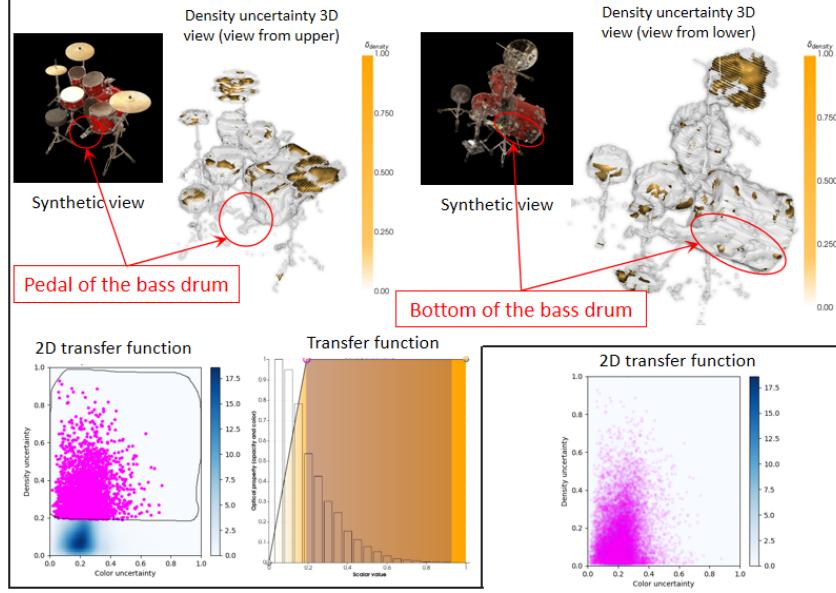
(b) Full-angle dataset.

Figure 7.15: The synthetic view, color uncertainty 3D views, transfer functions, and the 2D transfer functions for the drums scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.

vestigation in future work. The patterns in the mean density uncertainty heatmaps are distinct, as seen in Figure 7.17b. The mean density uncertainty values near the training images decrease from the limited-angle dataset to the full-angle dataset (**Q3**). The cells with relatively high mean density uncertainty in the blue-bordered and green-bordered rectangles represent viewing uncertainty from the upper and lower hemispheres, respectively, as depicted in Figure 7.17bA to Figure 7.17bD.



(a) Limited-angle dataset.

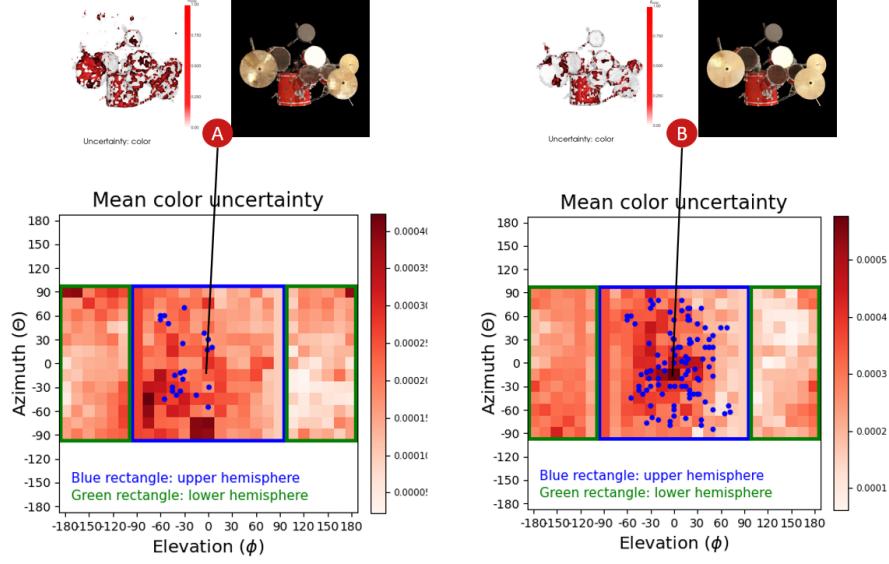


(b) Full-angle dataset.

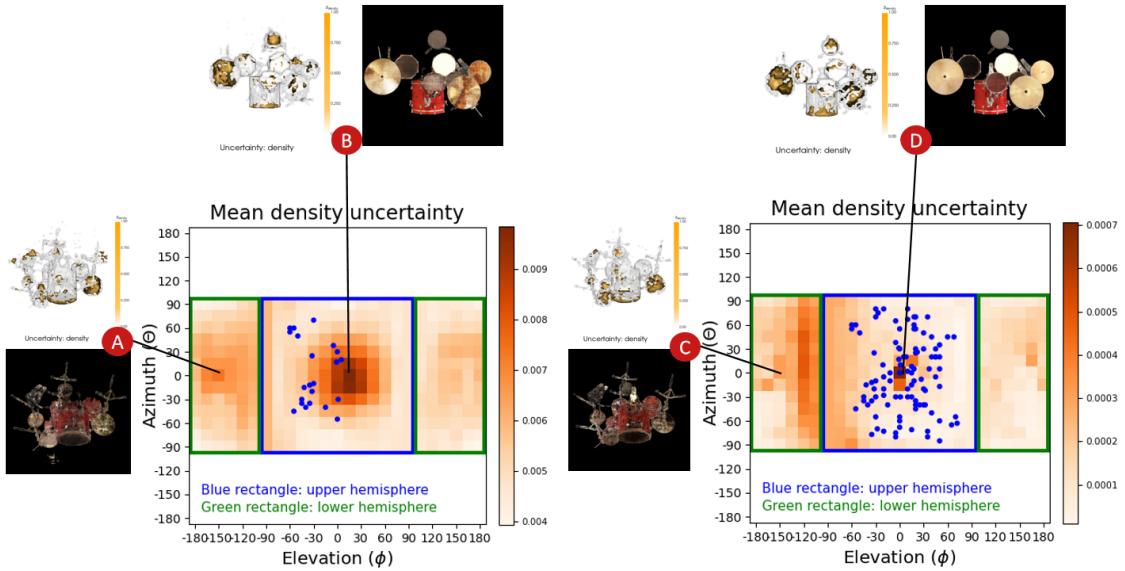
Figure 7.16: The synthetic view, density uncertainty 3D views, transfer functions, and the 2D transfer functions for the drums scene trained with the Ensemble model using (a) limited-angle dataset and (b) full-angle dataset.

Drums: Qualitative results comparison between Uncertainty NN and Ensemble

Figure 7.18 presents a comparison between the cross-sectional uncertainty 3D views and normalized mean uncertainty heatmaps of the Uncertainty NN and Ensemble models for the drums scene. As expected, the cross-sectional uncertainty 3D view of the drums from the Ensemble model correctly predicts relatively high uncertainty within the bass drum, as shown in Figure 7.18b. In contrast, the Uncertainty NN model does not exhibit high uncertainty within the bass drum, as



(a) Mean color uncertainty heatmaps for the limited-angle dataset (left) and the full-angle dataset (right).



(b) Mean density uncertainty heatmaps for the limited-angle dataset (left) and the full-angle dataset (right).

Figure 7.17: The (a) mean color uncertainty heatmaps and (b) mean density uncertainty heatmaps in selected viewing directions for the drums scene trained with the Ensemble model using limited-angle dataset (left heatmap) and full-angle dataset (right heatmap). The pattern in the standard deviation uncertainty heatmaps is similar to their respective mean uncertainty heatmaps. The standard deviation uncertainty heatmaps can be found in Figure A.2 in Appendix A.

shown in Figure 7.18a. This observation aligns with the chair and microphone scenes.

For the normalized heatmaps of both models in Figure 7.18, the observation is similar to the chair and microphone scenes. However, the pattern in the Ensemble's mean color uncertainty

heatmap is relatively unclear. This might indicate a more intricate relationship between color uncertainties and scene features in the drums scene, such as complex lighting conditions (**Q1**).

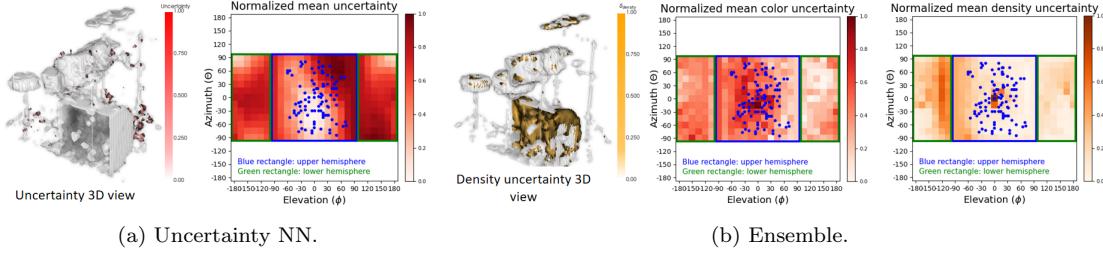


Figure 7.18: The uncertainty 3D views and mean uncertainty heatmaps from the (a) Uncertainty NN model and (b) Ensemble model for the drums scene. The mean uncertainty values in the heatmaps have been normalized.

7.3.4 Summary and discussion of findings

In the previous sections, we demonstrate the interactive exploration of uncertainty visualization within the NeRFDeltaView system, focusing on three distinct use cases. We correlate the exploration of uncertainty visualization with the three potential user questions (**Q1**, **Q2**, and **Q3**) outlined in Chapter 4.

Thanks to the designed visualization, we are able to explore and understand the model uncertainty by effectively gaining insight into the scene features (**Q1**), identifying the limitations of the model (**Q2**), and the correlation between the quantity and viewing direction of training images and model uncertainty (**Q3**). We will summarize the findings of the exploration in the subsequent paragraph.

The Uncertainty NN model demonstrated its sensitivity of uncertainty estimation to viewing directions in the mean uncertainty heatmap, a similar phenomenon for the color uncertainty heatmap of the Ensemble model. Overall, the density uncertainty estimation of the Ensemble model meets our expectations. It correctly identifies relatively high density uncertainty in regions of the chair and microphone scenes that were not visible in the training images. Specifically, the density uncertainty estimation of the Ensemble model accurately estimates relatively high density uncertainty within the interior of an isosurface for all three scenes. Both the Uncertainty NN and the Ensemble model encounter challenges in estimating uncertainties in intricate scenes, such as the drums scene with complex geometry and lighting conditions. For the findings that do not reflect our expectations, a more in-depth investigation is necessary to better understand the limitations of the model in these scenes. By acknowledging these limitations in the model uncertainty through our explorations, we realize the necessity for future work.

The developed visualization approach has demonstrated effectiveness in exploring uncertainty in a synthesized scene and understanding NeRF model uncertainty. The incorporation of interactive features such as uncertainty 3D views, cutting plane widget, and heatmaps enhance user exploration and understanding of the model uncertainty. For example, the heatmaps provide a comprehensive perspective on uncertainty patterns across the scene, enabling users to swiftly identify viewing directions with high uncertainty. Additionally, the adaptability of the NeRF-DeltaView system across different NeRF models and scenes emphasizes its strong versatility.

While the visualization approach has several strengths, there are also some potential limitations. In Section 3.3, we have discussed the sources of uncertainty in the visualization pipeline

according to Weiskopf [84]. The generation of uncertainty 3D grids aligns with the transformation stage in the visualization pipeline (Figure 3.4), addressing the uncertainty embedded in the inadequate or limited number of sample points in the grid used to represent the complex scene. The heatmaps are preprocessed before implementing the visualization tool, and the generation of heatmaps is aligned with the visual mapping and rendering stage in the visualization pipeline (Figure 3.4). The uncertainty statistics in the heatmaps could vary depending on the user-defined isovales and transfer function settings. These two stages added uncertainty to our visualization pipeline, and we are not able to fully account for their impact on the final uncertainty visualizations. Additionally, the effectiveness of the visualization approach may be sensitive to various parameter settings, such as the selection of isovales, transfer function settings, and selection of points in 2D transfer function. These settings could require expertise and involve a time-consuming process.

Chapter 8

Conclusion

Neural Radiance Field (NeRF) has demonstrated promising performances in synthesizing realistic three-dimensional (3D) scenes from a set of two-dimensional (2D) images, leveraging deep neural networks. However, it poses inherent challenges associated with uncertainty [61], arising from various sources, including a limited amount of training images, noisy image data, and scenes with intricate geometry. For instance, if a NeRF model is trained with a set of training images with a limited range of viewing angles, it may introduce errors in the synthesized scene that can be interpreted as uncertainty. Therefore, it is important for researchers and practitioners to understand and assess the reliability of the synthesized scenes created by NeRF. Incorporating uncertainty estimation in the synthesized scenes and implementing uncertainty visualization is a key approach to achieving this goal. Interactive uncertainty visualization enables users to explore the estimated uncertainty in a synthesized scene effectively, aiding in understanding NeRF model uncertainty.

The main objective of this research is to design an uncertainty visualization system that facilitates exploration and understanding of the NeRF model uncertainty. We used two uncertainty estimation approaches with NeRF. The first model, Uncertainty Neural Network (Uncertainty NN), employs a single deterministic method where uncertainty is directly predicted by a single deterministic neural network. The second model, Ensemble, leverages uncertainties estimated from multiple different deterministic neural networks during the inference process.

We have developed the NeRFDeltaView (Neural Radiance Field Delta View) system, which includes NeRF models with uncertainty estimation and interactive visualization designed to visualize the uncertainties estimated by these models. To effectively address the objectives of the NeRFDeltaView system, three potential user questions are outlined: (**Q1**) exploring the relationship between scene features and model uncertainty, (**Q2**) identifying model limitations through model uncertainty analysis, and (**Q3**) assessing the correlation between the quantity and viewing direction of training images with model uncertainty. Our visualization tool offers various interactive views to assist users in exploring uncertainty in a synthesized scene. These views are aligned with the three user questions; for example, the uncertainty 3D views facilitate users to explore specific scene features in the synthesized scene and relate them to model uncertainty (**Q1**) and identify the limitations in the model through uncertainty distribution in the synthesized scene (**Q2**). Moreover, the heatmaps offer a comprehensive perspective on the overall uncertainty pattern, facilitating users in investigating the influence of varying numbers and viewing directions of training images on the model uncertainty (**Q3**).

We created three use cases to verify NeRFDeltaView and enhance our understanding of the model based on the model uncertainty analysis. These three use cases cover synthetic scenes of varying complexity in geometry and lighting conditions. We demonstrated that NeRFDeltaView could be used to explore and analyze the uncertainty distribution in a synthesized scene from both an overview and volume rendering perspective, enabling us to understand the limitations of

the models. From the overview perspective, the Ensemble model’s results meet our expectations; specifically, the density mean and standard deviation uncertainties decrease with the addition of training images. On the other hand, the heatmaps from the Uncertainty NN model demonstrate its sensitivity to uncertainty estimation based on viewing direction (**Q3**). From the volume rendering perspective, compared to Uncertainty NN, the Ensemble model generally aligns with our expectations, particularly the density uncertainty estimation. That is, the relatively high density uncertainty is exhibited in the expected regions that were not visible in the training images (**Q1**). Both models encounter challenges in accurately estimating uncertainty when dealing with scenes with intricate details, such as complex geometry and lighting conditions. This highlights the need for further work to investigate the limitations in the model (**Q2**) and potential improvements to the model architecture. For example, refining the secondary multilayer perceptrons in the Ensemble model with a high number of parameters, and improving the uncertainty prediction approach in the Uncertainty NN model.

While NeRFDeltaView offers valuable insights for exploring uncertainty in the synthesized scene and understanding model uncertainty, we acknowledge certain inherent limitations regarding 3D regular grid generation and uncertainty visualization. Firstly, a limitation in generating the uncertainty 3D regular grid is that we set the input viewing direction to zero, even though color estimation is direction-dependent. This might lead to a less accurate representation of color uncertainty. While generating multiple 3D grids by considering the specific viewing direction could enhance accuracy, this approach could be time and memory-consuming. Secondly, our approach of mapping the 3D sphere as a 2D heatmap for the overview of the uncertainty pattern is considered a limitation. Although concise, this mapping may result in a loss of spatial context and distortion. Users might struggle to understand the distribution of uncertainties across various spatial directions in the perspective overview. A potential improvement could be a 2D circular heatmap [47] or a sphere heatmap inspired by Kwon et al. [40]. The uncertainty statistics could be projected on the interactive sphere surface. Both the 2D circular heatmap and sphere heatmap could overcome the limitation of loss of spatial context and distortion. Thirdly, the uncertainty statistics in the heatmaps could vary depending on the user-defined isovales and transfer function settings, and the number of sample points in the uncertainty 3D grids may be inadequate or limited to represent the complex scene. These variations add uncertainty to our visualization pipeline, and we are not able to fully account for their impact on the final uncertainty visualizations. A potential approach for the heatmaps could be to conduct sensitivity analyses by comparing the differences between several heatmaps generated from different isovales and transfer function settings. Lastly, the visualization may be sensitive to parameter settings, such as isovales and transfer functions, may require expertise and involve a time-consuming process. Implementing an automatic transfer function design [45, 37, 58, 83, 89] for visualizing volumetric data might help streamline this process.

In Chapter 3, we mentioned that Deep Ensemble [41] and Monte Carlo Dropout [16] are the two popular methods for incorporating uncertainty estimation into deep neural networks. We implemented the Ensemble model inspired by the Deep Ensemble. The uncertainty estimation using Monte Carlo Dropout could be implemented in future work. However, modifications to the implementation are necessary; the dropout layers are deactivated during training, but they are activated to estimate uncertainty during the inference phase. The advantage of adapting Monte Carlo Dropout is that only one MLP is trained, in contrast to multiple MLPs in the Ensemble model.

In this research, we have demonstrated that NeRFDeltaView effectively facilitates the exploration of the uncertainty in a synthesized scene and the analysis of model uncertainty through interactive uncertainty visualizations. Furthermore, we demonstrate our interactive volume rendering uncertainty visualization, which aids in identifying limitations in models, and an uncertainty overview to investigate the relationship between the amount and viewing direction of training images with model uncertainty. Researchers and practitioners could utilize this system to identify model limitations integrated with uncertainty estimation, contributing to the improvement of the

model architecture and facilitating informed and confident decision-making. This is particularly relevant in assessing potential risks in medical imaging and autonomous driving applications.

Bibliography

- [1] Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU), 2018. cite arxiv:1803.08375.
- [2] Charu C. Aggarwal. *Neural Networks and Deep Learning*. Springer, Cham, 2018.
- [3] James Ahrens, Berk Geveci, and Charles Law. ParaView: An end-user tool for large data visualization. In *Visualization Handbook*. Elesvier, 2005. ISBN 978-0123875822.
- [4] Rohith Aralikatti, D. Margam, Tanay Sharma, Abhinav Thanda, and S. Venkatesan. Global SNR Estimation of Speech Signals using Entropy and Uncertainty Estimates from Dropout Networks. In *INTERSPEECH*, 2018.
- [5] Arsenii Ashukha, Alexander Lyzhov, Dmitry Molchanov, and Dmitry P. Vetrov. Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning. In *ICLR*, 2020.
- [6] Gwangbin Bae, Ignas Budvytis, and Roberto Cipolla. Estimating and Exploiting the Aleatoric Uncertainty in Surface Normal Estimation. In *ICCV*, 2021.
- [7] Jiayang Bai, Letian Huang, Wen Gong, Jie Guo, and Yanwen Guo. Self-NeRF: A Self-Training Pipeline for Few-Shot Neural Radiance Fields, 2023.
- [8] Nathan Bell and Michael Garland. Efficient Sparse Matrix-Vector Multiplication on CUDA. NVIDIA Technical Report NVR-2008-004, NVIDIA Corporation, December 2008.
- [9] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight Uncertainty in Neural Networks. In *ICML*, 2015.
- [10] Eric Brachmann, Frank Michel, Alexander Krull, Michael Ying Yang, Stefan Gumhold, and Carsten Rother. Uncertainty-Driven 6D Pose Estimation of Objects and Scenes from a Single RGB Image. In *CVPR*, 2016.
- [11] A. Cedilnik and P. Rheingans. Procedural annotation of uncertain information. In *Proceedings Visualization 2000. VIS 2000 (Cat. No.00CH37145)*, pages 77–84, 2000.
- [12] Alexandre Coninx, Georges-Pierre Bonneau, Jacques Droulez, and Guillaume Thibault. Visualization of Uncertain Scalar Data Fields using Color Scales and Perceptually Adapted Noise. In *Applied Perception in Graphics and Visualization*, 2011.
- [13] Li Deng and John Platt. Ensemble Deep Learning for Speech Recognition. In *Proc. Interspeech*, September 2014.
- [14] Newton Der and Alex Pang. Direct Volume Rendering of Multi-Valued Data Sets. 08 2006.
- [15] Suzana Djurcicov, Kwansik Kim, Pierre Lermusiaux, and Alex Pang. Visualizing scalar volumetric data with uncertainty. *Computers Graphics*, 26(2):239–248, 2002.
- [16] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *ICML*, 2016.

BIBLIOGRAPHY

- [17] Jakob Gawlikowski, Cedrique Tassi, Mohsin Ali, Jongseok Lee, Matthias Humt, Jianxiang Feng, Anna Kruspe, Rudolph Triebel, Peter Jung, Ribana Roscher, Muhammad Shahzad, Wen Yang, Richard Bamler, and Xiao Zhu. A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56:1–77, 07 2023.
- [18] Yonatan Geifman, Guy Uziel, and Ran El-Yaniv. Bias-Reduced Uncertainty Estimation for Deep Neural Classifiers. In *ICLR*, 2019.
- [19] G. Grigoryan and P. Rheingans. Point-based probabilistic surfaces to show surface uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):564–573, 2004.
- [20] Fredrik K. Gustafsson, Martin Danelljan, and Thomas B. Schon. Evaluating Scalable Bayesian Deep Learning Methods for Robust Computer Vision. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1289–1298, 2020.
- [21] Giannis Haralabopoulos, Ioannis Anagnostopoulos, and Derek McAuley. Ensemble Deep Learning for Multilabel Binary Classification of User-Generated Content. *Algorithms*, 13(4), 2020.
- [22] D.D. Hearn, P. Baker, and W. Carithers. *Computer Graphics with Open GL: Pearson New International Edition*. Pearson Education, 2013.
- [23] Dan Hendrycks, Mantas Mazeika, Saurav Kadavath, and Dawn Xiaodong Song. Using Self-Supervised Learning Can Improve Model Robustness and Uncertainty. In *NeurIPS*, 2019.
- [24] T. Hengl. Visualisation of uncertainty using the HSI colour model : computations with colours. In D. Martin, editor, *Proceedings of the 7th International conference on GeoComputation, University of Southampton, United Kingdom, 8-10 September 2003. / ed. by. D. Martin. 8 p.*, United Kingdom, 2003. University of Southampton.
- [25] S. Hernández, Diego Vergara, Matias Valdenegro-Toro, and Felipe Jorquera. Improving predictive uncertainty estimation using Dropout–Hamiltonian Monte Carlo. *Soft Computing*, 24:4307–4322, 2020.
- [26] Matthew D. Hoffman, Tuan Anh Le, Pavel Sountsov, Christopher Suter, Ben Lee, Vikash K. Mansinghka, and Rif A. Saurous. ProbNeRF: Uncertainty-Aware Inference of 3D Shapes from 2D Images. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 10425–10444. PMLR, 25–27 Apr 2023.
- [27] Stephen C. Hora. Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management. *Reliability Engineering System Safety*, 54(2):217–223, 1996. Treatment of Aleatory and Epistemic Uncertainty.
- [28] Jing Hu, Qinrui Fan, Shu Hu, Siwei Lyu, Xi Wu, and Xin Wang. Uncertainty-aware Single View Volumetric Rendering for Medical Neural Radiance Fields, 2023.
- [29] John F. Hughes, Andries van Dam, Morgan McGuire, David F. Sklar, James D. Foley, Steven Feiner, and Kurt Akeley. *Computer Graphics: Principles and Practice*. Addison-Wesley, Upper Saddle River, NJ, 3 edition, 2013.
- [30] G. J. Hunter and M. F. Goodchild. Managing uncertainty in spatial databases: putting theory into practice. *URISA Journal*, 5(2):55–62, January 1993.
- [31] J. D. Hunter. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

- [32] Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110, 03 2021.
- [33] Liren Jin, Xieyuanli Chen, Julius Rückin, and Marija Popović. NeU-NBV: Next Best View Planning Using Uncertainty Estimation in Image-Based Neural Rendering, 2023.
- [34] C.R. Johnson and A.R. Sanderson. A Next Step: Visualizing Errors and Uncertainty. *IEEE Computer Graphics and Applications*, 23(5):6–10, 2003.
- [35] James Kajiya and Brian von herzen. Ray Tracing Volume Densities. *ACM SIGGRAPH Computer Graphics*, 18:165–174, 07 1984.
- [36] Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [37] Gordon Kindlmann and James W. Durkin. Semi-automatic generation of transfer functions for direct volume rendering. In *Proceedings of the 1998 IEEE Symposium on Volume Visualization*, VVS ’98, page 79–86, New York, NY, USA, 1998. Association for Computing Machinery.
- [38] Diederik Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.
- [39] Armen Der Kiureghian and Ove Ditlevsen. Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2):105–112, 2009. Risk Acceptance and Risk Communication.
- [40] Oh-Hyun Kwon, Chris Muelder, Kyungwon Lee, and Kwan-Liu Ma. A Study of Layout, Rendering, and Interaction Methods for Immersive Graph Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1802–1815, jul 2016.
- [41] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [42] Soomin Lee, Le Chen, Jiahao Wang, Alexander Liniger, Suryansh Kumar, and Fisher Yu. Uncertainty Guided Policy for Active Robotic 3D Reconstruction Using Neural Radiance Fields. *IEEE Robotics and Automation Letters*, 7(4):12070–12077, 2022.
- [43] M. Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37, 1988.
- [44] Jeremiah Liu, John William Paisley, Marianthi-Anna Kioumourtzoglou, and B. Coull. Accurate Uncertainty Estimation and Decomposition in Ensemble Learning. In *NeurIPS*, 2019.
- [45] Patric Ljung, Jens Krüger, Eduard Groller, Markus Hadwiger, Charles D. Hansen, and Anders Ynnerman. State of the Art in Transfer Functions for Direct Volume Rendering. *Computer Graphics Forum*, 35(3):669–691, 2016.
- [46] Claes Lundström, Patric Ljung, Anders Persson, and Anders Ynnerman. Uncertainty Visualization in Medical Volume Rendering Using Probabilistic Animation. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1648–1655, nov 2007.
- [47] Kirsten W. H. Maas, Nicola Pezzotti, Amy J. E. Vermeer, Danny Ruijters, and Anna Vilanova. NeRF for 3D Reconstruction from X-ray Angiography: Possibilities and Limitations. In Christian Hansen, James Procter, Renata G. Raidou, Daniel Jönsson, and Thomas Höllt, editors, *Eurographics Workshop on Visual Computing for Biology and Medicine*. The Eurographics Association, 2023.

BIBLIOGRAPHY

- [48] Andrey Malinin and Mark John Francis Gales. Predictive Uncertainty Estimation via Prior Networks. In *NeurIPS*, 2018.
- [49] Refik Can Malli, Mehmet Aygun, and Hazim Kemal Ekenel. Apparent Age Estimation Using Ensemble of Deep Learning Models. *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 714–721, 2016.
- [50] Steve Marschner and Peter Shirley. *Fundamentals of Computer Graphics, Fourth Edition*. A. K. Peters, Ltd., USA, 4th edition, 2016.
- [51] Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7210–7219, June 2021.
- [52] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.
- [53] Nelson Max. Optical Models for Direct Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.
- [54] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *ECCV*, 2020.
- [55] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. *ACM Trans. Graph.*, 41(4), July 2022.
- [56] Radford M. Neal. Bayesian Learning for Neural Networks. 1995.
- [57] Richard E. Neapolitan. Learning Bayesian networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining(KDD)*, 2007.
- [58] Binh Nguyen, Wei-Liang Tay, Chee-Kong Chui, and Sim Ong. Automatic Transfer Function Design for Volumetric Data Visualization using Clustering on LH Space. 06 2011.
- [59] NVIDIA, Péter Vingermann, and Frank H.P. Fitzek. CUDA, release: 10.2.89, 2020.
- [60] Andrés Ortiz, Jorge Munilla, Juan M. Górriz, and Javier Ramírez. Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer’s Disease. *International Journal of Neural Systems*, 26(07):1650025, 2016.
- [61] Xuran Pan, Zihang Lai, Shiji Song, and Gao Huang. ActiveNeRF: Learning Where to See with Uncertainty Estimation. In Shai Avidan, Gabriel Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner, editors, *Computer Vision – ECCV 2022*, pages 230–246, Cham, 2022. Springer Nature Switzerland.
- [62] Alex Pang, Craig Wittenbrink, and Suresh Lodha. Approaches to Uncertainty Visualization. *The Visual Computer*, 13, 10 1996.
- [63] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. In *CVPR*, 2019.
- [64] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in PyTorch. 2017.

- [65] Tobias Pfaffelmoser, Matthias Reitinger, and Rüdiger Westermann. Visualizing the positional and geometrical variability of isosurfaces in uncertain scalar fields. In *Proceedings of the 13th Eurographics / IEEE - VGTC Conference on Visualization*, EuroVis'11, page 951–960, Chichester, GBR, 2011. The Eurographics Association & John Wiley & Sons, Ltd.
- [66] Kristin Potter. *The Visualization of Uncertainty*. PhD thesis, 2010.
- [67] Kristin Potter, Joe Kniss, Richard Riesenfeld, and Chris R. Johnson. Visualizing Summary Statistics and Uncertainty. *Computer Graphics Forum*, 2010.
- [68] Bernhard Preim and Dirk Bartz, editors. *Visualization in Medicine*. Morgan-Kaufmann-Verlag, 2007.
- [69] Bernhard Preim and Charl P Botha. *Visual Computing for Medicine*. Morgan Kaufmann Publishers, 2013.
- [70] PyQt. PyQt Reference Guide. 2012.
- [71] Chao Qu, Wenxin Liu, and Camillo Jose Taylor. Bayesian Deep Basis Fitting for Depth Completion with Uncertainty. In *ICCV*, 2021.
- [72] R.G. Raidou. *Visual analytics for digital radiotherapy : towards a comprehensible pipeline*. Phd thesis 1 (research tu/e / graduation tu/e), Biomedical Engineering, March 2017. Proefschrift.
- [73] Yunlong Ran, Jing Zeng, Shibo He, Jiming Chen, Lincheng Li, Yingfeng Chen, Gimhee Lee, and Qi Ye. NeurAR: Neural Uncertainty for Autonomous 3D Reconstruction With Implicit Neural Representations. *IEEE Robotics and Automation Letters*, 8(2):1125–1132, 2023.
- [74] Penny Rheingans and Shrikant Joshi. Visualization of Molecules with Positional Uncertainty. In Eduard Gröller, Helwig Löffelmann, and William Ribarsky, editors, *Data Visualization '99*, pages 299–306, Vienna, 1999. Springer Vienna.
- [75] Hippolyt Ritter, Aleksandar Botev, and David Barber. A Scalable Laplace Approximation for Neural Networks. In *ICLR*, 2018.
- [76] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit (4th ed.)*. Kitware, 2006.
- [77] Jianxiong Shen, Antonio Agudo, Francesc Moreno-Noguer, and Adria Ruiz. Conditional-Flow NeRF: Accurate 3D Modelling with Reliable Uncertainty Quantification. In Shai Avidan, Gabriel Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner, editors, *Computer Vision – ECCV 2022*, pages 540–557, Cham, 2022. Springer Nature Switzerland.
- [78] Jianxiong Shen, Adria Ruiz, Antonio Agudo, and Francesc Moreno-Noguer. Stochastic Neural Radiance Fields: Quantifying Uncertainty in Implicit 3D Representations. In *2021 International Conference on 3D Vision (3DV)*, pages 972–981, 2021.
- [79] Ben Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, pages 336–343, 1996.
- [80] Niko Sünderhauf, Jad Abou-Chakra, and Dimity Miller. Density-aware NeRF Ensembles: Quantifying Predictive Uncertainty in Neural Radiance Fields. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9370–9376, 2023.
- [81] Mattias Teye, Hossein Azizpour, and Kevin Smith. Bayesian Uncertainty Estimation for Batch Normalized Deep Networks. In *ICML*, 2018.
- [82] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

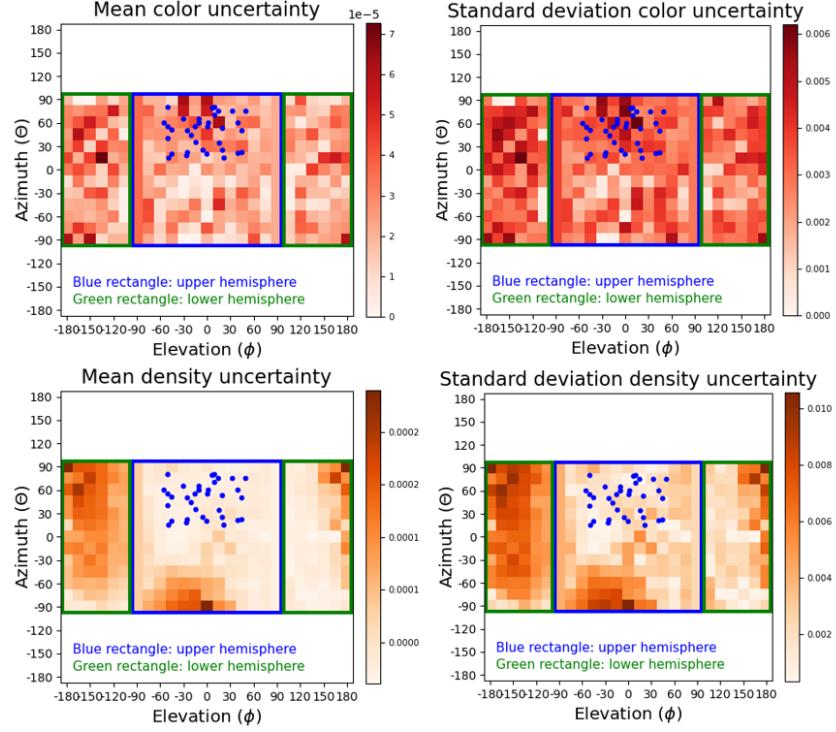
BIBLIOGRAPHY

- [83] Petr Šereda, Anna Vilanova, and Frans A. Gerritsen. Automating transfer function design for volume rendering using hierarchical clustering of material boundaries. In *Proceedings of the Eighth Joint Eurographics / IEEE VGTC Conference on Visualization*, EUROVIS'06, page 243–250, Goslar, DEU, 2006. Eurographics Association.
- [84] Daniel Weiskopf. Uncertainty Visualization: Concepts, Methods, and Applications in Biological Data Visualization. *Frontiers in Bioinformatics*, 2, 2022.
- [85] Hongbin Xu, Zhipeng Zhou, Yali Wang, Wenxiong Kang, Baigui Sun, Hao Li, and Yu Qiao. Digging into Uncertainty in Self-supervised Multi-view Stereo. In *ICCV*, 2021.
- [86] Steven Xu, HuiZhi Liang, and Timothy Baldwin. UNIMELB at SemEval-2016 Tasks 4A and 4B: An Ensemble of Neural Networks and a Word2Vec Based Model for Sentiment Classification. In Steven Bethard, Marine Carpuat, Daniel Cer, David Jurgens, Preslav Nakov, and Torsten Zesch, editors, *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 183–189, San Diego, California, June 2016. Association for Computational Linguistics.
- [87] Yang, Guo-Wei and Zhou, Wen-Yang and Peng, Hao-Yang and Liang, Dun and Mu, Tai-Jiang and Hu, Shi-Min. Recursive-NeRF: An Efficient and Dynamically Growing NeRF. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–14, 2022.
- [88] Huangying Zhan, Jiyang Zheng, Yi Xu, Ian Reid, and Hamid Rezatofighi. ActiveRMAP: Radiance Field for Active Mapping And Planning, 2022.
- [89] Tianjin Zhang, Zongrui Yi, Jinta Zheng, Dong Liu, Wai-Man Pang, Qiong Wang, and Jing Qin. A Clustering-Based Automatic Transfer Function Design for Volume Visualization. *Mathematical Problems in Engineering*, 2016:1–13, 01 2016.

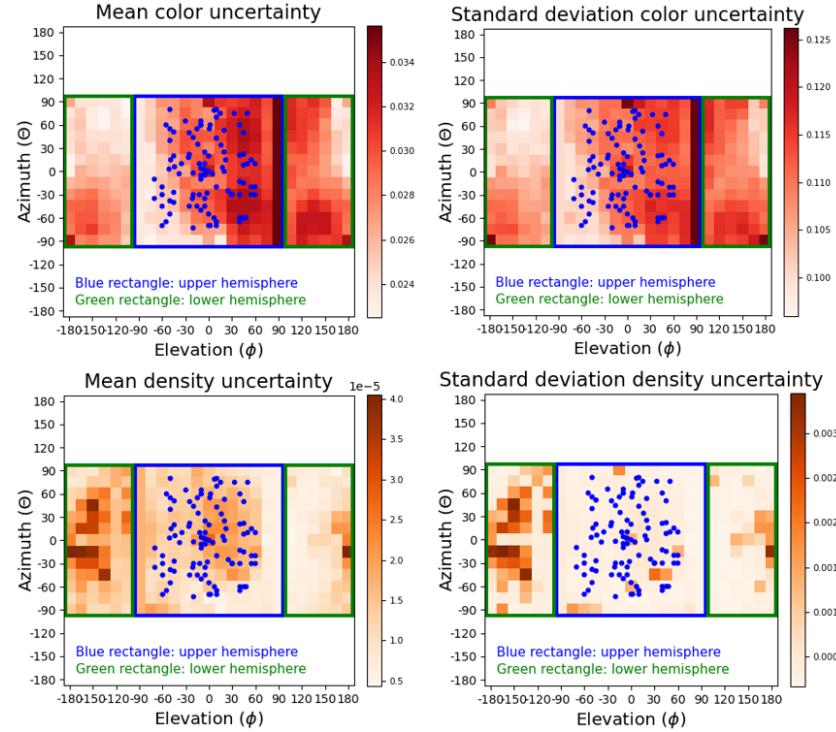
Appendix A

Experimental results

APPENDIX A. EXPERIMENTAL RESULTS

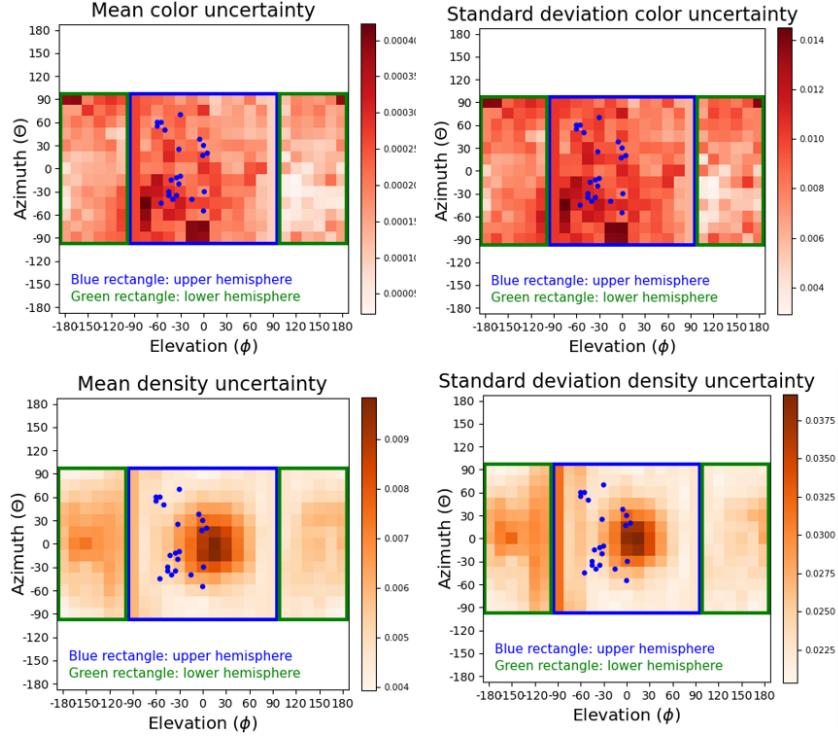


(a) Limited-angle dataset.

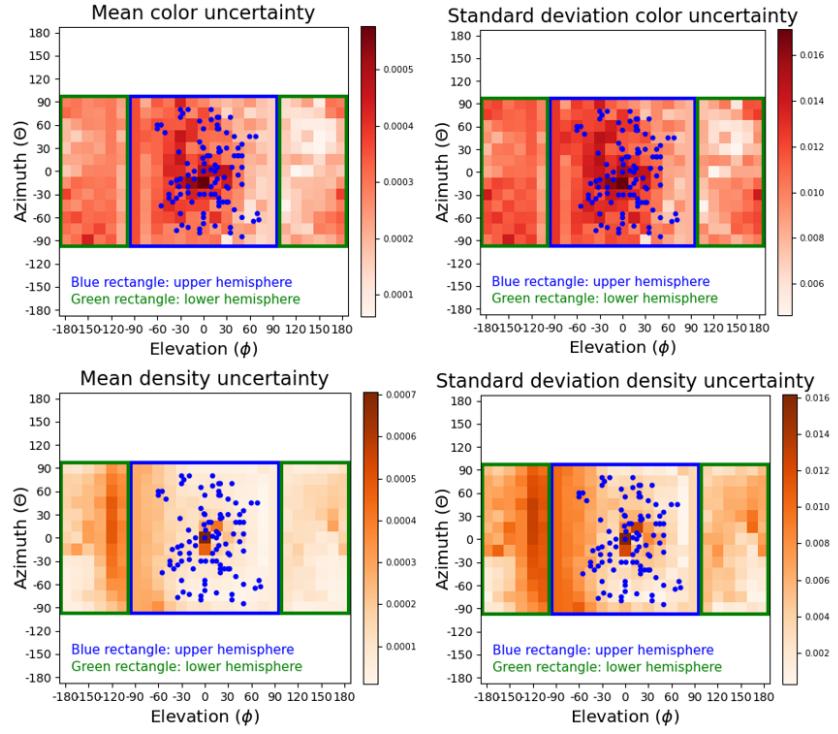


(b) Full-angle dataset.

Figure A.1: The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the microphone scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets.



(a) Limited-angle dataset.



(b) Full-angle dataset.

Figure A.2: The color and density mean and standard deviation uncertainty heatmaps in selected viewing directions for the drums scene trained with the Ensemble model using (a) limited-angle and (b) full-angle datasets.