A Framework for Inverse Scattering

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33840676">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33840676</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
A Framework for Inverse Scattering

A dissertation presented
by

Ioannis Athanasios Gkioulekas

to

Harvard John A. Paulson School of Engineering and Applied Sciences

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Engineering Sciences

Harvard University
Cambridge, Massachusetts
May 2016
Dissertation Advisor: 
Professor Todd Zickler

Author: 
Ioannis Athanasios Gkioulekas

A Framework for Inverse Scattering

Abstract

Scattering materials are ubiquitous: from our skin and food, to every day objects such as wax and soap, to industrial materials such as coatings and soft tissues. Common to all these materials is the complex way in which they interact with light. Their appearance is the result of photons that penetrate the material surface, and perform random walks inside the material before emerging towards a camera. Inverse scattering is, then, the problem of inverting this light transport process, in order to infer scattering parameters from images of a material.

We approach inverse scattering as an appearance matching problem: given a set of measurements (images) of a material, we search for the scattering parameters which, when used to computationally render new images, minimize the difference with the captured ones. In full generality, this is a very challenging optimization problem, due to the high-dimensional search space and the non-linear dependence of images on scattering parameters. We present several contributions for making this optimization problem tractable.

First, we present the results of a large-scale study of human perception of scattering material (translucent) appearance. Our study identifies a two-dimensional embedding of the physical scattering parameters in a perceptually-meaningful appearance space. Through our analysis of this space, we find uniform parameterizations of its two axes by analytical expressions of moments of the phase function, and provide an intuitive characterization of the visual effects that can be achieved at different parts of it. Our findings highlight the important role
phase function and mid-order scattering can have in controlling translucent appearance, motivating the development of inverse scattering algorithms that can handle these effects.

Second, we introduce a computational framework for efficiently solving inverse scattering appearance matching problems. Our framework is based on a combination of operator theory, stochastic gradient descent, Monte Carlo rendering, and material dictionary representations. It allows inverting the light transport process in a broad range of scattering materials, without having to rely on common approximations such as single scattering and diffusion. Additionally, it accommodates rich, high-dimensional material representations, enabling us to accurately measure parameters such as the scattering phase function shape, without having to rely on restrictive low-parameter models. To evaluate this framework experimentally, we create an acquisition setup that images thin material slabs under narrow-beam illumination from multiple lighting and viewing directions. Using measurements from this setup, we recover parameters of homogeneous (spatially-uniform) scattering materials, including arbitrary phase function shapes.

Third, we generalize our computational framework to address the heterogeneous inverse scattering problem, where the material parameters vary from point to point inside the volume. To this end, we make our algorithm applicable to measurements where photon contributions are decomposed based on criteria such as the distance they travelled inside the material, or their point of origin on the source illuminating the material. Additionally, we use path-space formulations of light transport, to allow our stochastic optimization framework to scale up to hundreds of thousands of scattering parameter unknowns. Through simulated experiments, we find that these extensions allow our algorithm to recover all spatially-varying scattering parameters for different types of scattering materials.

Fourth and finally, we present a computational imaging system that allows capturing the decomposed measurements used by our heterogeneous inverse scattering algorithms. Our system is based on interferometric techniques, and specifically on the optical coherence tomography framework. Our use of interferometry allows us to capture these decomposi-
ations at micron-scale resolutions, two to three orders of magnitude larger than previously possible. Such resolutions are necessary when collecting measurements for inverse scattering applications. We describe how to construct and optimize an optical assembly for this technique, and we build a prototype to measure and visualize scattering materials, and other optical phenomena.
## Contents

Abstract .............................................................. iii  
Acknowledgments .................................................... xxiv

1 **Introduction** .................................................. 1  
1.1 Fundamentals of Scattering ................................. 3  
1.2 Previous Approaches to Inverse Scattering .............. 4  
1.3 Outline and Contributions .................................. 7

2 **Motivation: Perception of Translucent Appearance** .... 9  
2.1 Background on Perception ................................. 12  
2.2 Phase Function Models ..................................... 13  
2.2.1 Henyey-Greenstein and von Mises-Fisher Lobes ..... 15  
2.2.2 Towards a Perceptual Space ............................ 16  
2.2.3 Characterizing Translucent Appearance .............. 17  
2.2.4 Sampling the Physical Parameter Space .............. 18  
2.2.5 Shape and Lighting ...................................... 19  
2.2.6 Design Choices .......................................... 21  
2.3 Computational Experiment ................................. 22  
2.3.1 Image-based Metrics .................................... 23  
2.3.2 Embeddings .............................................. 24  
2.4 Psychophysical Experiment ................................. 27  
2.4.1 Procedure ............................................... 27  
2.4.2 Stimuli .................................................. 27  
2.4.3 Implementation .......................................... 30  
2.4.4 Results .................................................. 30  
2.5 Analysis ....................................................... 36  
2.5.1 Parameterization of the 2D Perceptual Space ..... 37  
2.5.2 Phase Function Distance Metric ..................... 40  
2.6 Phase Function and Material Design ..................... 43  
2.6.1 Role in Translucent Appearance ..................... 43
<table>
<thead>
<tr>
<th>2.6.2 Phase Function Design</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7 Discussion</td>
<td>52</td>
</tr>
<tr>
<td>3 Radiative Transfer Background and Problem Setting</td>
<td>53</td>
</tr>
<tr>
<td>3.1 Notation and Setting</td>
<td>53</td>
</tr>
<tr>
<td>3.2 Radiative Transfer Equation</td>
<td>56</td>
</tr>
<tr>
<td>3.3 Volume Rendering Equation</td>
<td>59</td>
</tr>
<tr>
<td>3.4 Existence and Uniqueness</td>
<td>60</td>
</tr>
<tr>
<td>3.5 Imaging Scattering Media</td>
<td>61</td>
</tr>
<tr>
<td>3.6 Inverse Scattering by Appearance Matching</td>
<td>64</td>
</tr>
<tr>
<td>4 Homogeneous Inverse Scattering</td>
<td>66</td>
</tr>
<tr>
<td>4.1 Operator-theoretic Formulation of Volumetric Light Transport</td>
<td>67</td>
</tr>
<tr>
<td>4.1.1 Step-size Selection</td>
<td>71</td>
</tr>
<tr>
<td>4.2 Appearance Matching for Homogeneous Inverse Scattering</td>
<td>71</td>
</tr>
<tr>
<td>4.2.1 Dictionary Representation of Homogeneous Materials</td>
<td>72</td>
</tr>
<tr>
<td>4.2.2 Rendering Gradients of the Loss Function</td>
<td>74</td>
</tr>
<tr>
<td>4.2.3 Stochastic Optimization</td>
<td>77</td>
</tr>
<tr>
<td>4.2.4 Dictionary Design</td>
<td>79</td>
</tr>
<tr>
<td>4.3 Acquisition Setup Design</td>
<td>81</td>
</tr>
<tr>
<td>4.4 Experiments</td>
<td>83</td>
</tr>
<tr>
<td>4.4.1 Capture and Computation Time</td>
<td>84</td>
</tr>
<tr>
<td>4.4.2 Experiments with Synthetic Data</td>
<td>85</td>
</tr>
<tr>
<td>4.4.3 Validation Materials</td>
<td>89</td>
</tr>
<tr>
<td>4.4.4 Other Materials</td>
<td>91</td>
</tr>
<tr>
<td>4.4.5 Rendering of Measured Materials</td>
<td>94</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>94</td>
</tr>
<tr>
<td>5 Heterogeneous Inverse Scattering</td>
<td>97</td>
</tr>
<tr>
<td>5.1 Path-space Formulation of Volumetric Light Transport</td>
<td>98</td>
</tr>
<tr>
<td>5.2 Appearance Matching for Heterogeneous Inverse Scattering</td>
<td>100</td>
</tr>
<tr>
<td>5.2.1 Differentiation of the light transport matrix</td>
<td>100</td>
</tr>
<tr>
<td>5.2.2 Monte Carlo integration</td>
<td>101</td>
</tr>
<tr>
<td>5.2.3 Stochastic Optimization</td>
<td>102</td>
</tr>
<tr>
<td>5.3 Imaging Design</td>
<td>103</td>
</tr>
<tr>
<td>5.3.1 Local Ambiguities</td>
<td>104</td>
</tr>
<tr>
<td>5.3.2 Non-local Ambiguities</td>
<td>105</td>
</tr>
<tr>
<td>5.3.3 Noise Considerations</td>
<td>107</td>
</tr>
<tr>
<td>5.4 Initialization</td>
<td>108</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Parameterization of the vertical dimension of the embeddings of Figure 2.3
(“corr” refers to the absolute value of Pearson’s correlation coefficient). ...... 38
2.2 Absolute value of the Pearson’s correlation coefficient between functions of variance and the second moment of cosine, and the horizontal coordinate of the two-dimensional embeddings from Figure 2.3(a). ......................... 40
2.3 Parameterization of the horizontal dimension of the embeddings of Figure 2.3 ("corr" refers to the absolute value of Pearson’s correlation coefficient). .... 40
2.4 Mean value and 99% confidence interval for the image midpoint experiment described in Figure 2.11(a). In all cases, the mean response is close to the value 0.6364 predicted by our analysis. ................................. 51

4.1 Measurements of validation materials (controlled nano-dispersions). Values for $s$ are reported in (mm$^{-1}$). Phase function error is given as $L_2$ difference normalized by the $L_2$-norm of the reference phase function. All four validation materials have negligible absorption, resulting in both the predicted and measured values for $s_t$ to agree with those we report for $s_s$ to the third decimal. 90

4.2 Scattering parameters of materials measured using our proposed acquisition setup and inversion algorithm. Values for $s_s$ and $s_a$ are reported in (mm$^{-1}$). The average cosine of the measured phase functions is reported, while the entire phase functions are shown in Figure 4.9. Fitting and generalization errors are given as $\% L_2$ difference normalized by the $L_2$-norm of the reference image, averaged across the fitting and novel captured images respectively. 92

B.1 Pairwise Procrustes distances for the embeddings of Figure B.2. ............. 171
B.2 Pairwise Pearson’s correlation coefficient for the vertical coordinate of the embeddings of Figure B.2. ................................. 171
B.3 Pairwise Pearson’s correlation coefficient for the horizontal coordinate of the embeddings of Figure B.2. ................................. 171
B.4 Pairwise Procrustes distances for the embeddings of Figure B.3. ............. 172
B.5 Pairwise Pearson’s correlation coefficient for the vertical coordinate of the embeddings of Figure B.3. .............................................................. 173

B.6 Pairwise Pearson’s correlation coefficient for the horizontal coordinate of the embeddings of Figure B.3. .............................................................. 173

B.7 Parameterization of the vertical dimension of the full Procrustes mean of the embeddings of Figure 4 of the main paper. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.) .............................................. 182

B.8 Parameterization of the horizontal dimension of the full Procrustes mean of the embeddings of Figure 4 of the main paper. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.) .............................................. 182

F.1 List of major components used in the optical setup of Figure F.1. ............... 230
List of Figures

2.1 The role of phase function. (a) Completely different phase functions can result in very similar appearances, as shown by the two images (top), rendered with the phase functions shown in the polar plot (below). (b) Our analysis identifies a two-dimensional space for phase functions, where neighborhoods correspond to similar translucent appearance. (Numbers indicate the location in the two-dimensional space of the phase functions used to render the corresponding images in (a) and (c).) (c) We derive perceptually uniform axes that we can use to explore this two-dimensional space of translucent appearance. This enables us to design two-lobe phase functions to match the appearance of materials such as marble (left), but also others, such as white jade (right), that single lobe phase functions cannot match well (middle). . . 10

2.2 Phase functions for different materials. (a) Photographs under natural illumination of a “frog prince” model made of two different translucent materials, soap and wax. The middle two rows (b) and (c) show images, and rendered fits, of thick, homogeneous slabs made from these materials, under laser illumination. (b): soap, (c): wax. From left to right: photographs, best fit simulations using single HG lobes or their mixtures, best-fit simulations using single vMF lobes or their mixtures. HG lobes cannot reproduce the “apple” shaped scattering patterns observed in wax and other materials, whereas vMF can. (d) compares HG (blue) and vMF (red) phase functions for increasing values of average cosine, shown as increase in saturation. vMF lobes are wider and not as strongly forward-scattering as HG lobes. . . . . . . . . . 15
2.3 Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the cubic root metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them (see Section 2.5.1). To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

2.4 Computational analysis results. (a) Comparison of two-dimensional embeddings produced using different image-based metrics, for three representative scenes. Points on embeddings are colored according to the square of the average cosine of the corresponding phase function. (b) Comparison of consistency of two-dimensional embeddings across the nine scenes of Figure 2.3, for each of the three image metrics. (c)-(e) Measures of two-dimensional embedding similarity for each pair of the nine scenes, when the cubic root metric is used to produce the embeddings.

2.5 Setup used for psychophysical experiment. (a) Rendered Lucy images used as stimuli for the sidelighting (left) and backlighting (right) experiments, with the face masked out. (b) Screen capture of the interface used in the experiment.

2.6 Comparison of computational and perceptual embeddings for the sidelighting experiment. (a) Locations of the stimulus images selected by the clustering algorithm (overlaid on the two-dimensional computational embedding) are well-distributed. (b) Two-dimensional embedding learned from the data collected from the psychophysical experiment. The inset table shows values for Pearson’s correlation coefficient between corresponding coordinates of the cluster centers on two embeddings. (c) A subset of the stimulus images, arranged in a grid according to their positions in the computational embedding. The same grid is also consistent with the positions of the images on the perceptual embedding. Points on embeddings and images on the grid are colored according to the square of the average cosine of the corresponding phase function, and consistently labeled across (a)-(c).
2.7 Comparison of computational and perceptual embeddings for the backlighting experiment. (a) Locations of the stimulus images selected by the clustering algorithm, overlaid on the two-dimensional computational embedding. (b) Two-dimensional embedding learned from the data collected from the psychophysical experiment. The inset table shows values for Pearson’s correlation coefficient between corresponding coordinates of the cluster centers on the two embeddings. (c) A subset of the stimulus images, arranged in a grid according to their positions in the computational embedding. The same grid is also consistent with the positions of the images on the perceptual embedding. Points on embeddings and images on the grid are colored according to the square of the average cosine of the corresponding phase function, and consistently labeled across (a)-(c).

2.8 Parameterization of the computational two-dimensional embedding. (a) Embedded points are colored using the square of the average cosine of the corresponding phase function, which parameterizes the vertical dimension of the embedding. (b) Embedded points are colored using function (2.11), which parameterizes well the horizontal dimension.

2.9 Results of metric learning for phase functions. (a) compares the polar plots of the diagonal of the learned weight function (shown in blue, smoothed to 15 first Legendre polynomial coefficients for visualization), and the uniform one, corresponding to the $L_2$-distance (shown in red). (b) and (c) show the embeddings of phase functions using the $L_2$-distance and the learned metric, with points colored according to the square of the average cosine of the corresponding phase function (compare with Figure 2.8(a)).

2.10 Locations of points corresponding to different types of phase functions on the two-dimensional computational embedding. Embedded points corresponding to the specific type of phase function are shown colored according to the squared average cosine of the phase function, whereas the rest of the embedding is faded out. Top row: single lobe and single lobe + isotropic combinations are limited to the left part of the embedding. Bottom row: two lobe combinations achieve appearance towards the right in the embedding.
2.11 Perceptually uniform parameterization of a single-lobe HG. (a) Graphical user interface used to validate perceptually uniform interpolation. (b)-(c) Both rows: left image ($g = 0$), right image ($g = 0.9$). Middle image rendered using $g$ corresponding to the midpoint between these values: by linear sampling ($g = 0.45$, top row), or quadratic sampling according to $g^2$ ($g = 0.6364$, bottom row). By looking at each row from the left to the right, we observe that the appearance of the middle image in the bottom row is perceptually a better “middle point” than in the top row.

3.1 (a) Photons traveling inside scattering materials perform random walks that depend on the material parameters. (b) Computational imaging techniques capture measurements of such materials by collecting different subsets of photons, depending on the paths they follow. (c) Inverse scattering is the problem of inferring, potentially spatially-varying, scattering material parameters from such measurements.

4.1 2D fluence fields computed using the finite-difference approximation and the standard volume rendering equation: (top) uniform phase function $f_p(\theta) \equiv \frac{1}{4\pi}$; (bottom) von Mises-Fisher phase function with $\kappa = 5$.

4.2 Gradient computation (Equation (4.28)) for the dictionary-based appearance matching optimization problem (Equation (4.25)). For each pair of lighting and viewing directions, the gradient with respect to the weight of the $n$-th dictionary material requires computing a cascade of three rendering operations: a full volume rendering using the mixture material; then a single-step propagation using the corresponding atom material; then another full volume rendering using the mixture material. The resulting light fields are then sampled to produce the images used algebraically in Equation (4.28).

4.3 Phase functions in a tent dictionary with $N = 10$ atoms (each atom is colored uniquely for better visualization). To be valid probability distributions on the sphere, the atoms are normalized to satisfy Equation 3.19, and thus have varying magnitudes. The atoms centered at $0^\circ$ and $\pi^\circ$ are shown cropped.

4.4 Setup for scattering parameter acquisition. A sample cell is illuminated by collimated beams and imaged by a camera. Rotation stages to achieve arbitrary combinations of lighting and viewing directions. Left: schematic; Right: implementation.
4.5 The error surface for 2D optimization problems. We consider a dictionary of three phase functions whose mixing weights lie on the simplex. The simplex is parameterized by its first two coordinates. Columns show results for three reference \( \sigma_t \), and the rows show two different points as the correct reference in this space. ......................................................... 85

4.6 Accuracy in the recovery of material parameters. The plots show recovery errors for albedo, extinction coefficient, and mean error for phase function, as a function of different values of albedo \( a = \sigma_s / \sigma_t \) and mean free path \( d = 1 / \sigma_t \). 86

4.7 Measured materials in glass cells of width \( w = 1, 10, \) and \( 2.5 \) mm, from left to right. It is not necessary for all of the cell to be filled, as long as there exists a homogeneous region of size comparable to the beam diameter (e.g., hand cream). ......................................................... 88

4.8 Measurements of validation materials (controlled nano-dispersions). Left: For each material, we show for the green wavelength the theoretically predicted (dashed orange) and recovered (purple) phase functions, as well as the best Henyey-Greenstein (green) phase function fit. The recovered phase functions are in close agreement with the correct ones and the purple and orange curves tightly overlap. As another visualization, we show the images for a novel configuration: under frontal collimated laser illumination \( (\theta_f = 25^\circ, \theta_o = 0^\circ) \). We compare our phase function and the best Henyey-Greenstein fit (images are color-mapped for better visualization). The rightmost column shows a cross-section through the captured and re-rendered images for this configuration. 91

4.9 Phase functions of materials measured using our acquisition setup (purple), contrasted with the closest HG phase function (dashed green). See Table 4.2 for numerical values. ......................................................... 93

4.10 Top: Rendering of materials in natural illumination using our acquired material parameter values. Bottom: Same image with RGB vectors normalized to unit norm, to visualize color variations. ......................................................... 95

5.1 Inverse rendering algorithm. Left: We use the ADADELTA variant of stochastic gradient descent to minimize the appearance matching objective of Equation (5.8). Right: We use a modified Monte Carlo rendering algorithm to compute stochastic gradient estimates. When shading a path, we compute for each segment both the usual throughput terms \( f_m \) and the score terms \( S_{m,k} \). These are accumulated to determine the path’s contribution to the image \( I \) and its derivative \( G_k \) (Equations (5.4) and (5.11)). We repeat this process over many paths that are sampled to satisfy the sampling function \( W \) corresponding to the input image measurement (Equation (5.3)). .............. 104
5.2 (a) Configuration used for Claim 4. (b) Frontlighting allows decomposing the inverse scattering problem into sub-problems for depth-wise layers, where at each layer the material parameters of lower-depth layers are known. (c) Backlighting allows measuring deep layers at shorter pathlengths $\tau$, and thus higher SNR, than frontlighting. .......................... 106

5.3 Comparison of different imaging configurations. (a) Visualization of ground-truth and reconstructed material parameters as a function of location. (b) Mean relative reconstruction error for material parameters, averaged over multiple synthetic volumes. .......................... 112

5.4 (a) Renderings of a smoke volume using ground-truth and recovered material parameters under novel viewpoint and illumination conditions. (b) Ground-truth and recovered parameters for a vertical cross-section through the smoke volume. .......................... 113

6.1 Differentiating light paths. A pixel in a conventional image measures the sum of all light paths arriving at that pixel. Entries of a light transport matrix distinguish paths with different starting locations (orange vs. blue), and entries of a “pathlength-resolved” light transport matrix additionally distinguish paths of different lengths (e.g., light vs. dark blue). ......... 118

6.2 Michelson interferometer. An input beam is split by a beamsplitter into two copies that traverse to mirrors $M_r$ and $M_s$ at distances $d_r$ and $d_s$ from the splitter, respectively. The two copies reflect back and recombine at the beamsplitter before being imaged. Insets (i) and (ii) show images for two positions of mirror $M_r$ that either induce interference ($d'_r \approx d_s$) or do not ($d'_r \neq d_s$). Inset (iii) shows a setup that can be used to achieve oblique illumination. .......................... 122

6.3 Notation and coordinate system used in Sections 6.2 and 6.3. (a) The parts of the Michelson interferometer corresponding to the scene and reference arms. (b-e) The scene and reference arms shown unfolded in the same coordinate system. (b) and (c) show the target arms, with lighting and camera respectively. (d) and (e) show the reference arms, with lighting and camera respectively. .......................... 124
6.4 Measuring the temporal coherence length. The graph to the left shows the intensity measured by the camera in Figure 6.2 for different values of path-length difference \( \tau = |d_s - d_r| \), and the graph to the right shows a close-up over a range of 2 \( \mu \)m. When the path-length difference is zero, maximal constructive interference exists. A path-length shift of \( \lambda/2 \) results in destructive interference (zero intensity). As the path-length difference increases, correlation between the waves is reduced and the fringe contrast decays, becoming almost zero once \( \tau \) exceeds the temporal coherence length. The differently colored plots show measurements for different spectral bandwidths. Using a white light source with color filters of spectral bandwidth of 3, 10 and 25 nm produces temporal coherence lengths of 50, 25, and 10 \( \mu \)m, respectively. The measurements for this figure were taken at a resolution of 10 nm.

6.5 By equipping the Michelson interferometer with sources having different coherence properties, we can capture different light path decompositions. (a) Pathlength decomposition separates all paths in a scene \( (M = I) \), regardless of their endpoint locations, in terms of their optical length \( \tau \). It separates direct paths (blue), reflections (purple), retroreflections (orange), and all scattering (greens). (b) Spatial probing separates paths with certain start point locations while ignoring their lengths (by effectively measuring summations over pathlength). For a diagonal \( M \), it separates direct, retroreflection, and backscattering (dark green) paths. (c) Finally, it is also possible to combine the previous two cases and separate paths in terms of both pathlength and endpoint locations.

6.6 Reference mirror configurations in a Michelson interferometer induce different spatial probing patterns. Assuming co-axial source and camera: (a) A planar mirror probes the main diagonal of the light transport matrix, corresponding to direct, retroreflection, and backscattering paths. (b) A right-angle mirror pair probes the anti-diagonal, which includes \( 2n \)-bounce reflection paths. (c) Offsetting a right-angle pair allows probing away from the anti-diagonal. (d) Different combinations of these basic configurations probe blocks of the light transport matrix in different ways.
6.7 Capture and computational pipeline. We consider a toy scene consisting of a
diffuse plane, tilted relative to the optical axis of the camera and illumination,
as shown in the upper left. By translating the reference mirror to different
positions, we capture a sequence of frames corresponding to the intensity
\( I_{i,\tau}(x) \) measured for different values \( \tau \). By processing the frames using Equation (6.25), we separate the interference component \( \hat{T}_m(f(x), x, \tau) \), resulting
in the images at the bottom. As the target plane is tilted, different points
have different depths and therefore interfere with the target arm at different
pathlength values \( \tau \), corresponding to the scan lines shown in the processed
frames.

6.8 Optimizing the Michelson interferometer of Figure 6.2. The schematic to the
left describes our implementation. We consider the effect of the following
imaging parameters: source aperture area \( A_s \), camera aperture area \( A_c \), and
pixel pitch \( p \). The optimized settings correspond to camera F-number of \( f/32 \),
source aperture so that the source’s angular extent is approximately half that
of the camera, and pixel pitch of 4 μm. Different configurations are produced
by modifying only one parameter. For each configuration, the middle graph
shows the measured contrast for the tilted diffuse plane scene of Figure 6.7;
and the right graph compares contrast and total exposure time.

6.9 Scanning depth. For each scene from top to bottom: an unprocessed HDR
frame; scanned depth, visualized using pseudo-color and a periodic change
in luminance for a contouring effect; and rendering of scanned 3D mesh. All
scans are at a depth/pathlength resolution of 10 μm. Scenes from left to right:
US quarter, gnocchi, soap carving, plastic toy cup, dry pasta, and gummy bear.

6.10 Depth scans of complex scenes. From left to right: an unprocessed HDR
frame; scanned depth, visualized as in Figure 6.9; and rendering of 3D mesh.
Top row: Chess knight between a diffuse wall to the right and a mirror to the
left. The mirror-reflection is interpreted as a real object at a greater depth.
Bottom row: Gummy bear between two diffuse walls. The scene geometry is
recovered accurately, despite the conflation of paths from the diffuse walls
behind the near-transparent gummy bear.

6.11 Direct-global separation for a close-up of a strawberry.
6.12 Scanning a scene with materials of different scattering properties using an LED source. (a) A photograph of the scene: A turbine blade with a highly diffusive zirconia coating is placed in front of a spectrolon slab. (b) An unprocessed frame. (c) A processed frame, showing as an image the value of the measurement $T^m((x, y), (x, y), \tau_0)$ for some value of $\tau_0$ and all pixels $(x, y)$. (d) The function $T^m((x, y_0), (x, y_0), \tau)$ for pixels $(x, y_0)$ on the horizontal crosssection shown in (b), and for all values $\tau$.

6.13 Pathlength decomposition of scattering paths starting and ending at a single point on the surface of a scattering volume. All measurements use an LED source. Left: Measured scattering profiles $T^m((x_o, y_0), (x_o, y_0), \tau)$ at a single pixel $(x_o, y_0)$ as a function of $\tau$, for various materials. Right: Measured profiles for the parallel and cross-polarized scattering component of zirconia.

6.14 Visualizing dispersion in the pathlength domain. (a) A schematic of the scene, consisting of a glass container partially filled with water and placed in front of a mirror. As light travels through the glass and water, different wavelengths are delayed by different amounts due to dispersion. (b) A photograph of the scene we scan with an LED source. We consider the close-up shown in the box. (c) Frames acquired at a wavelength of 525 nm, and at three different pathlengths. Spatial fringes indicates interference of paths that travel through, from left to right, air only, air and glass, and air, glass, and water. (d) Visualization of the interference measured across a vertical cross-section of the frames, as a function of optical pathlength and for three different wavelengths. We observe that paths travelling through glass or through both glass and water are delayed more compared to paths travelling through air only. Additionally, the delay is a function of wavelength, resulting in the appearance of a “rainbow” as a function of optical pathlength.

6.15 Visualizing photoelasticity. (a) We use an LED source to scan a scene consisting of a plastic protractor in front of a mirror. We consider the close-up shown in the box. (b) A standard RGB frame under polarized light with no interference or scanning. The color patterns on the protractor correspond to wavelength and polarization-dependent variations of refractive index due to photoelasticity. (c) Low-resolution (1 mm) measurement of optical pathlength at every pixel, shown color-coded. (d) Zooming on the protractor, top row shows (left to right) the red, green and blue channels from (b); bottom row shows optical pathlength measurements for the same three channels and at a 10 $\mu$m pathlength resolution.
6.16 Probing different light transport matrix components. (a) We use an LED source to scan at three wavelengths a scene consisting of a thin glass slab placed between a mirror and a diffuser. A frame is shown at the bottom-right. The scene includes three types of paths: direct paths (blue), mirror-diffuser-mirror retroreflection paths (orange), and two-bounce reflection paths (purple). (b) Measurements of the primary diagonal of the light transport matrix, corresponding to direct and retroreflection paths. The left visualizes the measurements \( T_m((x, y_0), (x, y_0), \tau) \) over pathlength \( \tau \) for a horizontal line of pixels (a pixel-pathlength slice). The right shows image frames \( T_m((x, y), (x, y), \tau_0) \) for all pixels \((x, y)\) at two pathlength values. (c) Measurements of the anti-diagonal of the light transport matrix, corresponding to paths with start and end locations symmetric around the glass slab. These are two-bounce reflection paths that, for each wavelength, all have the same optical length. This produces horizontal stripes in the pixel-pathlength slice \( T_m((x, y_0), (-x, y_0), \tau) \) (left), and a short sequence of all-bright frames \( T_m((x, y), (-x, y), \tau_0) \) in the transient video (right, shown for the green wavelength). The rainbow patterns in (b) and (c) correspond to optical pathlength differences of 15 – 30 \( \mu \)m between different wavelengths, due to dispersion in the glass slab.

6.17 Capturing different pathlength decompositions. (a) We scan a scene consisting of a plastic toy cup, and show pixel-pathlength slices for a vertical cross-section of pixels through the visual field. (b) When we capture a pathlength decomposition of the entire light transport matrix for this scene using an SLD, the pixel-pathlength slice includes direct paths, scattering, caustics, and specular interreflections. These are also apparent in the 2D image frames of the transient sequence (orange and blue insets). (c) When we measure the pathlength decomposition only of the primary diagonal of the light transport matrix using an LED, the pixel-pathlength slice has only direct and backscattering paths.

6.18 Light-in-flight visualizations for two different scenes using an SLD source. For each row, the leftmost column shows a conventional photo of the scene, and columns two to four show representative frames. From top to bottom, the scenes are: chess knight between diffuse wall and mirror; zirconia layer on ground glass plate between diffuse walls.

B.1 Image of the scene “Lucy + Campus” rendered with phase functions of average cosine value \( \bar{\mathcal{C}} > 0.8 \).
B.2 Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the $L_1$-norm image metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

B.3 Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the $L_2$-norm image metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

B.4 Two-dimensional embeddings of images rendered with the 40 phase functions selected by the clustering algorithm, produced using the mean-opinion-score image metric. The embedded dots, one per rendered image for a total of 40 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

B.5 Comparison of two-dimensional embeddings produced by the cubic root and mean-opinion-score metrics, for images rendered with the 40 phase functions selected by the clustering algorithm. The embedded dots, one per rendered image for a total of 40 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them.

B.6 Member images of representative cluster 1.

B.7 Member images of representative cluster 2.

B.8 Member images of representative cluster 3.

B.9 Visualization of computational embedding for sidelighting experiment.

B.10 Visualization of perceptual embedding for sidelighting experiment.

B.11 Visualization of computational embedding for backlighting experiment.
B.12 Visualization of perceptual embedding for backlighting experiment. . . . . . 181
B.13 Bootstrapping analysis: comparison of (a) embedding learned from data provided from the subjects; and (b)-(f) embeddings learned from five random perturbations of the data. ......................................................... 181
B.14 Comparison between (a) the two-dimensional embedding for the scene “Lucy + campus”, produced using the cubic root metric; and (b) the full Procrustes mean of all the embeddings of Figure 4 of the main paper, as produced by generalized Procrustes analysis. .............................................. 182
B.15 Two-dimensional learned weight function for comparing phase functions, visualized as a matrix. Zooming in shows the high values of the diagonal, which is plotted separately in Figure 10(a) of the main paper. ............... 183

D.1 Light generation and control. Top: schematic; bottom: implementation. . 206

E.1 Schematic for proof of Claim 12 .................................................. 212

F.1 Schematic (top) and physical implementation (bottom) of our optical setup. 223
Acknowledgments

I would first like to thank Todd Zickler for the incredible mentorship he provided as my advisor. His intuition and instinct for identifying important problems; methodical approach to coming up with solutions; and unique ability to communicate ideas are all qualities I will aspire to emulate in my own career. I am indebted to him for his continued support and guidance throughout my seven years at Harvard.

I have been fortunate to have Kavita Bala and Anat Levin act as informal second advisors during the largest part of my doctorate studies. Kavita introduced me to a whole new research community, and her enthusiasm and clear view of the bigger picture were instrumental for turning rough ideas into successful projects. I spent many hours working with Anat, including late night hours in our dark room, and greatly benefited from her unique command of both theory and practice in several areas of computer vision and graphics. I have also had the privilege to collaborate with Edward Adelson and Frédo Durand. They both exposed me to diverse ideas and invaluable insights, on subjects ranging from all aspects of photography, to manufacturing interesting materials in one’s basement and using them to motivate research projects spanning several years. I am grateful to Steven Gortler, Na Li, Yue Lu, and Hanspeter Pfister, for inspiring discussions, and for the time they donated serving on my examination committee.

This work is the product of engaging and fruitful collaborations with Sanjeev Koppal, Bruce Walter, Bei Xiao, and Shuang Zhao, and this dissertation would not have been possible without them. I want to acknowledge my fellow members of Harvard’s vision and graphics groups, for the amazing work and social environment they created and shared with me: Emma Alexander, Moritz Baecher, Gaurav Bharaj, Guillermo Canas, Ayan Chakrabarti, Kevin Dale, Daniel Glasner, Qi Guo, Milos Hasan, Ritwik Kumar, Ruonan Li, Fabiano Romeiro, Ely Spears, Zak Stone, Kalyan Sunkavalli, Yuriy Vasilyev, Jialiang Wang, Ying Xiong, and Yuanchen Zhu. Special thanks go to Kathleen Masse, for putting up with my
never-ending administrative requests.

Finally, I am deeply obliged to my family; for their willingness to provide me with the means necessary to pursue my current career; and for their unconditional support throughout my life. I hope they will at least get to see a graduation ceremony this time.
Chapter 1

Introduction

Light travels in many ways through a scene. Photons emitted from a source follow multiple paths and bounce several times on the surfaces and interior of objects, based on their shape and material properties. We experience this process by using sensors, such as cameras and our own eyes, to capture some of these photons and convert their energy to measurements in the form of images. At a very high level, we can express this process using the following image formation equation,

\[
\text{image} = S_{\text{sensor}} \left( R_{\text{light transport}} \left( \text{geometry, material, illumination} \right) \right).
\]

Computer graphics and computer vision study two types of problems relating to the image formation equation. The first type are forward problems: Given a specification of the geometry, material, and illumination of a scene, can we infer the image that a real sensor would measure? This involves evaluating operators \( R_{\text{light transport}} \) and \( S_{\text{sensor}} \), a process referred to as rendering.

The second type is inverse problems: Given one or more images captured by physical sensors, can we infer the shape, material, and illumination that created it? This involves inverting \( R_{\text{light transport}} \) and \( S_{\text{sensor}} \), a process referred to as inverse rendering. The contributions of this
dissertation fall within the scope of this second category.

The complexity of light propagation makes general inverse rendering ill-posed. This forces us to consider solving this problem in more constrained settings, where we assume that we know some of the variables in the image formation process, and we seek to infer the rest. Depending on the assumptions we make, we can formulate three inverse rendering problems, as described in [96].

The first problem is inverse geometry rendering. Here, we assume that we know the material of all objects in a scene, and the illumination. Then, we are looking to recover the shape of all objects in the scene, as well as their relative positions. This category subsumes the various shape from X problems formulated in computer vision, where X can be shading, texture, defocus, multiple viewpoints, structured lighting, and so on.

The second problem is inverse lighting. In this category, we assume that we know the geometry of the scene and the material of all objects in it, and we seek to recover the incident illumination. This can be distant illumination, such as environment maps, or the shape and emittance of local sources, such as a lamp or a projector.

The third problem is, predictably, inverse material rendering, where we know illumination and geometry, and we are looking to recover the optical material properties of the objects in the scene. There are three types of such properties, corresponding to three further inverse rendering sub-problems:

- The reflectance and transmittance at the interface between different materials. These are described by the bidirectional scattering distribution function (BSDF), or the bidirectional reflectance distribution function (BRDF) for materials that are completely opaque. The inverse rendering problem of recovering the BRDF or BSDF is known as reflectometry.

- The index of refraction controlling the speed of light propagation at the interior of
materials. When the index of refraction is constant, this inverse rendering problem is referred to as refractometry; otherwise, it is known as inverse eikonal rendering.

- The scattering parameters inside materials that are translucent. This problem is referred to as inverse scattering or inverse volume rendering.

Within this taxonomy of inverse rendering problems, this dissertation focuses on inverse scattering.

1.1 Fundamentals of Scattering

Scattering occurs in translucent materials. These materials are ubiquitous, including: most biological tissues, minerals and gemstones, food, everyday chemicals like waxes and soaps, and industrial materials such as plastics and coatings. The term scattering refers to the common way in which they interact with light: photons penetrate the material surface and interact with material internal structures. As a result of these interactions, they perform random walks inside the material volume, before emerging back outside.

This behavior can be modelled using the radiative transfer equation, which describes the propagation of light inside materials using three wavelength-dependent parameters. The first is the mean free path, which determines the average distance photons travel between interactions with the material. The second is the volumetric albedo, which determines the probability that, at every interaction, a photon will be absorbed. The third parameter is the phase function, which is a spherical probability distribution determining the direction at which a photon travels after an interaction. Based on the spatial characteristics of these three parameters, we can classify scattering materials into two types: homogeneous materials, where the parameters are constant everywhere inside their volume; and heterogeneous materials, where the parameters change from point to point. We describe these parameters and radiative transfer in detail in Chapter 3.
1.2 Previous Approaches to Inverse Scattering

Because of the ubiquity of translucent materials, inverse scattering has been studied for decades not only in computer graphics and vision, but also in areas such as applied physics, remote sensing, biomedical imaging, and material sciences. Despite this, there is still no general solution to this inverse problem. The two primary challenges in solving inverse scattering are:

- The extremely multi-path nature of light propagation inside translucent materials. As a camera receives contributions from photons that have travelled along an infinite number of paths inside the material, the information available in the measurements about the material is greatly confounded.

- The very high-dimensionality of the parameter space. In homogeneous materials, this high-dimensionality is due to the phase function, which is a potentially infinite-dimensional probability distribution. In heterogeneous materials, this is exacerbated by the need to estimate different parameters at different points in the material.

To address the first problem, existing inverse scattering algorithms generally take one of three approaches.

Deep scattering and approximation by diffusion. Inverse problems in this category consider media that are optically-thick, so that photons scatter many times before being measured. Radiative transport is then modeled using the diffusion approximation, where the angular variation of the internal radiance is limited and the radiative transport equation reduces to a diffusion equation [67]. The advantage of this approach is that it simplifies the inference problem, allowing efficient acquisition and rendering systems [75, 29] and, as demonstrated by Wang et al. [145], the estimation of spatially-varying structure within a medium. The diffusion approximation is also employed in diffusing-wave spectroscopy [119], which is used for applications such as particle sizing or the measurement
of molecular weight; and diffuse optical tomography [13], which is used for imaging deep
tissues in medical applications. The diffusion approximation applies when high-order
scattering is dominant, causing the phase function to be confounded with other scattering
parameters [150, 156]. These ambiguities imply that parameters recovered using diffusion
do not generalize to other geometries that are less optically thick.

**Single scattering.** At the other extreme are optically-thin situations, where photons scatter
only once before being measured. Scattering parameters can often be measured directly in
these cases, using techniques like static or dynamic light scattering [76]. For graphics, Hullin
et al. [64] use fluorescent dyes to make qualitative observations on the scattering parameters
of optically thin media exhibiting mostly single scattering. Hawkins et al. [57] use a laser
to measure albedo and a tabulated phase function of a sparse homogeneous aerosol. For
liquids, Narasimhan et al. [104] successively dilute samples with water until they are sparse
enough to infer from single-scattering a one-parameter phase function. Single-scattering has
also been exploited to capture time-varying and spatially-varying wisps of smoke and sparse
mixing liquids [57, 43, 50]. All of these techniques rely on manipulating materials so that
single-scattering dominates, and while dilution can be used for aerosols and some liquids,
it cannot be easily applied to solids, or to liquids whose dispersing medium is unknown
and significantly different from water. This limitation motivates methods for suppressing
multiple scattering without dilution. Techniques for particle sizing or molecular weight,
for example, exploit cross-correlation properties of multiple temporal measurements [120],
but these are specific to those applications and do not easily extend to our problem. For
graphics, [102] use high frequency lighting patterns to isolate single-scattering effects [105],
allowing direct access to the mean free path and a good initialization for an indirect (multi-
scattering) optimization of an one-parameter phase function. Such lighting-based isolations
of single scattering are potentially quite useful, but as discussed in the context of 3D surface
reconstruction [62, 51], they provide only approximate isolation, and there is currently no
analysis of how this affects the accuracy of inferred scattering parameters.
**Surface representations of scattering.** A third class of techniques seek to recover surface-based descriptions of light transport through translucent objects [27, 45, 138, 117, 30]. These provide mappings between the input and output light on a specific object’s surface, and they do so without explicitly estimating all of the internal scattering parameters. Because of this, the material parameters recovered only apply to shapes of fixed shape, and they cannot be used to predict the appearance of an object of different shape made of the same material.

**Phase function models.** To address the problem of dimensionality, most existing approaches to inverse scattering either assume that the phase function is known, or they only seek to recover low-dimensional approximations to its shape. These approximations are typically based on the Henyey-Greenstein model [61], which reduces the infinitely-dimensional phase function space to a single scalar that can be used to continuously change from predominantly forward, to angularly uniform, and to predominantly backward scattering. Some extensions include low-dimensional models (two-three parameters) that use linear combinations of Henyey-Greenstein phase functions [125]. These parametric models can only be accurate for materials they represent well, and they will result in low-fidelity measurements when they cannot approximate well the real phase function shape.

Another common simplifying assumption is that the phase function is isotropic, meaning that it is invariant to rotations of the incident direction and cylindrically symmetric. While this assumption will also be used throughout this thesis, we must mention that it is inaccurate for materials that exhibit anisotropic internal structure, such as fibers in cloth. Accurately representing such materials requires using anisotropic phase function models [155, 70].
1.3 Outline and Contributions

From the overview of existing inverse scattering techniques presented above, we can identify two limitations shared by all algorithms: The first is the inability to perform inverse scattering in situations where mid-order scattering is significant. This corresponds to photons that scatter more than once, but not enough times to enter the diffusive regime. The second is the lack of techniques for recovering phase functions of arbitrary shape.

In this dissertation, we introduce a framework for inverse scattering that overcomes both of these limitations. Our contributions are as follows:

**Human perception of translucency.** To motivate the need for our inverse scattering framework, we perform a large-scale study of the effect different scattering parameters and phenomena have on how humans perceive translucent materials. Our study highlights that the shape of the phase function and mid-order scattering can strongly affect perceived appearance; and that commonly used low-dimensional parameterizations for the phase function are inadequate for many common materials.

**Homogeneous inverse scattering.** We introduce an algorithm for solving the inverse scattering problem for homogeneous materials, when provided with calibrated image measurements. Our algorithm operates by using Monte-Carlo rendering to compute stochastic estimates of the gradients of images with respect to material parameters; and combining these with stochastic optimization to solve an appearance matching problem for inferring the parameters. These features allow our algorithm to take into account all orders of scattering, single, mid, and high, in a physically accurate way, which enable its use for recovering phase functions of arbitrary shape and acquiring very broad classes of materials.

**Heterogeneous inverse scattering.** We generalize our inverse scattering algorithm in two ways: First, we make it applicable to measurements from more general types of imaging,
including transient imaging and other types of light transport decomposition. Second, we use a path-space formulation of light transport to make Monte-Carlo rendering of image gradients with respect to material parameters more efficient. These two generalizations allow our algorithm to scale up to the heterogeneous inverse scattering problem, enabling us to recover spatially-varying mean free path, albedo, and moments of the phase function.

**High-resolution light transport decompositions.** We introduce a computational imaging system that can capture the micron-resolution light transport decomposition measurements employed by our heterogeneous inverse scattering algorithm. Our system operates by using interferometry, and can additionally be used to produce measurements for shape recovery, light-in-flight visualizations, and inferring material parameters beyond scattering.

Each contribution is presented in its own chapter, in the order that they are listed above. In addition, Chapter 3 covers the mathematical foundations of scattering and radiative transfer, which are used in all subsequent chapters. Finally, each chapter has a corresponding appendix at the end of the thesis, which include additional results and proofs of formal claims we make throughout the thesis.
Chapter 2

Motivation: Perception of Translucent Appearance

To motivate the need for inverse scattering systems that can accurately handle mid-order scattering and arbitrarily shaped phase functions, in this chapter, we study the effect these two aspects of scattering have in the way human observers perceive translucency.

In particular, we study how changing the shape of this spherical scattering distribution affects appearance. The phase function can impact appearance in a perceptually important way near thin geometric structures, where light undergoes only a handful of scattering events before being emitted towards the observer. Figure 2.1(c) provides an example where, with all other material parameters being equal, certain phase functions induce a soft marble-like appearance (Image 4) while others create a “glassy” effect (Image 5) that has much sharper contrast near surface details, an appearance characteristic of materials such as white jade. Therefore, our study draws attention to this, often overlooked, crucial effect that the phase function can have on translucent appearance.

The space of phase functions is the space of probability distributions on the sphere of
The role of phase function. (a) Completely different phase functions can result in very similar appearances, as shown by the two images (top), rendered with the phase functions shown in the polar plot (below). (b) Our analysis identifies a two-dimensional space for phase functions, where neighborhoods correspond to similar translucent appearance. (Numbers indicate the location in the two-dimensional space of the phase functions used to render the corresponding images in (a) and (c).) (c) We derive perceptually uniform axes that we can use to explore this two-dimensional space of translucent appearance. This enables us to design two-lobe phase functions to match the appearance of materials such as marble (left), but also others, such as white jade (right), that single lobe phase functions cannot match well (middle).

directions, an infinite-dimensional function space. Typically, we regularize this space by only considering a family of analytic expressions controlled by a small number of parameters. The single-parameter Henyey-Greenstein (HG) distribution is used extensively because of its relative flexibility and sampling efficiency. This choice can, however, be restrictive, and it is known to be a poor approximation for some materials [101, 152], including common wax as shown in Figure 2.2. Therefore, we expand the space of analytic phase function models, by including von Mises-Fisher (vMF) distributions in addition to HG, and their linear combinations. In our results, we demonstrate the necessity of such more general phase function models, by showing that they considerably broaden the range of achievable scattering appearance.

Very different phase functions can produce the same visual appearance, even when the other material parameters remain fixed (see Figure 2.1(a)). This suggests that the perceptual space of translucency is much smaller than this expanded physical parameter space of phase functions. Our main goals were to obtain an intuitive understanding of this lower-dimensional perceptual space, and derive computational tools for navigating through it.
The most direct way to achieve our goals would be to densely sample the space of phase functions, and then design a psychophysical experiment to obtain perceptual distances between images rendered with all pairs of these phase functions. This has provided useful insights into the perception of certain aspects of opaque reflectance [118, 147]. In the present case this approach is not tractable, however: the number of dimensions of the physical parameter space implies that even a coarse sampling would require more than 60 million human judgments. Our solution was to combine psychophysics with computational analysis, and we did so in three steps.

First, we densely sampled the physical parameter space of phase functions to compute thousands of images, and employed computational analysis with various image distance measures to find low-dimensional embeddings of this space. These embeddings turned out to be two-dimensional (see Figure 2.1(b)), and very consistent across different distance metrics, shapes, and lighting.

Second, we ran a psychophysical study to determine if this computationally-determined embedding is consistent with perception. We sampled from our large image set to obtain a representative subset that was small enough to collect psychophysically exhaustive pairwise perceptual distances. We found that these distances are consistent with a two-dimensional embedding that is similar to that found computationally, thereby affirming the perceptual relevance of that embedding.

Third, we performed statistical analysis of the computationally-obtained and perceptually-consistent two-dimensional embedding, to investigate how its dimensions relate to phase function shape. We identified two linearly independent axes that are described as simple analytic expressions of generalized first and second moments of the phase function. We also obtained a learned distance matrix that can be used to compute an efficient approximation to the perceptual distance between any two tabulated phase functions.

Our results can facilitate manipulating the phase function to achieve a target appearance in
material design applications. We provide an intuitive characterization of the visual effects achieved at different parts of the derived appearance space. We show what parametric families inside our expanded set of phase functions can be used to reach each part of this space, verifying that models more general than single-lobe HG are necessary in many cases. We validate that the two identified analytic axes are perceptually uniform “knobs” that can be used to select exact physical parameter values.

2.1 Background on Perception

**Perception of opaque materials.** Most previous research in the perception of material properties has focused on opaque surfaces, and color and lightness in particular (e.g., [17, 151]). There have also been numerous studies of the perception of gloss, with results suggesting that real-world lighting is important for perception [38]; while simple image statistics correlate with perceived gloss in some situations [109, 100, 132], other situations seem to require explicit reasoning about the three-dimensional scene geometry [84, 95].

In graphics, the perception of materials has focused primarily on gloss perception. Pellacini et al. [118] and Westlund and Meyer [146] developed psychophysical models of gloss perception. Vangorp et al. [141, 140] studied the effect of object shape on gloss perception. Rushmeier et al. [127], Aydin et al. [6], and Ramanarayanan et al. [122] developed image-quality measures to characterize shape, material and lighting perception in CG images and animations [103]; these metrics have been used to assess the fidelity of rendering algorithms for material appearance [87]. Khan et al. [83] leveraged perception to develop tools for intuitive material editing. Kerr and Pellacini [82] studied the usability of various material design paradigms.

Pellacini et al. [118] used human similarity judgments to derive two perceptually-uniform knobs for the space of Ward bi-direction reflectance distribution functions (BRDF). Wills et al. [147] introduced a novel non-metric multidimensional scaling (MDS) algorithm of
human similarity judgments to find similar knobs, but for navigating a much larger space of tabulated isotropic BRDFs. An alternative approach is proposed by Ngan et al. [108], who built a BRDF navigation tool by, instead of using human observers, approximating the perceptual distance between two BRDFs using pixel-wise differences between images they produce for the same scene geometry and lighting. The successes of these important studies provide motivation for both the computational and psychophysical methods that we use in our study of translucency.

**Perception of transparent and translucent materials.** Traditional research on perception of non-opaque materials has focused on scenes with overlapping thin transparent sheets (e.g., [97, 4]). There has been relatively little research on the perception of translucent three-dimensional objects [136]. A first study was presented by Fleming et al. [40, 39], using synthetic images of three-dimensional objects to identify image cues, such as specular highlights, edge blurring, and luminance histograms, that seem to be correlated with the perception of translucency. The importance of highlights and shading on surfaces of three-dimensional objects has also been studied by Koenderink and van Doorn [86] and Motoyoshi [99]. Cunningham et al. [24] parameterized translucency scattering and absorption, along with gloss, with a focus on aesthetic judgments. None of these approaches attempts to parameterize the space of translucent materials numerically in a way that could be used for rendering, design and acquisition applications in graphics.

### 2.2 Phase Function Models

We restrict our attention to homogeneous media, which are described by a single set of parameters $\sigma_t$, $\sigma_s$, and $f_p$. The phase function describes the angular scattering of light at a small element of volume. Since solving the RTE usually involves drawing samples from this function, convenient analytic forms are desirable, and the form that is most often used is
due to Henyey and Greenstein [61]:

$$f_{p,HG}(\theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{3/2}}$$

(2.1)

$\theta \in [0, \pi]$. This is proposed as an approximation to the Mie solution [139] and is used widely because of its attractive combination of efficiency and relative flexibility. It is controlled by a single parameter $g \in [-1, 1]$, and can represent both predominantly forward and backward scattering ($g > 0$ and $g < 0$, respectively). The shapes of the HG phase function for three different positive values of $g$ are depicted in Figure 2.2(d).

The single $g$ parameter, however, limits the flexibility of the HG model and sacrifices some physical plausibility, including not reducing to the Rayleigh phase function in the small-particle regime [21]. However, it is a common approximation in graphics; renderers, using similarity theory [150], often use $g = 0$ with appropriately scaled coefficients for simulating deep scattering.

To model more complex types of scattering, linear combinations of two HG lobes—one forward and one backward—can be used [66, 81, 74], but there are still materials for which this is a limited approximation. Consider, for example, the photographs of soap and wax in Figure 2.2(b)-(c), which we acquired from homogeneous material samples of thickness 4 cm with a viewing direction normal to the surface and laser illumination from an angle of 45° (halfway between the view direction and the surface tangent plane). These photographs have been coarsely quantized to a small number of gray levels (i.e., *posterized*) to emphasize the shapes of the emergent iso-brightness contours. They clearly show that these two materials exhibit very distinct scattering behaviors. We distinguish these patterns as being “teardrop” versus “apple”. The second column of the figure shows the best-fit simulation results that can be obtained when rendering using either a single HG or linear combinations of two HG phase functions, and these are obviously qualitatively very different. In particular, we have found that the HG phase function—or any two-lobe mixture of HG phase functions—is unable to produce the apple-shaped iso-brightness contours that can be observed in the wax.
photographs. In Figure 2.2(a), we also show photographs under natural illumination of the same “frog prince” model made from these two materials, soap and wax; we can observe that the translucent appearance of the two objects is very different.

### 2.2.1 Henyey-Greenstein and von Mises-Fisher Lobes

One possible explanation for apple-shaped scattering is a phase function that is forward scattering, but with significant probability mass away from the forward direction, so that enough photons entering the material “turn around” after multiple bounces. HG lobes do not have this property, as different values of parameter $g$ only move between isotropic and strongly-forward lobes.

To augment the usual space of phase functions, we propose adding lobes that are shaped according to the von Mises-Fisher distribution on the sphere of directions [36]. This distribution has many of the properties of the normal distribution for real-valued random variables, and is often used as its analogue for spherical statistics, for example to represent BRDFs [52], or lighting environments [54]. Its probability density function is given as:

$$f_{p,vMF} (\theta) = \frac{\kappa}{2\pi \sinh \kappa} \exp (\kappa \cos \theta),$$

(2.2)
\[ \theta \in [0, \pi], \text{ and has a single shape parameter } \kappa: \]

\[ \kappa = \frac{2\bar{C}}{1 - M_C}, \text{ where,} \]

\[ \bar{C} = 2\pi \int_{\theta=0}^{\pi} \cos \theta p(\theta) \sin \theta \, d\theta, \quad \text{(2.4)} \]

\[ M_C = 2\pi \int_{\theta=0}^{\pi} (\cos \theta)^2 p(\theta) \sin \theta \, d\theta. \quad \text{(2.5)} \]

So, \( \kappa \) is a measure of concentration, being directly related to the average cosine \( \bar{C} \) and inversely related to the second moment of the cosine of directions on the sphere. Note that, like the HG distribution, the vMF distribution is suitable for rendering as it can be sampled efficiently [69].

Figure 2.2(d) compares HG and vMF lobes that have the same average cosine \( \bar{C} \)—equal in value to the shape parameter \( g \) for the HG distribution—, where we observe that the vMF distribution has more mass in off-frontal directions. We found that by using a phase function that is either a single forward vMF lobe or a linear combination where the forward lobe is a vMF lobe, one can successfully reproduce the apple-shaped scattering patterns demonstrated by the wax sample (Figure 2.2(c)).

### 2.2.2 Towards a Perceptual Space

We consider the augmented space of analytic phase functions consisting of linear mixtures of two lobes. An instance of a phase function is created by: deciding on the type of each of the forward and backward lobes (vMF or HG); selecting values for their respective shape parameters (\( \kappa \) or \( g \)); and selecting a value for the mixing weight between the two lobes. Our space of phase functions is thus a union of three-dimensional subspaces embedded in a five-dimensional ambient space.

The procedure just described involves selecting five parameters. This is inconvenient for design tasks such as adjusting scattering parameters to match a photograph to a desired appearance. The same is true even if one only allows linear mixtures of two HG lobes,
which requires adjusting three numerical values. The problem is that there is no intuitive relationship between the physical parameters and the appearance they induce, so one is essentially left to choose parameters blindly through exhaustive search.

While the physical parameter space is daunting, the perceptual space seems more manageable. As shown in Figure 2.1(a), different physical parameters can produce the same image, suggesting that the perceptual space of phase functions is smaller. Motivated by this belief, we seek to parameterize this smaller perceptual space, which we do in three steps. In the first step (Section 2.2.3), we design a large set of scenes with physical parameters that characterize the space of translucent appearance, under material, object shape, and lighting variations. The second step (Section 2.3) is to perform a large-scale computational analysis to examine the geometry of this space, by rendering thousand of images and using image-driven metrics to compare them. The third step (Section 2.4) is to use the findings of the computational analysis to design and perform a psychophysical study involving a manageable subset of scenes and phase functions. We process the results of these experiments in Sections 2.5 and 2.6, to discover a two-dimensional space of translucent appearance; we find perceptually uniform parameterizations of the two dimensions by analytic expressions of the first and second moments of the phase function, and show that they correspond to intuitive changes in the appearance of a translucent object. Finally, we conclude in Section 6.6 with discussion of future directions.

2.2.3 Characterizing Translucent Appearance

We begin our study by exploring the space of translucent appearance to determine the image conditions to use for our experiments. This involves examining rendered images corresponding to a large number of combinations of object shape, lighting, and material (i.e., physical parameter values).

To guide our design choices, we ran many pilot studies, each involving 3 to 6 human subjects
running a paired-comparison experiment for stimulus sets of 16 images and participating in informal exit interviews. The experimental and data analysis methodologies of these studies are identical to those described in Section 2.4 for our large-scale psychophysical experiments. We now describe our design choices based on our findings from these pilot studies.

2.2.4 Sampling the Physical Parameter Space

We require a sampling of the useful region of the physical parameter space that is dense enough to be visually representative. We formed such a set by considering phase functions with a single HG or vMF lobe, and those with linear mixtures of a forward and a backward lobe, each of which can be one of either HG or vMF. For the HG lobes, we considered the range of $g$ values often used in the rendering literature, and selected $g$ from $\{0, \pm 0.3, \pm 0.5, \pm 0.8, \pm 0.9, \pm 0.95\}$; for the vMF lobes we found through experimentation that $\kappa > 100$ produced unrealistic results, and selected $\kappa \in \{\pm 1, \pm 5, \pm 10, \pm 25, \pm 75, \pm 100\}$. We also chose values of the mixing weight of the first lobe from $\{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$, with the corresponding weight on the second lobe being one minus this value. From the set of phase functions produced by all these combinations, we remove those with negative average directions (i.e., materials that are more backward than forward scattering).

In pilot studies, using images rendered with different subsets from the above set of phase functions, subjects tended to cluster materials into unrealistic glass-like, and more diffuse translucent ones, and this variation dominated their responses. Through experimentation, we found that such appearance corresponded to large average cosine values ($\bar{C}$). By examining the image sets ourselves, we decided to remove phase functions with $\bar{C} > 0.8$, to arrive at a final set of 753 phase functions.

As we restrict our attention to phase function effects, we kept the scattering and absorption coefficients, $\sigma_s$ and $\sigma_a$, fixed for the majority of our simulations, except for the cases
mentioned in the following paragraph, and constant across color channels. For the values of scattering and absorption, we chose marble as a representative material often used in rendering. The parameters, selected from [75], are those of the green channel of marble: $\sigma_s = 2.62 \text{ mm}^{-1}$ and $\sigma_a = 0.0041 \text{ mm}^{-1}$. The choice of marble was also motivated by the fact that a large enough portion of the phase function space we used produced realistic appearance for the corresponding optical parameters. Subjects in pilot studies expressed preference for images rendered with these parameters than others we obtained through appearance matching experiments for the materials shown in Figure 2.2(a) ($\sigma_s = 0.99 \text{ mm}^{-1}$ and $\sigma_a = 0.01 \text{ mm}^{-1}$; and, $\sigma_s = 1.96 \text{ mm}^{-1}$ and $\sigma_a = 0.04 \text{ mm}^{-1}$).

Though the purpose of fixing other scattering parameters is to isolate the effect of the phase function, the obvious issue arises that the conclusions reached from our study may only be applicable to the selected coefficients $\sigma_s$ and $\sigma_a$. For this reason, in some of our simulations we also used perturbations of the scattering parameters around the values for marble we selected. Specifically, we considered an “optically thin material” with $\sigma_s = 1.31 \text{ mm}^{-1}$ and $\sigma_a = 0.00205 \text{ mm}^{-1}$ (same albedo as marble and twice its mean free path); and a “low absorption material” with $\sigma_s = 2.6241 \text{ mm}^{-1}$ and $\sigma_a = 0.00041 \text{ mm}^{-1}$ (same mean free path as marble and absorption decreased by a factor of 10).

In terms of surface properties, we modeled the material as a smooth dielectric, effectively reducing its BSDF to a Dirac delta function. The motivation for this choice is to have a glossy surface with strong specular highlights, which in the past have been reported to be important for the perception of translucency [39, 99].

### 2.2.5 Shape and Lighting

We considered a number of candidate scenes, each using one of four shapes: the Lucy, Dragon, and Buddha models from the Stanford 3D Scanning Repository [134]; and our own star-shaped Candle (Figure 2.3). This set includes complex and simple geometries.
There are thick parts whose appearance is dominated by multiple scattering and is therefore not very sensitive to the shape of the phase function, and there are parts for which dense approximations are not appropriate and low-order scattering is important.

We placed the selected shape with a manually-selected pose in a subset of the HDR environment maps presented by Debevec [26]: Campus, Eucalyptus, St. Peter’s, and a rotated version of Dining Room. This set includes the environment maps previously recommended for material perception [38], and represents different lighting conditions such as sidelighting and backlighting. We used a red-and-blue checkerboard as background. The checkerboard gave a patterned background to modulate the rays of light that scatter through the shape, and its use was also motivated from feedback from subjects in pilot studies, reporting difficulty segmenting grayscale objects from a uniform background. We did not use the checkerboard for the backlighting case, as we judged that in that case there was enough contrast with the background due to strong brightness changes at the border of the foreground object (see “Lucy + Dining room” in Figure 2.3). Overall, our scenes were designed to feature image cues that have been related to the perception of translucency by others [39, 99] or that we ourselves believe to be related. These cues include fine geometric detail, specular highlights, sharp edges, intensity gradients, color bleeding effects, and strong backlighting.

Based on subject feedback in pilot studies comparing the different objects under the same illumination, we determined that: the Candle and Buddha geometry do not induce a sufficient diversity and number of image cues for translucency; and, the Dragon induces cues that were non-uniformly distributed between the front and rear body, and has confounding features (head versus body parts). The Lucy provided the best compromise between cue diversity and uniform spatial arrangement. As shown in Figure 2.5(a), the dress exhibits fine details and sharp edges, specular highlights are present at the chest and the base, and a low-frequency intensity gradient is observed moving from the middle to the edge; furthermore, there is a plethora of thin (hands, wings, folds) and thick parts (body and
base), which are very prominent under backlighting conditions.

2.2.6 Design Choices

Based on the discussion in the previous two subsections, we selected nine combinations of scattering parameters ($\sigma_s$ and $\sigma_a$), geometry, and lighting to characterize translucent appearance, with the intention of studying, for each of these nine scenes, the image variation that is induced by changes in phase function. Our nine scenes can be grouped into three categories, based on which of the three components is variable.

1. **Scattering parameter variation.** Includes fixed shape (Lucy) and lighting (Campus), and varying $\sigma_s$ and $\sigma_a$ parameters: marble, optically thin material, and low absorption material.

2. **Shape variation.** Includes fixed lighting (Campus), $\sigma_s$ and $\sigma_a$ parameters (marble), and varying shape: Lucy, Buddha, Dragon, Candle.

3. **Lighting variation.** Includes fixed shape (Lucy), $\sigma_s$ and $\sigma_a$ parameters (marble), and varying lighting: Campus, Eucalyptus, St. Peter’s, and Dining room.

These are shown in Figure 2.3. As discussed above, the “marble + Lucy + Campus” scene is the “base” configuration, and used in all three categories. In the following, whenever not explicitly specified, the $\sigma_s$ and $\sigma_a$ parameters of marble are used.

Conducting psychophysical experiments with each of these nine scenes and each of the 753 phase functions discussed in Section 2.2.4 is possible nowadays through the use of crowdsourcing marketplaces, which facilitate large-scale access to human workers. However, in a number of pilot studies we ran using Amazon MTurk [3], we found that workers routinely failed simple quality control tests. The data we collected from these studies exhibited large inconsistencies across different workers, and because of the amount of
Figure 2.3: Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the cubic root metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them (see Section 2.5.1). To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

noise only provided us with coarse clusterings of images into two groups (roughly, more diffuse and more glass-like ones). On the other hand, doing psychophysical experiments at this scale in a carefully controlled laboratory environment is not tractable. We addressed this problem by combining computation with psychophysics, as described in the next two sections.

2.3 Computational Experiment

We determined the axes of the perceptual space of phase functions in two steps. The first step (this section) was to render thousands of images with different phase functions sampled from our physical parameter space, and to use many instances of MDS with
different between-image distance metrics to choose a single metric that provides consistent embeddings across our nine different scenes. The second step (Section 2.4) was to use our selected metric to generate a smaller set of stimulus images for a psychophysical study that produces a perceptual embedding.

2.3.1 Image-based Metrics

The geometry of a space can be described by the distance metric used to compare the objects within it, and one should choose a metric that is representative of the properties considered important. In our setting, objects are phase functions, and we are interested in their effects on translucent appearance. We therefore adopt the following approach, motivated by previous work on opaque BRDFs [108]: we compare phase functions by comparing images they induce for the same scene geometry and lighting.

How do we measure distance between two images? Our choices inevitably introduce bias to our subsequent characterization of the space of phase functions, and if we choose poorly, our conclusions will not generalize to different scenes and metrics. A good metric, then, should be able to isolate translucency variations and characterize translucent appearance in a consistent way across different “nuisance” conditions, to borrow from the statistics literature terminology, such as lighting and shape. At the same time, consistency alone is not enough: a metric collapsing all images to a single point is guaranteed to be consistent across all imaginable variations! A good metric should be descriptive enough to represent the richness of variations in translucent appearance, and prevent such degeneracies.

To ameliorate these issues we took the following approach. We performed our computational analysis multiple times, each with a different combination of scene and image metric, and compared the embeddings produced by all of them. Then, we selected as our final metric that which produces the most consistent embeddings across different scenes. We see later in Sections 2.4-2.6 that the metric selected by this strategy does indeed possess perceptual
relevance. In particular, it leads to predictions that can be validated psychophysically.

This approach required rendering thousands of images, and it was made possible by the availability of utility computing services. Specifically, for each of the scenes we selected in Section 2.2.4, we rendered 753 images, one with each of the phase functions discussed in Section 2.2.6. We used volumetric path tracing and the Mitsuba physically based renderer [68] on Amazon EC2 clusters [2]. Overall, this produced a final set of 6777 high dynamic range images.

For candidate image metrics, we considered three: pixel-wise $L_1$-norm, and $L_2$-norm differences of high dynamic range (HDR) images; as well as the perceptually-motivated metric proposed by Ngan et al. [108], the $L_2$-norm difference between the cubic root of HDR pixel values. In the following we refer to this last metric as the cubic root metric. All of our distance calculations ignored pixels of the background. In the appendix, we further report results from small-scale experiments using the image quality (mean-opinion-score) metric proposed by Mantiuk et al. [94], which we found to be consistent with those contained here.

### 2.3.2 Embeddings

Choosing a single scene and image metric induces a numerical distance between every pair of phase functions, and these distances can be used to find a low-dimensional embedding of the phase function space by classical MDS [23]. (Experimentation with metric MDS with different normalizations of the stress function produced almost identical embeddings.) We did this separately for each scene/metric pair, producing a set of such embeddings. Each embedding is a projection of the manifold occupied by phase functions in our ambient five-dimensional physical parameter space down to a Euclidean two-dimensional space where distances correspond to appearance differences as captured by the corresponding scene and metric. For all image metrics, the first two eigenvalues of the Grammian of the MDS output—the inner product matrix of the feature vectors produced by classical
Figure 2.4: Computational analysis results. (a) Comparison of two-dimensional embeddings produced using different image-based metrics, for three representative scenes. Points on embeddings are colored according to the square of the average cosine of the corresponding phase function. (b) Comparison of consistency of two-dimensional embeddings across the nine scenes of Figure 2.3, for each of the three image metrics. (c)-(e) Measures of two-dimensional embedding similarity for each pair of the nine scenes, when the cubic root metric is used to produce the embeddings.

MDS—captured at least 99% of the total energy. This indicates that the space of the phase functions, as described by these metrics, is close to being two-dimensional, so we can compare the different embeddings by visualizing them in two dimensions.

The two dimensional embeddings produced using the cubic root metric for all of the scenes are shown in Figure 2.3. For visualization, we first align all of the embeddings with the embedding for the “Lucy + campus” scene shown in (a), through a similarity transformation (rotation, uniform scaling, and translation) computed using full Procrustes superimposition [32]. In Figure 2.4(a), we show the embeddings produced from the different computational metrics we considered, for a representative subset of the scenes.

By visually comparing the embeddings shown in Figures 2.3 and 2.4, we can observe that the structure of all of the embeddings is remarkably consistent, in particular across different
scattering and absorption coefficients (Figure 2.3(b)) and shapes (Figure 2.3(c)). Larger differences are observed across lighting, and especially for the backlighting scene (“Lucy + dining room”) where there is a visible shear in the shape of the embedding. This is caused by the large changes in the brightness of the body of the Lucy across the different images; we discuss this effect in greater detail in Section 2.6.1. Even in this case though, the embeddings are quite consistent.

To quantify our observations, we compared embeddings using three measures of similarity: full Procrustes distances, and Pearson’s correlation coefficients between each of the two corresponding dimensions. In Figure 2.4(b), we report the root-mean-square of the full Procrustes distances, and average correlation values across the embeddings for the nine scenes of Figure 2.3, and for each of the three computational measures. Smaller variability, and therefore better consistency, corresponds to smaller values of the full Procrustes distances and larger correlation values. Based on this table, we found that the cubic root metric produces the most consistent results across scenes, and therefore we chose this metric as our measure of between-image distance.

In Figures 2.4(c)-(e), we report numerical values for each of the three embedding similarity measures we considered and all the scene pairs, for the case when the cubic root metric is used to produce the embedding. In agreement with our observations above, the largest differences occur for the backlighting scene. Even for this scene though, we can see in Figures 2.4(d)-(e) that there is very strong correlation between the corresponding coordinates of embeddings for different scenes, either vertical or horizontal. For each pair of the nine scenes, we performed a linear regression hypothesis test [19] for their corresponding coordinates, and found that the hypothesis of linear relation between them is statistically significant at the 99% confidence level—the same is true when embeddings are produced using the other two image metrics we considered.
2.4 Psychophysical Experiment

The computational experiment of the previous section provides a consistent metric for measuring distance between same-scene images of different translucent materials. Our next steps were to select a set of stimulus images using this metric, and then recover a perceptual embedding through psychophysical analysis. Our psychophysical analysis parallels that used for the study of gloss by [147]: we collected relative similarity judgments using a paired-comparison approach and then found an embedding through kernel learning.

2.4.1 Procedure

We designed a paired-comparison experiment in which each observer was shown a series of triplets of rendered images and was required to choose an image-pair that is more similar. An example is shown in Figure 2.5(b); the observer was required to indicate whether the center image is more similar to the image on the left or the image on the right. At the beginning of the experiment, observers were told that all images were rendered with the same shape and lighting, but with differing material properties.

In each trial (i.e., each triplet), the observer indicated their choice (left or right) by pressing the keys Q or P on a standard keyboard. The observers had unlimited viewing time for each trial but were told in advance that the average viewing time per trial was expected to be two or three seconds. Each subsequent trial followed immediately after the observer pressed a key to indicate their choice.

2.4.2 Stimuli

Though the analysis of Section 2.3 provides clear motivation for one particular image metric (the cubic root metric), it does not provide clear choices for shape, lighting, and scattering and absorption coefficients. For the reasons described in Sections 2.2.4, we
chose the $\sigma_s$ and $\sigma_l$ values of marble for generating the stimuli for psychophysical analysis. We also selected the Campus and Dining room lighting environments, corresponding to sidelighting and backlighting conditions, which have been demonstrated to be important for translucency [39].

In terms of shape, we mentioned in Section 2.2.5 that the Lucy has simultaneously the desirable properties of diversity of cues important for translucency, and uniform spatial arrangement. However, in pilot studies such as those used in Section 2.2.3, we found that subjects invariably attended only to the facial region of the model, which alone does not have the full set of cues. We tried both cropping and occluding the face, and selected the latter because it preserved the context of the original scene. Subjects indicated that, as a result, they examined all of the body parts, responding to a more diverse collection of cues, and that they did not consider the blocked-out region distracting.

Based on the above discussion, we settled on two different scenes, shown in Figure 2.5(a): the “Lucy + Campus” and “Lucy + Dining room”. We used these to perform two separate experiments of the kind described above, which we refer to as the sidelighting experiment and backlighting experiment, respectively. We emphasize that the consistency in Figures 2.3 and 2.4 suggests that these choices of scene are not critical, and that our image-sets can be reasonably well-characterized by any of the scenes.

Having settled on scene geometry and lighting, the number of stimulus images was decided by balancing the competing needs of covering the physical parameter space and reducing our reliance on mixing judgments by different observers. The number of triplets requiring judgments increases combinatorially with the size of the stimulus set, and reducing the set to a point where a single observer can provide judgments for all possible triplets would not allow adequate coverage of the parameter space. Conversely, when the size of the stimulus set is allowed to grow, we end up with many different observers providing judgments for distinct sets of triplets, making inconsistency a real concern. Pilot studies suggested that an observer could complete 2000 trials in approximately two hours, and that increasing this
number would result in increased fatigue and a drop in data quality. With this in mind, we set the stimuli set size to 40, implying a total of 30,000 possible triplets for each of the two experiments.

We used a clustering approach to select the 40 stimulus images from the “Lucy + Campus” dataset in Section 2.3. Using affinity propagation [41], images were grouped into clusters having large intra-group affinity, and simultaneously, an exemplar image was selected from each cluster. We used, as input to the affinity propagation, pairwise distances between images computed by the cubic root metric, selected based on the results of Section 2.3. Figures 2.6(a) and 2.7(a) show where the resulting exemplars reside in our previous computational embeddings for “Lucy + Campus” and “Lucy + Dining room”; by visual inspection,
we observe that they are relatively uniformly spread across the entire embedding. Additionally, we verified that the intra-cluster variation was small by visually inspecting the members of each cluster.

Each observer performed a subset of the 30,000 total triplets of either the sidelighting or the backlighting experiments; to create these subsets, we randomly permuted the 30,000 triplets and divided them into 15 parts with 2000 triplets per part. Each of the parts was assigned to one of 15 different observers for each experiment, for a total of 30 observers.

2.4.3 Implementation

The stimulus images were tone-mapped using a linear response and clipping at the same value across the dataset. They were presented in a dark room on a LCD display (ViewSonic HDMI 1080P, contrast ratio 1500:1, resolution 1920 × 1090, sRGB color space, max luminance 700 cd/m², 60:1 dynamic range, gamma 2.3). At a visual distance of 75 cm, images subtended 11.6 ° of visual angle.

Overall, 38 naive observers participated (15 observers for each of the two experiments, and 8 more for an inter-subject consistency test described below). All observers had normal or corrected-to-normal vision. Each observer completed 2000 trials, broken into 20 blocks of 100 trials each. Observers had the opportunity to rest between blocks, and before data collection commenced, each observer completed a small practice session of 20 trials.

2.4.4 Results

Consistency. Each observer in our two experiments provided responses for a unique set of 2000 triplets, and in order to obtain a perceptual embedding, we need to collect all 30,000 responses into a single pool for each experiment. For individual differences not to factor into the interpretation of the results, it is important that different observers agree on the
response for any particular triplet. To measure inter-observer consistency, we performed a separate small-scale study in which each of 8 observers evaluated the same set of 2000 randomly chosen triplets from the sidelighting experiment. By comparing their responses, we found consistency—defined as the number of observers that agree with the majority vote for each triplet—averaged across the 2000 triplets to be 75.94%. Both Pearson’s chi-squared and the likelihood ratio tests [19] show that the consistency of responses is statistically significant at the 99% confidence level, against the null hypothesis that observers decide based on chance.

**Learning from paired comparisons.** Having established consistency, we move on to computing a two-dimensional embedding using the pooled responses for each of the two experiments. An additional analysis step is required before we can compute this embedding using a method like MDS, because instead of having direct access to between-image distances as in the computational experiment of Section 2.2.3, we only have relative constraints of the form \( d(i, j) < d(i, k) \). Motivated by Wills et al. [147] and Tamuz et al. [135], we address this by using the relative constraints to learn a *kernel Grammian matrix* \( K \) for the stimulus set. This is a \( 40 \times 40 \) positive semi-definite matrix whose entries \((K)_{ij}\) represent inner-product values between images \(i\) and \(j\) in some feature space [129]. Once this matrix is learned, the between-image distances needed for computing our embedding are given by

\[
d_K(i, j) = \sqrt{(K)_{ii} + (K)_{jj} - 2(K)_{ij}},
\]

in direct analogy to the expression of Euclidean distance in terms of the standard dot product. By learning the kernel Grammian from the observers’ paired comparisons, our intention is that distances given by (2.6) will provide a good match to perceptual distances.

The learning algorithm we use is inspired by [147]. We learn \( K \) by solving the optimization problem

\[
\min_{K \succeq 0} \lambda \|K\|_s + \frac{1}{S} \sum_{s=1}^{S} L(d_K(i_s, k_s) - d_K(i_s, j_s) + b),
\]

where summation is over the set \( S = \{(i_s, j_s, k_s), s = 1 \ldots S\} \) of triplets for which a human
observer has judged that $d(i_s,k_s) < d(i_s,j_s)$. In this equation, $\| \cdot \|_*$ is the nuclear norm of a matrix, equal to the sum of its singular values; $L(\cdot)$ is a loss function, described below; $d_K(i_s,k_s)$ and $d_K(i_s,j_s)$ are shorthand for entries of $K$ via (2.6); and $b$ is a margin parameter, set equal to 1, used to resolve the scale ambiguity in relative comparisons.

The objective of (2.7) provides a balance between an empirical loss term penalizing violations of the constraints in $S$, and a regularization term to prevent overfitting the data. The use of the nuclear norm for regularization is to encourage low-rank solutions $K$ [147], and also due to its provable generalization and robustness to noise properties [123]. In terms of the loss function, a natural choice would be to use the hinge loss,

$$L(x) = \max(x,0).$$ (2.8)

In this case, (2.7) is the Lagrangian formulation of the optimization problem of (2.6), with the slack variables and linear constraints of the latter replaced by the empirical loss term. In our implementation, we use a modification proposed by [126] called Huberized squared hinge loss,

$$L(x) = \begin{cases} \sqrt{4 + 4x}, & x > 2, \\ [\max(x,0)]^2, & \text{otherwise}. \end{cases}$$ (2.9)

This loss function shares the attractive properties of the hinge loss—convexity, linear penalization of violated constraints, and zero loss for satisfied constraints; but is also smooth. Together with the Lagrangian formulation, these modifications to the optimization problem proposed by [147] provide attractive scalability and speed, by avoiding the need for semidefinite programming solvers. In our implementation, we solve (2.7) using the accelerated proximal gradient algorithm of [137], described in Algorithm 1. We set $\lambda$ through cross-validation as in [147], and $\rho$ and $\delta$ (parameters controlling the speed of convergence of the algorithm) empirically.

**Perceptual embeddings.** Figure 2.6(b) visualizes the final two-dimensional perceptual embedding of the stimulus set, obtained by first learning the kernel matrix and then using it
Algorithm 1 Kernel learning from paired comparisons.

Require: \( S = \{(i_s, j_s, k_s), s = 1 \ldots S\}, \lambda. \)

\[ K_0 = 0; K_{-1} = 0; t_0 = 1; t_{-1} = 1; \lambda_0 = \delta \lambda. \]

\textbf{while} not converged \textbf{do}

\{Update low-rank kernel Grammian.\}
\[ Y_k = K_k + \frac{t_k}{t_{k-1}} (K_k - K_{k-1}). \]
\[ G_k = Y_k - \frac{1}{\rho_k} \sum_{s=1}^{S} \nabla Y_k L (d_{Y_k} (i_s, k_s) - d_{Y_k} (i_s, j_s) + b). \]
\[ (Q, e) = \text{eig} (G_k). \]
\[ K_{k+1} = Q \text{ diag} \left( \max \left( e - \frac{\lambda_k}{\rho_k}, 0 \right) \right) Q^T. \]

\{Prepare for next iteration.\}
\[ t_{k+1} = \frac{1 + \sqrt{4t_k^2 + 1}}{2}. \]
\[ \lambda_{k+1} = \max (\eta \lambda_k, \lambda). \]
\[ k = k + 1. \]

\textbf{end while}

Ensure: \( K = K_k. \)

---

as input to kernel PCA \[129\], for the sidelighting experiment (“Lucy + Campus”). Similarly, Figure 2.7(b) shows the perceptual embedding for the backlighting experiment (“Lucy + Dining room”). In both cases, the perceptual embedding can be directly compared to the two-dimensional embedding of the same 40 images that was obtained computationally in Section 2.3. The geometry of the perceptual and computational embeddings is very similar. In particular, we observe through close inspection that the overall ordering of the images is very consistent, though there are some notable differences. We note, for instance, that the ordering of images 2 and 6 in both embeddings is reversed. These two stimulus images appear significantly different than the rest of the dataset. This was reflected in subject responses for triplets including both images: we observed that when one image was in the middle, the two were very consistently judged as similar; whereas when neither of them was in the middle, there was no consistent trend in the responses. As a consequence, in the perceptual embedding they were both correctly placed at the upper-right corner of the embedding, but their relative position was inferred reversed. Differences are also seen in the upper-left corner, where all stimulus images have a comparatively similar appearance resembling marble. It is important to note that, due to the use of relative constraints, the
ordering of the points of the embedding is more important than their absolute locations. We use the drawing in the inset figure to the left to provide some intuition for why this is true:

Consider the case of placing a point (point 1) on a plane, subject to constraints on its distances from two other points (points 2 and 3) on the plane. The right and left part of the drawing shows all the admissible configurations for compatible absolute and relative constraints, respectively. Absolute constraints permit only two isolated locations as solutions (shown as blue circular points, on the right), whereas with relative constraints...
Figure 2.7: Comparison of computational and perceptual embeddings for the backlighting experiment. (a) Locations of the stimulus images selected by the clustering algorithm, overlaid on the two-dimensional computational embedding. (b) Two-dimensional embedding learned from the data collected from the psychophysical experiment. The inset table shows values for Pearson’s correlation coefficient between corresponding coordinates of the cluster centers on the two embeddings. (c) A subset of the stimulus images, arranged in a grid according to their positions in the computational embedding. The same grid is also consistent with the positions of the images on the perceptual embedding. Points on embeddings and images on the grid are colored according to the square of the average cosine of the corresponding phase function, and consistently labeled across (a)-(c).

possible solutions span almost an entire half-plane (shown in blue, on the left).

In Figures 2.6(c) and 2.7(c), a subset of the embedded images are shown arranged in a grid consistent with their positions in both the computational and perceptual embedding. In both figures, we also report correlation values between corresponding coordinates of the perceptual and computational embeddings, for each of the two scenes. In all four cases, a linear regression hypothesis test showed that the hypothesis of linear relation between corresponding coordinates is statistically significant at the 99% confidence level.

As we mentioned at the beginning of this section, we found that there was significant
consistency in the responses provided by the study subjects. In order to further examine the stability of the perceptual embeddings to noise in the responses, we performed a bootstrapping analysis: we randomly selected a percentage of the 30,000 triplets from the sidelighting experiment equal in size to the percentage of inconsistent responses, switched the (binary) response provided by the subjects for those, and then learned an embedding using the perturbed data. We report the results of this analysis in the appendix, where we observe that the embeddings learned from such perturbed data sets are very similar to the embedding learned from the actual user data. We believe that this robustness to noise is due to the strong and correlated constraints the full set of relative comparisons imposes on admissible geometric configurations, as well as the generalization properties of the nuclear-norm regularization we use in the learning algorithm.

2.5 Analysis

In this section, we characterize the geometry of the two-dimensional perceptual embedding, and we derive computational tools for translucent material design, including analytic expressions for two distinct perceptual axes and a perceptual metric for comparing tabulated phase functions. First, though, it is worth discussing the significant implications of the consistency between the perceptual embeddings of the previous section and the computational embeddings of Section 2.3.

Consistency between perceptual and computational embeddings implies that any geometric structure we discover in the computational embedding—including, for example, expressions for axes and learned metrics—is likely to exist in the perceptual space as well. This is a very useful property, as it means that we can leverage powerful data mining techniques to discover geometric structure in the large computational datasets of Section 2.3, and then interpret this as structure in the perceptual space. Without consistency between embeddings, such tools would not be applicable, because it is not feasible to run psychophysical paired
comparison experiments at the scale required to generate embeddings of a sufficiently large set of images. (Perceptually embedding a stimulus set of even 500 images would require 60 million human judgments.) Without this consistency, we would be left to draw conclusions from dozens of embedded points instead of thousands, and we would be limited to using a much weaker set of tools.

A similar benefit comes from the consistency of computational embeddings across different scenes (Section 2.3.2). These results suggest that structure we discover in one of our two-dimensional embeddings is likely to generalize, at least to some degree, to different scene geometries and lighting conditions.

In light of this discussion, most of the analysis in this section is based on the computational two-dimensional embedding obtained with the “Lucy + Campus” scene (Figure 2.3(a)), with the expectation that results also apply to the perceptual space of phase functions.

2.5.1 Parameterization of the 2D Perceptual Space

We begin by characterizing the two dimensions of the computational embedding in terms of functionals of the phase function. We seek uniform parametric representations of the two dimensions, meaning that: 1) a fixed increment to any of the two parameter values induces an equal-length movement in the two-dimensional space regardless of the initial parameter values before the increment was applied; and 2) by adjusting both parameters, we can get from any point in the two-dimensional space to any other. Parameters defined in this way facilitate designing a phase function to correspond to any location on the embedding.

To identify our two parameters, we considered a large candidate set of statistical functionals of probability distributions, such as mean, variance, kurtosis, skewness, moments of different orders, and different analytic functions of these quantities. We then examined the linear dependence between these candidate functionals and the two coordinates of the embedding, using a combination of intuition and exhaustive computational search. To quantify the linear
dependence, we used Pearson’s correlation coefficient [19], which takes values $r \in [-1, 1]$. Larger absolute values indicate stronger linear dependence, while lower ones imply either lack of dependence or absence of a linear relationship. Since any functional can be negated, we are not concerned with the sign of the coefficient, and for the remainder of this section reported values correspond to the absolute value $|r|$.

**Vertical dimension.** We observed through experimentation that the vertical dimension of the computational embedding is parameterized well by the average cosine $\bar{C}$ of the phase function, defined in (2.4). Based on this observation, we performed an exhaustive search over powers of the average cosine, $\bar{C}^a$, for $a \in [1.5, 2.5]$ and found that correlation is maximized by $a = 2.05$, where $|r| = 0.9939$. For the average cosine and its square, $a = 1$ and 2, the correlation is equal to 0.9621 and 0.9939, respectively. The square of the average cosine is very strongly correlated to the vertical coordinate, and the average cosine slightly less so. A linear regression test showed that the hypothesis of linear relation between $\bar{C}^2$ and the vertical coordinate is statistically significant at the 99% confidence level. Figure 2.8 shows the two-dimensional embedding, colored using $\bar{C}^2$. This result is the reason why we use $\bar{C}^2$ to color the majority of two-dimensional embeddings throughout this chapter, most notably Figure 2.3.

In Table 2.1, we report the Pearson’s correlation coefficient values between powers of the

<table>
<thead>
<tr>
<th>scene</th>
<th>corr with $\bar{C}$</th>
<th>corr with $\bar{C}^2$</th>
<th>best power $a \in [1.5, 2.5]$</th>
<th>corr with $\bar{C}^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucy + Campus</td>
<td>0.9621</td>
<td>0.9939</td>
<td>2.05</td>
<td>0.9939</td>
</tr>
<tr>
<td>Lucy + Dragon</td>
<td>0.9480</td>
<td>0.9871</td>
<td>2.22</td>
<td>0.9879</td>
</tr>
<tr>
<td>Lucy + Buddha</td>
<td>0.9365</td>
<td>0.9766</td>
<td>2.24</td>
<td>0.9774</td>
</tr>
<tr>
<td>Lucy + Candle</td>
<td>0.9388</td>
<td>0.9765</td>
<td>2.15</td>
<td>0.9768</td>
</tr>
<tr>
<td>optically thin material</td>
<td>0.9701</td>
<td>0.9950</td>
<td>1.89</td>
<td>0.9952</td>
</tr>
<tr>
<td>low absorption material</td>
<td>0.9699</td>
<td>0.9958</td>
<td>1.93</td>
<td>0.9959</td>
</tr>
<tr>
<td>Lucy + Eucalyptus</td>
<td>0.9586</td>
<td>0.9917</td>
<td>2.07</td>
<td>0.9918</td>
</tr>
<tr>
<td>Lucy + St. Peter’s</td>
<td>0.9399</td>
<td>0.9772</td>
<td>2.17</td>
<td>0.9776</td>
</tr>
<tr>
<td>Lucy + Dining room</td>
<td>0.9658</td>
<td>0.9913</td>
<td>1.96</td>
<td>0.9914</td>
</tr>
</tbody>
</table>

Table 2.1: Parameterization of the vertical dimension of the embeddings of Figure 2.3 ("corr" refers to the absolute value of Pearson’s correlation coefficient).
average cosine $\bar{C}$ and the vertical coordinate of the cubic-root-metric embeddings for all nine scenes of Figure 2.3. We observe that, in all cases, the square of the average cosine is strongly correlated with the vertical coordinate, and the best power $a \in [1.5, 2.5]$ is close to 2.

**Horizontal dimension.** By visually examining phase functions at different points across horizontal lines of the embedding, we observed that they differ in terms of the spread of their probability mass. We therefore considered a number of functions of the second moment of the cosine $M_C$, defined in (2.5), as well as its variance,

$$V_C = 2\pi \int_0^{\pi} (\cos \theta - \bar{C})^2 p(\theta) \sin \theta d\theta. \quad (2.10)$$

Table 2.2 shows the correlation values of the horizontal coordinate with different functions of these quantities, the largest being achieved by

$$D_C = 1/\sqrt{1 - M_C}, \quad (2.11)$$

which is related inversely with the second moment of the cosine $M_C$. The correlation value for this case, $|r| = 0.9122$, indicates that $D_C$ is strongly correlated with the horizontal coordinate. Figure 2.8 shows the two-dimensional embedding, colored using $D_C$. We observe that the axes corresponding to changes in $D_C$ are not exactly horizontal, but instead diagonal with small slope. As before, a linear regression test showed that the hypothesis of linear relation between $D_C$ and the horizontal coordinate is statistically significant at the 99% confidence level.

In Table 2.3, we report the Pearson’s correlation coefficient values between the quantity $D_C = 1/\sqrt{1 - M_C}$ and the horizontal coordinate of the cubic-root-metric embeddings for all nine scenes of Figure 2.3. We found that, for all scenes, the correlation with this quantity is the largest among all the other quantities we show in Table 2.2, though the order after the first was not identical across scenes.
Table 2.2: Absolute value of the Pearson’s correlation coefficient between functions of variance and the second moment of cosine, and the horizontal coordinate of the two-dimensional embeddings from Figure 2.3(a).

<table>
<thead>
<tr>
<th>scene</th>
<th>corr with $1/\sqrt{1 - M_C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucy + Campus</td>
<td>0.9122</td>
</tr>
<tr>
<td>Lucy + Dragon</td>
<td>0.9020</td>
</tr>
<tr>
<td>Lucy + Buddha</td>
<td>0.9172</td>
</tr>
<tr>
<td>Lucy + Candle</td>
<td>0.9181</td>
</tr>
<tr>
<td>optically thin material</td>
<td>0.9129</td>
</tr>
<tr>
<td>low absorption material</td>
<td>0.8912</td>
</tr>
<tr>
<td>Lucy + Eucalyptus</td>
<td>0.9256</td>
</tr>
<tr>
<td>Lucy + St. Peter’s</td>
<td>0.8757</td>
</tr>
<tr>
<td>Lucy + Dining room</td>
<td>0.9101</td>
</tr>
</tbody>
</table>

Table 2.3: Parameterization of the horizontal dimension of the embeddings of Figure 2.3 (“corr” refers to the absolute value of Pearson’s correlation coefficient).

2.5.2 Phase Function Distance Metric

The image-driven metric adopted in Section 2.3 is very cumbersome because comparing two phase functions requires first rendering two images and then comparing those. In this section, we learn an efficient distance metric that operates directly on tabulated phase functions and provides an approximation to this image-driven metric.

For two phase functions $p_1, p_2$, we considered metrics of the form

$$d_w(p_1, p_2) = \int_{\theta_1=0}^{\pi} \int_{\theta_2=0}^{\pi} \left( p_1(\theta_1) - p_2(\theta_2) \right)^2 w(\theta_1, \theta_2) \, d\theta_1 \, d\theta_2,$$

with symmetric and non-negative weight function $w$. Note that if

$$w(\theta_1, \theta_2) = \begin{cases} 1, & \theta_1 = \theta_2, \\ 0, & \theta_1 \neq \theta_2 \end{cases}, \quad \theta_1, \theta_2 \in [0, \pi],$$

then
the metric $d$ is the usual squared distance. Even though the squared distance is a straightforward way to compare two phase functions, it does not predict differences in translucent appearance well. To demonstrate this, we show in Figure 2.9(b) the two-dimensional embedding obtained using squared distance between phase functions as input to MDS. The geometry of the obtained embedding is very different from the embedding produced using the image-driven metric that we are trying to approximate.

To improve this approximation, we learned the weight function $w$ by finely discretizing the phase function domain $\theta \in [0, \pi]$ and optimizing the weight matrix representation of $w$ of the resulting Mahalanobis distance. We used the metric learning algorithm of Davis et al. [25] with the objective of having distances $(p_1 - p_2)^T w (p_1 - p_2)$ between discretized phase functions $p_1, p_2$ be close to those given by the image-driven metric. We found that the learned weight function had large values on its diagonal, that is when $\theta_1 = \theta_2$, and much smaller values otherwise. Figure 2.9(a) provides a visualization, in the form of a polar plot, of a smoothed version of the diagonal of the learned weight function (shown in blue), and how it compares with the squared distance from (2.13) (shown in red). (A visualization of the full two-dimensional function as a matrix is provided in the appendix.) The metric emphasizes differences in the backward scattering lobe. To assess the quality of the approximation to the image-based metric afforded by this learned weight function, we also show a two-dimensional embedding of the phase functions produced by MDS with distances given by the learned metric in Figure 2.9(c). This embedding indeed can be observed to have qualitatively similar structure to that obtained using the image-driven metric (Figure 2.8(a)). Quantitatively, their full Procrustes distance is equal to 0.3381; the absolute value of the Pearson’s correlation coefficient between corresponding dimensions of the two embeddings is equal to 0.96 and 0.89, for the vertical and horizontal coordinates respectively; and in both cases the hypothesis of linear relation between the corresponding coordinates is statistically significant at the 99% confidence level. For comparison, the full Procrustes distance of the embedding of Figure 2.9(b), produced using the simple squared distance, from that of Figure 2.8(a) is equal to 0.7760; and the Pearson’s correlation coefficient
Figure 2.8: Parameterization of the computational two-dimensional embedding. (a) Embedded points are colored using the square of the average cosine of the corresponding phase function, which parameterizes the vertical dimension of the embedding. (b) Embedded points are colored using function (2.11), which parameterizes well the horizontal dimension.

Figure 2.9: Results of metric learning for phase functions. (a) compares the polar plots of the diagonal of the learned weight function (shown in blue, smoothed to 15 first Legendre polynomial coefficients for visualization), and the uniform one, corresponding to the $L_2$-distance (shown in red). (b) and (c) show the embeddings of phase functions using the $L_2$-distance and the learned metric, with points colored according to the square of the average cosine of the corresponding phase function (compare with Figure 2.8(a)).

between corresponding coordinates is 0.3044 and 0.5729, for the vertical and horizontal coordinates respectively.
2.6 Phase Function and Material Design

Our analysis so far has shown that our expanded space of phase functions can produce a two-dimensional perceptual appearance space, with the two axes being controlled by functionals of the first and second moments of the phase function. In this section, we explore this space further with material design applications in mind, and seek to answer the following questions: what types of visual effects can be produced and with what phase function in different parts of the appearance space; and how to select a phase function for achieving a target material appearance.

2.6.1 Role in Translucent Appearance

**Single-lobe phase function models** that consist of either one HG or one vMF lobe provide a single-parameter family of distributions. Thus, by varying their shape parameter ($g$ or $k$ respectively) one can only access materials that lie on a one-dimensional curve within the two-dimensional embedding. Moreover, we find that for both the single-HG and single-vMF cases, the resulting one-dimensional curve of materials is almost vertical and near the left-most part of the embedding. (Dense samplings of these one-dimensional curves are depicted in Figure 2.10(a)-(b).) This shows that using only these single-lobe phase function models limits the set of reachable translucent appearances to those on the left side of the embedding.

**Mixtures of single-lobe and isotropic phase function models** are often used as an extension of the simple single-lobe model [74]. Such phase functions are parameterized by two parameters, one for the shape of the forward scattering lobe ($g$ or $k$), and another for the weight of the lobe in the linear mixture. Despite this added flexibility, the set of translucent appearances that can be reproduced is still restricted to the left side of the embedding, as shown in the top row of Figure 2.10.
Figure 2.10: Locations of points corresponding to different types of phase functions on the two-dimensional computational embedding. Embedded points corresponding to the specific type of phase function are shown colored according to the squared average cosine of the phase function, whereas the rest of the embedding is faded out. Top row: single lobe and single lobe + isotropic combinations are limited to the left part of the embedding. Bottom row: two lobe combinations achieve appearance towards the right in the embedding.

Double-lobe phase function models are necessary, therefore, for accessing the complete two-dimensional space of appearances. To show which parts of the two-dimensional space can be accessed using different combinations of HG and vMF lobes, the second row of Figure 2.10 highlights materials that are created when the forward lobe of the mixture is one of either HG (Figure 2.10(c)) or vMF (Figure 2.10(d)) type. While most of the embedding can be reached using either an HG or vMF forward lobe, there is a region on the right that can only be accessed using a vMF forward lobe. The importance of the backward lobe is also evident from Figure 2.9(a), where we can see that the learned weights for comparing phase functions place strong emphasis on the backward scattering part.
Visualization of the appearance corresponding to the middle and right parts of the embedding, can be achieved by looking at Figures 2.6(c) and 2.7(c). These correspond to phase functions that reside in an approximately-uniform grid of positions within the two-dimensional embedding, and the position of each image within Figures 2.6(c) and 2.7(c) is indicative of where the corresponding phase function exists in that grid.

We see that moving from top to bottom (across the vertical dimension parameterized by $\bar{C}_2$) results in more diffusion, an effect similar to that achieved by increasing the mean free path. Moving from left to right (across the horizontal dimension parameterized by $D_C$) results in greater surface detail and more “glassy” appearance. By selecting different points in the embedding, we can achieve different trade-offs between diffusion and sharpness. In the middle of the embedding we see the emergence of very distinctive appearance, part translucent, part glassy with sharp details, that is characteristic of materials such as high quality white jade.

To demonstrate the importance of the middle part of the embedding, Figure 2.1(c) shows renderings of the “Lucy + Campus” scene (with the checkerboard replaced by a concrete background), where we have matched the appearances of marble and white jade. For the marble image (Figure 2.1(c)-(3)), a broad single-lobe phase function (HG with $g = 0$) is used to produce the characteristic diffusion in the body of the object. In contrast, matching white jade requires much less diffusion and greater preservation of surface detail. Attempting to match this appearance with only HG, that is by staying at the left part of the embedding only, produces the middle image in Figure 2.1(c)-(4), which has most detail blurred out. On the other hand, by using a double-lobe phase function to move towards the middle of the embedding, we achieve the result shown at the right of Figure 2.1(c)-(5), which captures both of the desired characteristics.

Thus, a result of our study is that including a secondary small backward lobe in the phase function is necessary to achieve glassy translucent appearance, like that of white jade. This is a visually interesting region of translucent materials that is neither highly diffusive nor
highly transparent.

**The effect of material, lighting, and shape.** Our results in Section 2.3 show that the geometry of the appearance space is consistent across some variations of material (in the sense of different scattering and extinction coefficients), shape and lighting. This indicates that our conclusions in the above discussion generalize to the nine scenes of Figure 2.3, and to novel scenes. The computational embedding for the “Lucy + Dining room” scene, corresponding to backlighting, is the one exhibiting the largest structural differences from the other embeddings, and for this reason we discuss it in greater detail. In particular, the embedding shows a skewing, being “stretched” across the diagonal from bottom left to top right. This effect can be understood by examining the grid of images in Figure 2.7(c), and specifically all the diagonals from bottom left to top right. We observe that this diagonal direction corresponds to a decrease in brightness, both in the thin parts (wings, hands) and the thick body. In a lighting environment as dramatic as backlighting, this change in overall brightness may be an important cue that strongly affects the responses of observers, confounding other cues. Such large changes in overall brightness also strongly affect the image metrics used to produce the computational embedding.

This strong change in brightness when moving from bottom to top, and from left to right on the embedding, can be related, by considering the physics of subsurface scattering, to the parameterization we derived in Section 2.5.1 for the vertical and horizontal dimensions of the embedding. Moving from bottom to top on the embedding results in less forward scattering and more isotropic phase functions (smaller $\bar{C}$), implying that less light reaches from the back of the Lucy, where it enters the medium, to the front, where the camera observes it. Moving from left to right on the embedding results in phase functions with larger variance (larger $M_C$ and $D_C$), implying that more light gets scattered towards side-directions, and therefore less light exits from the front for the camera to observe. Strong backlighting conditions emphasize this effect of phase function in brightness, confounding its other effects in translucent appearance. Geometrically, this manifests itself as a shear
stress of the two-dimensional embedding, suppressing to an extent its two dimensions to a single diagonal dimension.

Despite this strong confounding effect, our numerical analysis in Sections 2.3.2 and 2.5.1 demonstrates that the two-dimensional structure of the embedding is very consistent with those for the other scenes. Furthermore, by looking at Figure 2.7(c), we can observe the same qualitative changes in translucent appearance for vertical and horizontal movements on the embedding (changes in amount of diffusion and surface sharpness, respectively), as those we described earlier in this section for the “Lucy + Campus” scene (Figure 2.6(c)). To highlight this consistency, we show at the bottom row of Figure 2.1(c) renderings of the same three materials as in the top row, using the backlighting environment. We observe again that, using the same navigation tools as before to utilize the full two-dimensional embedding, we can control the trade-off between diffusion and detail, and achieve the part translucent, part glassy with high sharpness, appearance of white jade.

The importance of phase function. We can summarize the above discussion as a list of salient points relating the phase function to translucent appearance:

1. Our expanded phase function space, of double-lobe models with HG and vMF distributions, creates a perceptual two-dimensional space of translucent appearance.

2. Our expansion of the phase function space is non-trivial; the expanded space is necessary for producing all of the appearance space, and simpler single-lobe or forward-isotropic mixture models can only produce a one-dimensional subspace.

3. The expansion of the appearance space itself is physically important, allowing to achieve realistic appearance corresponding to materials such as white jade.

4. The two dimensions of the appearance space can be intuitively described as controlling diffusion and sharpness.
5. The appearance space remains relatively unaffected by changes in scattering parameters other than the phase function, shape, or illumination, the last having the strongest effect.

The above points highlight the fact that the phase function can crucially affect the appearance of translucent objects. In the past, this effect has often been overlooked, in favor of emphasizing the role of the other scattering parameters (mean free path, albedo). This approach is justified when simulating higher-order scattering, which according to similarity theory can be modeled accurately using the isotropic phase function. In turn, this implies that phase function is less important for thick objects and optically dense materials, where higher-order scattering dominates appearance. However, most interesting real objects have both thin and thick parts, and many materials exist that are not as dense. In these cases, the contribution of the low-order scattering component, which is strongly affected by the shape of the phase function, to the overall appearance becomes non-negligible and, as our results demonstrate, perceptually important.

2.6.2 Phase Function Design

As we established in the previous subsection, selecting the right phase function can be critical for achieving a desired appearance in a material design application. This can be formulated as a dual task of selecting a convenient parametric family of phase functions, and then finding appropriate values for the corresponding parameters.

The qualitative observations of Section 2.6.1 and quantitative analysis in Section 2.5 can facilitate this process in a two-fold way. First, by considering the intuitive description of vertical and horizontal movements on the two-dimensional translucent appearance space (changes in amount of diffusion and sharpness, respectively), we can approximately localize the target appearance on this space. This, in turn, allows us to use Figure 2.10 to guide our selection of parametric family of phase functions.
Second, having settled on a parametric family, we can reparameterize it in terms of the analytic expressions $\tilde{C}^2$ and $D_C$ of the first and second moment of the phase function, respectively, that we derived in Section 2.5.1. These quantities uniformly parameterize the vertical and horizontal dimensions of the appearance-based two-dimensional embedding of phase functions. This further implies that they can be used as quantitative “knobs” to control the appearance effects of movements on the translucent appearance space in a perceptually uniform way: specifically, $\tilde{C}^2$ controls amount of diffusion, and $D_C$ amount of sharpness. Writing the selected parametric family in terms of these quantities enables us to do perceptually uniform interpolation directly in the phase function space, which can be helpful for fine-tuning the parameter values.

As an instructive example, consider the case of single-lobe phase functions. Based on the discussion of the previous subsection, these families span an almost vertical one-dimensional subspace of the appearance space, which is uniformly parameterized by $\tilde{C}^2$. For the case of HG lobes, where the shape controlling parameter $g$ is equal to $\tilde{C}$, this suggests that a perceptually uniform reparameterization can be obtained by simply using $g^2$ instead of $g$. Similarly we can interpolate appearance in a more perceptually uniform way for the vMF family, through the equation,

$$\tilde{C}^2 = (-1/\kappa + \coth \kappa)^2.$$  

(2.14)

In order to further validate these findings, we designed a simple graphical user interface, shown in Figure 2.11(a). The three images shown are all rendered with a single-lobe HG phase function; those to the left and right correspond to values $g = 0$ and $g = 0.9$, respectively, whereas the $g$ value for the middle image can be increased or decreased by the user, by pressing the corresponding buttons. This is achieved internally using pre-rendered images for a dense set of $g$ values in the interval $(0, 0.9)$. We then used this interface to perform the following psychophysical experiment: subjects were asked to adjust the middle image until they find the one that is, in their opinion, the best visual midpoint between
the images shown to the left and right, at which point their selection was recorded. Our previous findings predict that this middle image should correspond to a $g$ value different from the arithmetic mean 0.45 and near the quadratic mean 0.6364 of the endpoints 0 and 0.9.
<table>
<thead>
<tr>
<th>Parameterization</th>
<th>Lucy + Campus</th>
<th>Buddha + Campus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.6583 ± 0.1390</td>
<td>0.6786 ± 0.0476</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.6507 ± 0.0900</td>
<td>0.6945 ± 0.0612</td>
</tr>
</tbody>
</table>

**Table 2.4:** Mean value and 99% confidence interval for the image midpoint experiment described in Figure 2.11(a). In all cases, the mean response is close to the value 0.6364 predicted by our analysis.

We used two different scenes for the experiment, “Lucy + Campus” and “Buddha + Campus” from Section 2.2.3. For selecting the g values used to create the interpolation sequence, we considered both linear (uniformly in g) and quadratic (uniformly in g²) sampling from the interval (0, 0.9). Then, each subject performed the experiment using either only the linear or only the quadratic sampling setting for both scenes. We emphasize here that the reason for considering both parameterization settings was not to compare their utility for design applications, which was not the purpose of the experiment; rather, we wanted to make sure that the results would not be biased by the choice of sampling. In Table 2.4, we report the mean g values selected by the subjects for each sampling and scene, as well as 99% confidence intervals. We observe that, in all cases, the mean of the selected g values is significantly skewed towards the quadratic midpoint of 0.6364 predicted by our analysis, and the exact numerical value is always inside the confidence intervals; using ANOVA [19], we found that there is no statistically significant difference between the four sample means. To the contrary, using two-sided t-tests [19], we found that in all cases the sample mean is statistically significant from the value 0.45 predicted by linear scaling. These results are in agreement with the predictions of our analysis.

For visualization, in Figure 2.11(b)-(c), we compare renderings of the “Lucy + Campus” scene of Section 2.2.3, with single-lobe HG functions. The left-most and right-most renderings of each row correspond to a phase function with g = 0 and g = 0.9, respectively. Renderings in between use g values sampled from this interval, either uniformly in the parameter space (top row, g = 0.45), or uniformly according to g² (bottom row, g = 0.6364). We observe that transitions from left to right are more perceptually uniform in the second case.
2.7 Discussion

In this chapter, we focused on the effect of phase functions in the appearance of translucent materials, and highlighted that for many realistic geometries and materials this effect is perceptually important. We explored an expanded space of phase functions, double-lobe models with both HG and vMF distributions, and showed that it can produce a much richer set of realistic translucent appearance. We also took first steps towards parameterizing the perceptual dimensions of translucency, by deriving computational tools for describing and navigating this expanded space of phase functions in a perceptually meaningful way. We achieved our results using a new methodology that combines computational analysis and psychophysical experiments, and we found strong evidence that our results generalize to novel scene shapes and lighting environments. That said, additional tests will be required to assess generalization more completely.

Our findings can be used for the development of interfaces for design of translucent materials by novice users, without the need for understanding the underlying physical models. Inspiration for how to proceed towards that direction can be drawn from several analogous studies for BRDF editing [82, 38, 141, 83]. The utility of perceptually intuitive instead of physical parameters in interfaces for translucent appearance editing needs to be further evaluated, as has been done in the past for BRDF editing [82]. Such design tools should also correctly convey appearance to the user, and the appearance navigation tools we derived can potentially be used to discover canonical objects and lighting conditions suitable for this purpose.
Chapter 3

Radiative Transfer Background and Problem Setting

In this chapter, we provide background on the radiative transfer framework for describing light propagation inside scattering materials. This is an approximate ray optics model that assumes incoherent, unpolarized illumination. After introducing the radiative transfer framework, we will use it at the end of this chapter to formulate inverse scattering as an optimization problem involving radiometric measurements and material parameters. Our notation and terminology borrow extensively from [18, 142, 5, 8]. Other comprehensive treatments of the subject include [67, 20, 98].

3.1 Notation and Setting

We will use lower-case bold letters to denote points $x$ in the Euclidean space $\mathbb{R}^3$ and directions $\omega$ in the two-dimensional unit sphere $\mathbb{S}^2$. We will also use capital-case bold letters to denote point-direction pairs $X = (x, \omega)$, and the notation $x (X)$ and $\omega (X)$ to refer to the point or direction component of such pairs, respectively.
Figure 3.1: (a) Photons traveling inside scattering materials perform random walks that depend on the material parameters. (b) Computational imaging techniques capture measurements of such materials by collecting different subsets of photons, depending on the paths they follow. (c) Inverse scattering is the problem of inferring, potentially spatially-varying, scattering material parameters from such measurements.

We assume that $\mathcal{M}$, a subset of $\mathbb{R}^3$, is occupied by a scattering medium with uniform index of refraction $\eta$, corresponding to speed of light $c = c_0/\eta$ inside the medium. We use $\partial \mathcal{M}$ for the boundary of $\mathcal{M}$, and at every boundary point $x \in \partial \mathcal{M}$ we use $\hat{n}(x)$ for the outward normal vector. We also define sets $\Gamma_i = \{ (x, \omega) \in \partial \mathcal{M} \times S^2 : \omega \cdot \hat{n}(x) < 0 \}$ and $\Gamma_o = \{ (x, \omega) \in \partial \mathcal{M} \times S^2 : \omega \cdot \hat{n}(x) > 0 \}$ of position-direction pairs on the boundary $\partial \mathcal{M}$ pointing in or out, respectively.

In the radiative transfer framework, light propagation inside a scattering medium $\mathcal{M}$ is described in terms of idealized light particles, often called photons in an abuse of terminology, despite not having wave or quantum properties. As shown in Figure 3.1(a), these photons perform random walks consisting of stochastic scattering events that occur due to their interaction with the medium, causing the photons to either get absorbed or change direction of travel. These events can be of two types: surface events, which occur at the medium boundary, and volume events, which occur at points inside the medium.

The scattering events are determined by a set of material parameters. At every boundary point $x \in \partial \mathcal{M}$, the bidirectional scattering distribution function (BSDF) $f_s : \partial \mathcal{M} \times S^2 \times S^2 \to \mathbb{R}_{\geq 0}$ controls refraction and reflection events. Intuitively, the BSDF $f_s(x, \omega_o, \omega_i)$ at $x \in \partial \mathcal{M}$ determines the amount of illumination that is scattered towards direction $\omega_o \in S^2$, given illumination reaching that point from direction $\omega_i \in S^2$.  

54
At every interior point \( x \in \mathcal{M} \), the medium is characterized by a set of scattering parameters \( k = \{ \sigma_s, \sigma_a, f_p \} \), shown in Figure 3.1(a), where \( \sigma_s, \sigma_a : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0} \), and \( f_p : \mathcal{M} \times S^2 \times S^2 \rightarrow \mathbb{R}_{\geq 0} \). Intuitively, the scattering coefficient \( \sigma_s (x) \) and the absorption coefficient \( \sigma_a (x) \) determine the amount of illumination that is scattered (changes direction) or absorbed, respectively, at every volume event at point \( x \in \mathcal{M} \). Similar to the BSDF, the phase function \( f_p (x, \omega_o, \omega_i) \) determines the amount of illumination that is scattered towards direction \( \omega_o \in S^2 \), given illumination reaching point \( x \in \mathcal{M} \) from direction \( \omega_i \in S^2 \). Finally, the extinction coefficient \( \sigma_t : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0} \), defined at every point \( x \in \mathcal{M} \) as \( \sigma_t (x) \triangleq \sigma_a (x) + \sigma_s (x) \), determines the spatial frequency of volume events. Complementary to the above scattering parameters, we also define the mean free path and volumetric albedo \( d_m, \alpha : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0} \), respectively. At every point \( x \in \mathcal{M} \), the mean free path \( d_m (x) \triangleq 1 / \sigma_t (x) \) determines the average distance a photon travels between volume events; and the albedo \( \alpha (x) = \sigma_s (x) / \sigma_t (x) \) determines the portion of scattered to total incoming illumination. Given the redundancy in the definition of the scattering parameters, in the remaining, we will use \( k \) to refer to any non-redundant and complete subset of the above scattering parameters, for instance, \( \{ \sigma_s, \sigma_a, f_p \} \), \( \{ \sigma_t, \sigma_s, f_p \} \), and \( \{ \sigma_t, \alpha, f_p \} \). Additionally, we will use the notation \( k (x) \) to collectively refer to the values of the above scattering parameters at a point \( x \in \mathcal{M} \).

Finally, we mention that the phase function is typically assumed to be isotropic, meaning that it is invariant to rotations of the incident direction and cylindrically symmetric. This implies that it is a function of only \( \theta = \arccos (\omega_i \cdot \omega_o) \in [0, \pi] \). We will adopt this assumption everywhere in this thesis and, by abuse of notation, often write \( f_p (x, \theta) \) to mean \( f_p (x, \omega_o, \omega_i) \). For a discussion of the light transport framework in cases where the phase function is anisotropic, we refer to [71].
3.2 Radiative Transfer Equation

The fundamental quantity describing the propagation of illumination in the light transport model is the radiance $L : M \times S^2 \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$. Intuitively, the radiance $L(x, \omega, t)$ is equal to the rate of photons at point $x$ travelling towards direction $\omega$ at time $t$. For a detailed presentation of radiance and related radiometric concepts, we refer to [16].

We now assume that a temporally-varying light source is applied at the boundary of the medium. Then, the propagation of illumination through the medium can be described in terms of radiance through the time-dependent radiative transfer equation (RTE),

$$\frac{1}{c} \frac{\partial L(x, \omega, t)}{\partial t} + \omega \cdot \nabla L(x, \omega, t) = -\sigma_l(x) L(x, \omega, t) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) L(x, \psi, t) \, d\psi,$$

subject to the boundary conditions

$$L(x, \omega, t) = L_i(x, \omega, t), \quad (x, \omega) \in \Gamma_i, \quad (3.2)$$

$$L(x, \omega, t) = \int_{S^2} f_s(x, \omega, \psi) L(x, -\psi, t) |\psi \cdot \hat{n}(x)| \, d\psi, \quad (x, \omega) \in \Gamma_o, \quad (3.3)$$

$$L(x, \omega, t) = 0, \quad t < 0, \quad (3.4)$$

where $L_i$ describes radiance injected at the boundary of the medium by the light source. Equation (3.3) is often referred to in computer graphics as the rendering equation [79].

The time-dependent Green’s function $T(X_o, X_i, t)$ is the solution of the RTE at $X_o \in \Gamma_o$ and time $t$, for an input pulse of infinitesimal duration $\delta(t)$ and unit radiance at $X_i \in \Gamma_i$, that is

$$L_i(x, \omega, t) = \delta(x - x(X_i)) \delta(\omega(X_i) - \omega) \delta(t). \quad (3.5)$$

The Green’s function can also be defined inside the medium, but we restrict its point-direction arguments $X_o, X_i$ to $\Gamma_o, \Gamma_i$, as we assume that we can only inject and measure light at the medium’s boundary. For the same reason, we omit volumetric sources from Equation (3.1).
Throughout this thesis, we will frequently use the change of variables $\tau = ct$, to convert time to optical pathlength. Then, Equation (3.1) becomes

$$\frac{\partial L(x, \omega, \tau)}{\partial \tau} + \omega \cdot \nabla L(x, \omega, \tau) = -\sigma_t(x) L(x, \omega, \tau) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) L(x, \psi, \tau) \, d\psi,$$

(3.6)

where, in an abuse of notation, we use $L(x, \omega, \tau)$ to refer to $L(x, \omega, t/c)$. The boundary conditions in Equations (3.2)-(3.4) change accordingly. As with radiance, we will also abuse notation and use $T(X_o, X_i, \tau)$ to denote the corresponding pathlength-resolved Green’s function. This change of variable allows us to define $T$ in two equivalent ways. First, $T(X_o, X_i, t)$ describes the temporal propagation of a pulse of light of infinitesimal duration through the scattering medium. Second, $T(X_o, X_i, \tau)$ is equal to the radiance produced by accumulating contributions only from photons that travel paths starting at boundary point $x(X_i)$ with direction $\omega(X_i)$, ending at boundary point $x(X_o)$ with direction $\omega(X_o)$, and having total length $\tau$.

In the computer vision and graphics literature, more commonly encountered is the steady-state version of the RTE

$$\omega \cdot \nabla L^s(x, \omega) = -\sigma_t(x) L^s(x, \omega) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) L^s(x, -\psi) \, d\psi,$$

(3.7)

subject to the boundary conditions

$$L^s(x, \omega) = L^s_i(x, \omega), \quad (x, \omega) \in \Gamma_i,$$

(3.8)

$$L^s(x, \omega) = \int_{S^2} f_s(x, \omega, \psi) L^s(x, -\psi) |\psi \cdot \hat{n}(x)| \, d\psi, \quad (x, \omega) \in \Gamma_o.$$  \hspace{1cm} (3.9)

Compared to the time-dependent version, these equations are expressed in terms of the steady-state radiance $L^s: \mathcal{M} \times S^2 \to \mathbb{R}_{\geq 0}$, which is time-independent. We can define a Green’s function $T^s(X_o, X_i)$ for the steady-state case exactly analogously the time-dependent case. The steady-state formulation corresponds to the case where the illumination applied at the boundary of the medium remains constant over time. Alternatively, the illumination changes at a frequency much smaller than the frequency of light waves, making the partial
derivative with respect to time in Equation (3.1) vanish. Therefore, the steady-state case describes the radiance our eyes sample when we look at real scattering materials under standard, everyday light sources.

We mention that the Green’s functions $T_s$ and $T$ are the continuous version of the steady-state and time-resolved (or pathlength-resolved) light transport matrices used in the computer graphics and computer vision literature [107, 110]. In the following we use the terms Green’s function and light transport matrix interchangeably to refer to $T$.

In the following, we will only consider incident radiances of the form

$$L_i(x, \omega, t) = L^s_i(x, \omega) \delta(t), \quad (x, \omega) \in \Gamma_i.$$  \hfill (3.10)

That is, we will assume that time-varying incident radiances take the form of pulses of infinitesimal duration, with some time-independent spatio-angular profile $L^s_i$ on the inwards boundary of the medium. (See the appendix for a note on the units of the two sides of the above equation). Then, because the time-dependent RTE (3.1) is linear and shift-invariant with respect to time (and analogously for pathlength), we have that

$$L^s(x, \omega) = \int_0^\infty L(x, \omega, t) \, dt,$$  \hfill (3.11)

where $L$ and $L^s$ are the solutions to the time-dependent and time-independent RTE, respectively. We provide a detailed proof of this relationship in the appendix. Similarly, since the incident radiance (3.5) used to define the time-resolved Green’s function is of the form of Equation (3.10), we also have that

$$T^s(X_o, X_i) = \int_0^\infty T(X_o, X_i, t) \, dt,$$  \hfill (3.12)

and similarly for pathlength. Therefore, $T$ decomposes the steady-state radiance $T^s$ into components, each corresponding only to photons that travel a specific pathlength $\tau$ inside the medium.
3.3 Volume Rendering Equation

Both the time-dependent and steady-state forms of the radiative transfer equation are
integro-differential equations: they involve both derivatives and integrals. This is inconvenient
for analysis, and for this reason, the two equations are often encountered in equivalent
forms that use only integrals. These are referred to as the time-dependent and steady-state
volume rendering equation. We will also use these equivalent forms in the rest of the section,
and for this reason we provide them here. Before stating the equations, we introduce some
notation.

For any point \( x \in M \) and direction \( \omega \in S^2 \), we denote by \( y_r (x, \omega) = x - r\omega, r \geq 0 \) the
points of a half-line (ray) starting at \( x \) and having direction vector \( \omega \in S^2 \). Because \( M \) is
convex, we know that every such half-line intersects the boundary \( \partial M \) at a unique point.
We denote this point by \( y_{\partial M} (x, \omega) \), and the corresponding \( r \) by \( r_{\partial M} (x, \omega) \). Finally, given
two points \( x, z \in M \), we define the volumetric attenuation \( a (x, z) \) as

\[
a (x, z) = \exp \left( - \int_0^{\|x - z\|} \sigma_t (x - r\omega (z \rightarrow x)) \, dr \right). \tag{3.13}
\]

where \( \omega (z \rightarrow y) = (z - x) / \|x - z\|_2 \) is the direction of the vector \( (z - x) \).

Using this notation, the time-dependent volume rendering equation is

\[
L (x, \omega, t) = a (x, y_{\partial M} (x, \omega)) L^x_t (y_{\partial M} (x, \omega), \omega) \delta \left( t - \frac{r_{\partial M} (x, \omega)}{c} \right) \\
+ \int_0^{r_{\partial M} (x, \omega)} \sigma_t (x, y_r (x, \omega)) \sigma_t (y_r (x, \omega)) \\
\int_{S^2} f_p (y_r (x, \omega), \psi^T \omega) L (y_r (x, \omega), \psi, t - \frac{r}{c}) \, d\psi \, dr, \tag{3.14}
\]
and the steady-state volume rendering equation is

\[
L^s(x, \omega) = a(x, y_{\partial M}(x, \omega)) L_i^s(y_{\partial M}(x, \omega), \omega)
+ \int_{0}^{\tau_{\partial M}(x, \omega)} a(x, y_r(x, \omega)) \sigma_s(y_r(x, \omega))
\int_{S_2} f_p(y_r(x, \omega), \psi^T \omega) L^s(y_r(x, \omega), \psi) \, d\psi \, dr,
\]

where in both cases the solution is subject to the same boundary conditions as in the corresponding radiative transfer equation.

The derivation of the steady-state version is shown in [115] and involves integrating the two parts of the steady-state RTE. The derivation of the time-dependent version is not presented in previous literature, to the best of our knowledge, but follows by trivial adjustments to the derivation of the steady-state case.

### 3.4 Existence and Uniqueness

The conditions on the shape of \( \partial M \), the material \( k \), and the incident radiance \( L_i \) under which the time-dependent and steady-state radiative transfer equations (or the corresponding volume rendering equations) have a unique solution have been studied extensively in transport theory. We refer to [8] for a comprehensive review. Interestingly, for a given medium \( M \) and analogous incident radiance, the set of materials \( k \) for which the time-dependent RTE has a unique solution is larger than the corresponding set for the steady-state RTE.

We mention here only one set of sufficient conditions on \( k \), which corresponds to most real-world scattering materials. Namely, for all \( M \) and \( L_i \), the time-independent and steady-state
radiative equations have a unique solution if \( k \) satisfies

\[
\sigma_t(x) , \sigma_s(x) , f_p(x, \theta) \geq 0, \quad \forall x \in \mathcal{M}, \theta \in [0, \pi], \quad (3.16)
\]

\[
\sigma_t \in \mathcal{L}^\infty (\mathcal{M}) \quad (3.17)
\]

\[
\sigma_s(x) < \sigma_t(x), \quad \forall x \in \mathcal{M}, \quad (3.18)
\]

\[
2\pi \int_{\theta=0}^{\pi} f_p(x, \theta) \sin(\theta) \, d\theta = 1, \quad \forall x \in \mathcal{M}, \quad (3.19)
\]

where \( \mathcal{L}^\infty (\mathcal{M}) \) is the set of bounded functions on \( \mathcal{M} \). The first condition is true by definition. The rest of the conditions imply that the extinction coefficient must remain bounded, the absorption coefficient must be non-zero anywhere where the extinction coefficient is, and the phase function at every point must be a probability distribution. We will refer to materials \( k \) satisfying the conditions of Equations (3.16)-(3.19) as *admissible materials*, and we will denote the set of all such materials by \( \mathcal{A} \).

### 3.5 Imaging Scattering Media

The measurements produced using different imaging techniques can be described as different ways to sample the pathlength-resolved light transport matrix. If we consider a *sampling function* \( W : \Gamma_o \times \Gamma_i \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \) then all such measurements can be written in the form

\[
S(T, W) \triangleq \int_{\Gamma_i} \int_{\Gamma_o} \int_{0}^{\infty} W(X_o, X_i, \tau) \, T(X_o, X_i, \tau) \, d\tau \, dX_i \, dX_o. \quad (3.20)
\]

In this equation, we assume that \( T_m \) and \( W \) are regular enough to allow changing the integration order. The sampling function \( W \) can typically be decomposed into three components: First, an *emittance function* \( W_i : \Gamma_i \to \mathbb{R}_{\geq 0} \), which is non-zero on a subset of the medium inward boundary \( \Gamma_i \) and describes the incident illumination. (This corresponds to the boundary condition (3.2).) Second, an *importance function* \( W_o : \Gamma_o \to \mathbb{R}_{\geq 0} \), which is non-zero on a subset of the outward boundary \( \Gamma_o \) and corresponds to the rays accumulated by the sensor. Third, a *pathlength sampling function* \( W_\tau : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \), which is non-zero only
for a subset of possible pathlength values. Then,

\[ W(X_o, X_i, \tau) = W_o(X_o) W_i(X_i) W_r(\tau). \] (3.21)

Note that the emittance function \( W_i \) is equivalent to the incident radiance term \( L_i^s \) we use in the boundary conditions for the radiative transfer equation.

**Steady-state imaging.** Conventional imaging sensors measure all photons, regardless of the distance they have traveled inside the medium. This corresponds to using a pathlength sampling function \( W_r(\tau) = 1 \), for all values of \( \tau \). Equivalently, using Equation (3.12), we can think of conventional imaging sensors as devices that sample the steady-state light transport matrix \( T^s \).

**Pathlength decomposition.** Pathlength decomposition techniques include transient imaging [143] and optical coherence tomography, which we discuss in more detail in Chapter 6. These techniques discriminate between photons based on the pathlength they travel (Figure 3.1(b): low-saturation, short paths versus high-saturation, long paths). Ideally, they sample pathlength slices of the light transport matrix, that is,

\[ W_r(\tau) = \delta(\tau - \tau_c), \] (3.22)

for some \( \tau_c > 0 \). Real systems instead have finite pathlength resolution, with \( W_r(\tau) \) being, say, a Gaussian or a square function. Typically, pathlength decomposition techniques densely capture multiple such pathlength slices, each centered at a different \( \tau_c \).

**Spatial probing.** Cameras typically have multiple sensor elements (pixels) that capture parallel sets of measurements \( \{ S(T, W_p^p \cdot W_i \cdot W_r), p = 1, \ldots, P \} \), with \( P \) the number of pixels. Conventionally, the measurements in such a set use different importance functions \( W_p^p \) and a common illumination \( W_i \). An alternative is to use spatial probing techniques [113] that allow different pixels on the same sensor to capture measurements corresponding to different importance-emittance pairs, \( \{ S(T, W_o^p \cdot W_i^p \cdot W_r), p = 1, \ldots, P \} \). With reference
to Figure 3.1(b), this allows the orange and blue camera pixels to only measure photons that begin at the orange and blue source pixels, respectively. Note that equivalent measurements can be obtained by capturing multiple images sequentially, each time using a different source $W^p_i$ and discarding unneeded pixels. Therefore, rather than providing a fundamentally different way to sample the light transport matrix, spatial probing allows reducing acquisition time through temporal multiplexing. We discuss spatial probing in more detail in Chapter 6.

**Types of sources.** It is useful to define an ideal source that emits an ideally collimated and narrow beam, corresponding to emittance

$$W_i(X_i) = \delta(x(X_i) - x_i) \delta(\omega(X_i) - \omega_i),$$

(3.23)

for some $(x_i, \omega_i) \in \Gamma_i$. Note that such a source cannot be realized physically: The wave nature of light implies that a source cannot be perfectly concentrated in both the spatial and angular domains simultaneously [46]; and any source ideally concentrated in either domain would have zero etendue and power [16].

The utility of ideal sources lies in their convenience for analysis [8], and for synthesizing other, more realistic sources. For instance, a perfectly collimated area source can be created by combining ideal sources with different $x_i$ but sharing the same $\omega_i$. A diffuse point source is created by combining ideal sources that share the same $x_i$, for each $\omega_i \in S^2$. Finally, a physical source with non-zero spatio-angular extent $E_i \subset \Gamma_i$ can be created as a weighted combination of ideal sources for each $(x_i, \omega_i) \in E_i$, with the weights depending on intrinsic (source power distribution) and extrinsic (geometry) factors.

**Types of cameras.** Similar to the ideal pointer source, we can define an ideal sensor element with importance function

$$W_o(X_o) = \delta(x(X_o) - x_s) \delta(\omega(X_o) - \omega_s),$$

(3.24)
for some \((x_s, \omega_s) \in \Gamma_o\). Arrays of such sensors can be used to construct different types of cameras. An \textit{orthographic camera} has multiple ideal sensor elements, each with a different \(x_s\) but all sharing the same \(\omega_s\). As with sources, real sensors cannot be perfectly concentrated in either the spatial or angular domain, but can still be expressed as weighted combinations of ideal sensor elements.

Simultaneously using an ideal sensor, pointer source, and pathlength decomposition eliminates the integrals in Equation (3.20), sampling a single value of the pathlength-resolved light transport matrix, \(T ((x_s, \omega_s), (x_l, \omega_l), \tau_c).\)

### 3.6 Inverse Scattering by Appearance Matching

We can now formally define the inverse scattering problem (Figure 3.1(a)). We assume we are given a set \(\{I^n, n = 1, \ldots, N\}\) of calibrated measurements with known corresponding sampling functions \(\{W^n, n = 1, \ldots, N\}\), of a medium with known 3D shape, BSDF \(f_s\) and index of refraction \(\eta\), but unknown scattering material parameters \(k\).

We cast the problem of inferring the material properties in an \textit{appearance matching} framework: find the material parameters \(k\) that best reproduce the measurements in the least-squares sense. Let \(\pi\) be an appropriate \(M\)-dimensional parameterization of the material \(k\), and \(k(\pi)\) be the corresponding material values. Then, appearance matching corresponds to solving the optimization problem,

\[
\min_{\pi, k(\pi) \in \mathcal{A}} \sum_{n=1}^{N} \frac{1}{2} (I^n - S (T_{\pi}, W^n))^2, \tag{3.25}
\]

where we use \(T_{\pi}\) to make explicit the dependence of the light transport matrix on the material parameterization \(\pi\). Throughout this thesis, we will denote the loss function of the above optimization problem by \(E(\pi)\).

The primary difficulty in solving the above optimization problem comes from the need
to evaluate the light transport matrix $\mathcal{T}_\pi$, which in turn requires solving the RTE for all possible materials $k(\pi)$. In Chapters 4 and 5, we develop two Monte-Carlo algorithms for solving this optimization problem efficiently. We compare the two algorithms in the appendix.
Chapter 4

Homogeneous Inverse Scattering

In this chapter, we develop an optimization procedure for solving the appearance matching problem of Equation 3.25 under two assumptions:

1. The medium $\mathcal{M}$ is homogeneous, meaning that the scattering material parameters $k = \{\sigma_t, \sigma_s, f_p(\theta)\}$ are the same everywhere in $\mathcal{M}$ and independent of position $x$.

2. All measurements of the medium are steady-state measurements, that is, $W_\tau(\tau) = 1$.

As discussed in the introduction, even this simplified version of the inverse scattering problem has only been addressed for two specific imaging regimes. The first is when image measurements can be assumed to contain only single scattering. The second is when all measured photons have undergone so many scattering events, that we can assume images to be in the diffusion regime. The requirement for single scattering or diffusion means that existing inverse scattering algorithms can only be used for either very optically thin or very optically thick materials. Alternatively, they algorithms can be applied to specific geometries where single scattering or diffusion apply, but the obtained measurements will not generalize to different geometries.

Another limitation of existing homogeneous inverse scattering techniques is that only recover
low-dimensional approximations to the phase function, which as discussed in Chapter 2 can be inadequate for many materials.

Motivated by the above limitations and the discussion in Chapter 2, in this chapter we develop an algorithm that can solve Equation 3.25 for homogeneous materials, overcoming both of the above limitations:

1. Our algorithm takes into account all orders of scattering, single, middle, and high, in a physically accurate way. This makes it more broadly applicable, allowing us to use it to measure materials such as turbid liquids and solids that could not be measured before.

2. Our algorithm can recover general phase functions \( f_p (\theta) \), and is not restricted to HG or other low parameter models. These general phase functions are particularly visually important for accurate visual appearance of objects with thin features.

In Chapter 5, we develop an alternative inverse scattering algorithm that we apply to heterogeneous materials. We compare the two algorithms in the appendix.

4.1 Operator-theoretic Formulation of Volumetric Light Transport

We begin our derivation by presenting the operator-theoretic formulation of radiative transfer, which we will use to setup a tractable optimization procedure for the appearance matching objective. The specific formulation we use is tailored toward our optimization algorithm, but approximations of similar nature have been considered for rendering applications [128, 12].

We begin by considering a discretization of the time-dependent volume rendering equation (3.14). We can approximate the integrals in this equation using the rectangle method [91].
Assuming a step size \( h = r_{\partial M} (x, \omega) / N \), we have

\[
L \left( x, \omega, N \frac{h}{c} \right) \approx \exp \left( -\sigma_t Nh \right) W_i \left( x - Nh\omega, \omega \right) \\
+ h \sum_{n=1}^{N} \exp \left( -\sigma_t nh \right) \sigma_s \int_{S^2} f_p \left( \Psi \cdot \omega \right) L \left( x - nh\omega, \Psi, (N - n) \frac{h}{c} \right) \, d\Psi. \tag{4.1}
\]

In the above, we have used the fact that, for constant \( \sigma_t \), the volumetric attenuation of Equation (3.13) becomes

\[
a (x, z) = \exp \left( -\sigma_t \| x - z \| \right), \tag{4.2}
\]

and therefore the attenuation terms in (3.14) simplify to

\[
a (x, y_r (x, \omega)) = \exp \left( -\sigma_t r \right). \tag{4.3}
\]

Equation (4.1) can be rewritten equivalently in recursive form as

\[
L \left( x, \omega, N \frac{h}{c} \right) \approx \exp \left( -\sigma_t h \right) L \left( x - h\omega, \omega, (N - 1) \frac{h}{c} \right) \\
+ h \exp \left( -\sigma_t h \right) \sigma_s \int_{S^2} f_p \left( \Psi \cdot \omega \right) L \left( x - h\omega, \Psi, (N - 1) \frac{h}{c} \right) \, d\Psi, \tag{4.4}
\]

with the initial condition

\[
L (x, \omega, 0) = W_i (x, \omega), \quad (x, \omega) \in \Gamma_i, \tag{4.5}
\]

\[
L (x, \omega, 0) = 0, \quad \text{otherwise}. \tag{4.6}
\]

Finally, using the first-order approximations

\[
\exp (-\epsilon) \approx 1 - \epsilon, \tag{4.7}
\]

\[
\epsilon \exp (-\epsilon) \approx \epsilon, \tag{4.8}
\]

and after some algebraic manipulation, Equation (4.4) becomes

\[
L \left( x, \omega, N \frac{h}{c} \right) \approx \left( 1 - h\sigma_t \right) L \left( x - h\omega, \omega, (N - 1) \frac{h}{c} \right) \\
+ h\sigma_s \int_{S^2} f_p \left( \Psi \cdot \omega \right) L \left( x - h\omega, \Psi, (N - 1) \frac{h}{c} \right) \, d\Psi. \tag{4.9}
\]
To simplify notation and make the discretization in time more explicit, we move the temporal variable of the radiance $L$ to a subscript and write:

$$L_n (x, \omega) \triangleq L \left( x, \omega, n \frac{h}{c} \right).$$  \hspace{1cm} (4.10)

Additionally, we define the following linear operator that acts on real functions on $\mathcal{M} \times S^2$

$$K_k (f) (x, \omega) \triangleq (1 - h\sigma_t) f (x - h\omega, \omega) + h\sigma_s \int_{S^2} f_p (\omega \cdot \psi) f (x - h\omega, \psi) \, d\psi,$$  \hspace{1cm} (4.11)

where the subscript $k$ is used to make explicit the dependence of this operation on the scattering material parameters. We can now rewrite Equation (4.9) in the form

$$L_n = K_k L_{n-1}. \hspace{1cm} (4.12)$$

Intuitively, the action of $K_k$ can be viewed as a single, constant-length step in the temporal propagation of a photon inside the medium. After traveling a distance of length $h$, the photon will transit in one of the following ways: 1) continue travelling in the same direction unaffected by the medium (probability $1 - h\sigma_t$); 2) undergo a volume event and continue travelling towards a new direction determined by the phase function $p$ (probability $h\sigma_s$); 3) undergo a volume event and be absorbed (probability $h (\sigma_t - \sigma_s)$). Consecutive applications of $K_k$ describe the time-resolved random walk the photon performs as it travels through the medium. By aggregating photons based on the distance they have travelled inside the medium, we can arrive at equation (4.12), which shows that the operator $K_k$ describes how the discretized time-resolved radiance evolves over time.

The above derivation has focused on time-dependent radiative transfer, and time-resolved radiance. For the steady-state radiance, by discretizing the integral in Equation (3.11), we have,

$$L^s \approx \sum_{n=0}^{\infty} L_n. \hspace{1cm} (4.13)$$
Using Equation (4.12), we can write the above in the form
\[ L^s \approx \left( \sum_{n=1}^{\infty} K^n_k \right) L^s_i = R_k L^s_i, \] (4.14)
where we define
\[ L^s_i = L_0, \] (4.15)
and
\[ R_k \triangleq (I - K_k)^{-1} = \sum_{n=0}^{\infty} K^n_k. \] (4.16)
The second equality in Equation (4.16) follows by using the Neumann series expansion on the inverse of \( I - K_k \). This assumes that \( I - K_k \) is invertible, or equivalently that the series converges. In the appendix, we prove that, for every admissible homogeneous material \( k \), this assumption is valid. We refer to \( K_k \) as the single-step propagation operator and to \( R_k \) as the rendering operator for material \( k \).

We note here that we derived Equation (4.14) starting from the time-dependent volume rendering equation (3.14), discretizing in time, applying the rectangle rule, using the first-order approximations (4.7)-(4.8), and finally using Equation (3.11). However, it is easy to see that we can arrive at (4.14) by applying the rectangle rule and using the first-order approximations (4.7)-(4.8) directly on the steady-state volume rendering equation (3.15). Therefore, this formulation provides another way to prove Equation (3.11), that is, that the time-resolved radiance provides a decomposition of the steady-state radiance into terms corresponding only to photons that have travelled for specific distance \( nh \) inside the medium.

Finally, given a steady-state sampling function
\[ W(x_o, x_i, \tau) = W_o(x_o) W_i(x_i), \] (4.17)
we can use Equation (4.14) to rewrite the corresponding measurement \( S(T, W) \) of Equation (3.20) as
\[ S(T, W) = D(I - K_k)^{-1} L^s_i, \] (4.18)
where $L^s_i$ is defined from $W_i$ using Equations (4.5), (4.6) and (4.15), and $\mathcal{D}$ is a \textit{sampling operator} defined as
\[
\mathcal{D}(f) \triangleq \int_{\Gamma_o} W_o(X_o) f(X_o) \, dX_o.
\] (4.19)

4.1.1 Step-size Selection

In the limit that $h$ goes to zero, the steady-state radiance $L^s$ of Equation (4.14) converges to the exact solution of the steady-state radiative transfer equation; and Equation (4.14) becomes equivalent to the steady-state volume rendering equation (3.15).

However, in general, the validity of the above operator-theoretic formulation relies on the accuracy of the first-order approximations (4.7)-(4.8) we used in its derivation. For the approximation to be accurate, we need to select a step-size $h = \epsilon / \sigma_t$ such that $0 < \epsilon \ll 1$. The effect of the choice of $h$ is shown in Figure 4.1, where we observe that, for small enough $h$, the above formulation produces equivalent renderings to the exact radiative transfer equation.

4.2 Appearance Matching for Homogeneous Inverse Scattering

Using the results of the previous section, we can now rewrite the appearance matching optimization problem (3.25) as
\[
\min_{k \in \mathcal{A}} \sum_{n=1}^{N} \frac{1}{2} \left( I^n - \mathcal{D}^n(I - K_k)^{-1} L^{s,n}_i \right)^2.
\] (4.20)

In the rest of this section we derive an optimization algorithm for this appearance matching problem. We represent the material parameters as a convex linear combination of a material dictionary. We then differentiate the appearance matching error with respect to this material representation and derive an efficient optimization scheme based on stochastic gradient descent and Monte-Carlo rendering. Despite the highly non-linear problem, we show
through simulations that, for an adequate set of observations, the error surface is smooth without local minima and allows accurate reconstruction of material parameters.

4.2.1 Dictionary Representation of Homogeneous Materials

To better parameterize the search space in Equation (4.20), we use a dictionary representation for the materials. Specifically, we consider a dictionary set of admissible homogeneous materials $D = \{k_m, m = 1, \ldots, M\}$, where $k_m = \{\sigma_{1,m}, \sigma_{s,m}, f_{p,m}(\theta)\}$, and their corresponding single-step propagation operators $\{K_{k_m}\}$. Then, for any weight vector $\pi$ in the $M$-dimensional simplex $\Delta^M$,

$$\pi = [\pi_m] \in \mathbb{R}^M, \quad \pi_m \geq 0, m = 1, \ldots, M, \quad \sum_{m=1}^{M} \pi_m = 1,$$  \hspace{1cm} (4.21)
we can represent a novel mixture material \( k(\pi) = \{ \sigma_t(\pi), \sigma_s(\pi), f_p(\theta, \pi) \} \), in terms of the dictionary atoms as

\[
\sigma_t(\pi) \triangleq \sum_{m=1}^{M} \pi_m \sigma_{1,m},
\]

\[
\sigma_s(\pi) \triangleq \sum_{m=1}^{M} \pi_m \sigma_{s,m},
\]

\[
f_p(\theta, \pi) \triangleq \frac{\sum_{m=1}^{M} \pi_m \sigma_{s,m} f_{p,m}(\theta)}{\sum_{m=1}^{M} \pi_m \sigma_{s,m}}.
\]

It follows trivially that, if all the dictionary materials satisfy the conditions (3.16)-(3.19), then so does the mixture material \( k(\pi) \). In other words, if all the dictionary materials are admissible, then all materials of the form \( k(\pi) \) are admissible as well.

In the following, we denote by \( K(\pi) \) and \( R(\pi) \) the single-step propagation (Equation (4.11)) and rendering (Equation (4.16)) operators, respectively, for the material \( k(\pi) \). We can then rewrite the appearance matching optimization problem (4.20) as

\[
\min_{\pi \in \Delta^M} \sum_{n=1}^{N} \frac{1}{2} \left( \tilde{I}^n - D^n (I - K(\pi))^{-1} L^n \right)^2.
\]

We now present the following claim about \( K(\pi) \).

**Claim 1.** For any vector \( \pi \in \Delta^M \), the single-step propagation operator \( K(\pi) \) for the mixed material \( k(\pi) \) defined by Equations (4.22)-(4.24), is a convex combination of the single-step propagation operators of the individual atoms with the same mixing weights,

\[
K(\pi) = \sum_{m=1}^{M} \pi_m K_{k_m}.
\]

We prove this claim in the appendix. Claim 1 shows that a convex combination of materials \( k_n \) can be directly identified with a convex combination of single-step propagation operators \( K_{k_n} \), and vice-versa. As a result, the objective function of the optimization problem of Equation (4.25) has a much simpler functional dependence on the parameters \( \pi \) we optimize over, allowing us to derive a tractable optimization strategy as discussed in the following subsection. This property is the key motivator for our use of the operator-theoretic formulation of
Section 4.1, and the use of convex combinations $\pi$ with the mixing Equations (4.22)-(4.24).

### 4.2.2 Rendering Gradients of the Loss Function

To minimize the appearance matching error of Equation (4.25), we begin by obtaining an operator-theoretic expression for the derivative of the appearance matching error of Equation (4.25) with respect to the mixing weights $\pi$. We use the following claim.

**Claim 2.** For the operator $\mathcal{K}(\pi)$ defined in Equation (4.26), the following differentiation rule holds

$$
\frac{\partial}{\partial \pi_m} (\mathcal{I} - \mathcal{K}(\pi))^{-1} = (\mathcal{I} - \mathcal{K}(\pi))^{-1} \mathcal{K}_m (\mathcal{I} - \mathcal{K}(\pi))^{-1}.
$$

(4.27)

This is a well-known result in the case of finite-dimensional matrices. In the appendix, we provide a precise statement and proof of the claim for the case of infinite-dimensional linear operators.

Using Equations (4.26) and (4.27), we can write the gradient of $E(\pi)$ with respect to each
coordinate of the mixture weight vector $\pi$ as

$$
\frac{\partial E}{\partial \pi_m} (\pi) = \sum_{n=1}^{N} 2 \left( \nabla^* \left( \mathcal{I} - \mathcal{K}(\pi) \right)^{-1} L^{s,n}_i \right) \nabla^* \left( \left( \mathcal{I} - \mathcal{K}(\pi) \right)^{-1} \mathcal{K}_{km} \left( \mathcal{I} - \mathcal{K}(\pi) \right)^{-1} L^{s,n}_i \right).
$$

(4.28)

Equation (4.28) shows that the gradient computation simplifies to a sequence of rendering and sampling operations:

1. Render radiance $L^n_1$ from incidence radiance $L^{s,n}_i$ using the rendering operator $\mathcal{R}(\pi)$.

2. Apply the single-step propagation operator $\mathcal{K}_{km}$, corresponding to the $m$-th material $k_m$ in the dictionary $D$, on $L^n_1$ to produce radiance $L^n_2$.

3. Render radiance $L^n_3$, by applying the full rendering operator $\mathcal{R}(\pi)$ on $L^n_2$.

4. Apply the sampling operator $\mathcal{D}^n$ on $L^n_1$ and $L^n_3$, and evaluate their product (Equation (4.28)).

These gradient evaluation steps are summarized in Algorithm 2 and visualized in Figure 4.2.

The fact that the loss function gradient can be expressed as a sequence of rendering steps has an important practical implication: We can efficiently compute stochastic estimates of the gradient using Monte-Carlo rendering techniques [35]. For the first stage, given that the rendering operator is asymptotically equivalent to standard volumetric rendering, we use a Monte-Carlo particle tracing process to estimate $I_1$, while simultaneously caching all intermediate particle positions in a set $C_1$ to form an approximation of $L_1$. This process is described in Algorithm 3. Then the application of $K_k$ on $L_1$ is stochastically approximated as described in Algorithm 4: particles are uniformly sampled from $C_1$, propagated by $h$, and then either scattered or absorbed. The results are cached in a set $C_2$ as an approximation of $L_2$. Finally, samples from $C_2$ are used as sources for another full particle tracing process that directly estimates $I_3$, without further caching. This is performed similar to Algorithm 3,
Algorithm 2 ComputeGradient.

Require: \( \pi \in \Delta^M \).
1: for \( m = 1 \) to \( M \) do
2: \( g_m \leftarrow 0. \)
3: for \( n = 1 \) to \( N \) do
4: Render \( L_1^n \leftarrow R(\pi) L_{s,m}^n \).
5: Apply single-step propagation \( L_2^n \leftarrow K_m L_1^n \).
6: Render \( L_3^n \leftarrow R(\pi) L_1^n \).
7: Sample \( I_1^n \leftarrow \mathcal{D}^n L_1^n \).
8: Sample \( I_3^n \leftarrow \mathcal{D}^n L_3^n \).
9: \( g_m \leftarrow g_m + 2 (I_1^n - I_3^n) n_3^n \).
10: end for
11: end for
12: return \( g \).

Algorithm 3 Adjoint particle tracing for computing \( L_1 \) and \( I_1 \).
1: Let \( x_0 \) be the location where the laser hits \( \partial M \).
2: \( x \leftarrow x_0, \omega \leftarrow \omega_L, t \leftarrow 1, C_1 \leftarrow \emptyset \).
3: while true do
4: \( t \leftarrow t \cdot a \).
5: Sample \( s \) from pdf \( p(s) = \sigma_1 \exp(-\sigma_1 s) \).
6: \( x' \leftarrow x + s \cdot \omega \).
7: if \( x' \notin S \) then
8: break
9: end if
10: Cache the particle location \( C_1 \leftarrow C_1 \cup \{ (x, \omega) \} \).
11: Let \( \psi_1 \) be the direction connecting \( x' \) and the camera.
12: Let \( y_1 \) be the intersection of ray \( (x', \psi_1) \) and the image sensor.
13: \( \psi_1 \leftarrow t \cdot p(\psi_1^T \omega) \cdot \exp(-\sigma_1 \|x' - y_1\|) \cdot P_0 c / A \),
   where \( c \) is the total number of pixels on the sensor, \( A \) is the sensor’s surface area, and \( P_0 \) is the source power.
14: Add \( \psi_1 \) to the corresponding pixel on the image sensor.
15: Sample a direction \( \psi_2 \) according to the phase function \( f_p \).
16: \( x \leftarrow x', \omega \leftarrow \psi_2 \).
17: end while

but with the initialization of \( x \) and \( \omega \) in Step 2 replaced by an initialization from a particle drawn uniformly from \( C_2 \), and with the caching Step 10 omitted.
Algorithm 4 Importance sampling for computing $L_2$.

1: $C_2 \leftarrow \emptyset$
2: Uniformly sample a pair $(x_0, \omega_0) \in C_1$.
3: $x \leftarrow x_0$, $\omega \leftarrow \omega_0$.
4: $x \leftarrow x + h\omega$.
5: Sample $u$ uniformly in $(0, 1)$.
6: if $u < h (\sigma_{t,n} - \sigma_{s,n})$ then
7: terminate
8: end if
9: if $h (\sigma_{t,n} - \sigma_{s,n}) < u < h\sigma_{t,n}$ then
10: Sample a direction $\psi$ according to the phase function $f_{p,n}$.
11: $\omega \leftarrow \psi$.
12: end if
13: Cache the particle location $C_2 \leftarrow C_2 \cup \{(x, \omega)\}$.

4.2.3 Stochastic Optimization

The availability of stochastic estimates of the gradient makes stochastic gradient descent (SGD) algorithms attractive for minimizing Equation (4.25). Similar to standard gradient descent, SGD algorithms perform iterations of steps proportional to the negative of the stochastic estimates of the gradient. These algorithms only require that the estimates be unbiased; even if they are otherwise noisy, there exist convergence guarantees analogous to those of standard gradient descent, with the noise variance only affecting convergence speed. This behavior has an important practical implication: we can reduce the number of particles in Monte-Carlo evaluations of the gradient, and speed computation at the cost of introducing noise to the gradient estimate. As long as the rendering algorithm is unbiased such noisy gradient estimates are valid inputs to SGD. The noise due to the reduced number of particles somewhat increases the number of iterations. However, it is still possible to reduce the number of particles quite drastically, and achieve a significant speedup in terms of overall computation time [15]. Note that known convergence guarantees for SGD algorithms generally require at least convexity of the objective function. However, even when applied to non-convex problems, they have been empirically shown to converge to good solutions (for instance, when used for supervised back-propagation in multi-layer networks [90]).
Algorithm 5 Solve appearance matching optimization problem.

1: Initialize $p_m(0) \leftarrow 1/M$.
2: while not converged do
3: \hspace{1em} $g(t) \leftarrow \text{ComputeGradient} \left( \pi(t) \right)$.
4: \hspace{1em} $\pi(t+1) \leftarrow \mathcal{P}_{\Delta^M} \left( \pi(t) - \eta(t) g(t) \right)$.
5: end while
6: return $\pi^{\text{opt}} = \frac{1}{T} \sum_{t=0}^{T} \pi(t)$.

As the vector $\pi$ is constrained to lie on the simplex, we use projected stochastic gradient descent (PSGD). We denote by $g \in \mathbb{R}^M$ consecutive noisy estimates of the gradient of $E(\pi)$ obtained through Monte-Carlo rendering such that

$$E[g_m] = \frac{\partial E(\pi)}{\partial \pi_m}.$$  \hspace{1em} (4.29)

We use them to iterate

$$\pi(t+1) = \mathcal{P}_{\Delta^M} \left( \pi(t) - \eta(t) g(t) \right),$$  \hspace{1em} (4.30)

where $\mathcal{P}_{\Delta^M}$ denotes the Euclidean projection operator to the simplex $\Delta^M$ [34]. The step size is often chosen equal to $\eta(t) = \frac{c}{\sqrt{t}}$. Though the speed of convergence depends on the proportionality constant $c$, in practice SGD is known to be robust to this selection. To cancel noise in individual steps, SGD returns as its final output the average of all $T$ iterations, $\pi^{\text{opt}} = \frac{1}{T} \sum_{t=0}^{T} \pi(t)$. This procedure is summarized in Algorithm 5.

The optimization problem we solve is highly non-linear and essentially involves inversion of the photon random walk process. Despite the non-linear formulation, our simulations in Section 4.4.2 show it allows an accurate reconstruction of material parameters. While an exact proof of this property is a subject for further research, all our simulations indicate that, for an adequate set of observations, the error surface is very smooth and does not suffer from local minima, explaining the good convergence we are able to achieve.
4.2.4 Dictionary Design

The dictionary-based formulation of the appearance matching problem in Equation (4.25) can be used with any dictionary choice\(^1\). To complete the description of our homogeneous inverse scattering algorithm, we describe below our own simple dictionary \(D\). We start with the phase function component of the materials and then address the extinction and scattering components.

**Phase functions.** To allow the dictionary to be as general as possible, we aim to be able to express any isotropic phase function, up to some angular resolution. Hence, we model the phase function as a piecewise linear function of \(\theta\). Then, we can use a set of tent (triangular) functions, spaced equally over angular domain \(\theta \in [0, \pi]\), to approximate it. Denoting the bin number by \(M\) and the bin spacing by \(\theta_s = \pi / (M - 1)\), we use tent functions of width \(2\theta_s\) and centered at points \(\theta_{c,n} = (0, \theta_s, 2\theta_s, \ldots, \pi)\). Each of the tent functions is normalized to satisfy the constraint of Equation (3.19). In Figure 4.3, we show the phase functions in a tent dictionary with \(M = 10\). In our experiments, we use \(M = 200\), which corresponds to a discretization step of \(0.9^\circ\).

To avoid high frequency artifacts in the phase function solution we expand the loss function in Equation (4.25) to include a Laplacian smoothness term,

\[
\min_{\pi \in \Delta^M} \sum_{n=1}^{N} \left( D^n (I - K(\pi))^{-1} L^n - I^n \right)^2 + \lambda \pi^T (I - W) \pi, \tag{4.31}
\]

where \(I\) is the \(M \times M\) identity matrix, and \(I - W\) is the unnormalized Laplacian matrix corresponding to the adjacency matrix \(W\) with entries \(W_{ij} = \exp(-a |\theta_{c,i} - \theta_{c,j}|)\).

**Extinction and scattering coefficients.** The definition of the atoms’ extinction and scattering coefficients should ensure that the dictionary can represent materials with a wide range of

---

\(^1\)We use the term “dictionary” because the phase function sets we use can be under- or over-complete and not strictly “bases” in the technical sense.
\(\sigma_t, \sigma_s\) values. We select an upper bound \(\sigma_{t,\text{max}}\) on the desired extinction coefficients. Note that each admissible material \(\{\sigma_t, \sigma_s, f_p(\theta)\}\) can be represented in terms of convex combinations of three materials: \(\{\sigma_{t,\text{max}}, \sigma_{t,\text{max}}, f_p(\theta)\}\), the purely absorptive atom \(\{\sigma_{t,\text{max}}, 0, \varnothing\}\), and an atom of the form \(\{0, 0, \varnothing\}\) describing scattering-free propagation in vacuum. We use the symbol \(\varnothing\) to indicate that the last two atoms are independent of phase function (the phase function is undefined for these two media).

Based on the above considerations, we adopt a material dictionary including 200 atoms of the form \(\{\sigma_{t,\text{max}}, \sigma_{t,\text{max}}, f_{p,m}(\theta)\}\) with the \(f_{p,m}(\theta)\) defined above, plus the two purely absorptive atoms \(\{\sigma_{t,\text{max}}, 0, \varnothing\}\) and \(\{0, 0, \varnothing\}\). In our experiments we set \(\sigma_{t,\text{max}} = 200\text{ mm}^{-1}\), based on the results from Section 4.4.2.

**Other parameterizations:** Our specific choice of dictionary aims to represent any general phase function shape. Depending on the application, other dictionaries may be more appropriate, and some examples include: zonal spherical harmonics for low-frequency phase functions, phase functions derived from Mie theory [14, 42] when measuring dispersions, and compact dictionaries such as the set of Henyey-Greenstein and von Mises-Fisher functions of Chapter 2 when a simple phase function model is sufficient. A small adaptation can also allow differentiating directly with respect to the single parameter of a Henyey-Greenstein function (the average cosine). Our optimization framework is quite attractive even for retrieving such simpler phase functions, since it alleviates the need for input.
measurements which isolate single scattering events.

4.3 Acquisition Setup Design

The optimization strategy described above is general enough to be applied to captured data with any geometry, as long as we can calibrate the 3D shape of the material, the relative position of the camera and light source, and the indices of refraction of the scattering material and its surroundings. Below we describe the physical acquisition setup we built, which is motivated by the simplicity of this calibration process and by some considerations related to the stability of the optimization problem. As mentioned at the beginning of this section, we also limit ourselves to the use of only steady-state measurements. Inspiration for the design is also drawn from analogous designs in [75, 45, 145] and physics [76]. A schematic and a photograph of our acquisition setup are shown in Figure 4.4. Further implementation and calibration details are provided in the appendix.

**Geometry.** We cast the material we are interested in measuring into glass cells of variable thickness \( w \). This allows us to create box-shaped material samples whose exact shape is known with very high accuracy. Furthermore, using micron-accurate smooth glass surfaces
means that the BSDF functions \( f_s \) describing transition and refraction at the various material interfaces (material and glass, glass and air) can be easily simulated using Fresnel refraction and reflection laws (the so-called “smooth dielectric” BSDF). In our experiments, we use cells of widths \( w = 1, 2.5, 5, \) and \( 10 \) mm.

**Imaging and lighting.** We use an approximately orthographic camera with a high magnification macro lens (4.3° subtended angle and 1 : 3 reproduction ratio) to sample the light field produced by the material volume. We use narrow (1 mm diameter) collimated beams to illuminate the sample. We use a configuration that allows illuminating either the sample surface imaged by the camera (frontlighting), or its opposite (backlighting). Through a combination of two motorized rotation stages, we can achieve different combinations of front and back lighting directions \( \theta_f, \theta_b \) and viewing directions \( \theta_o \). In our experiments, we use all possible combinations of \( \theta_f, \theta_b, \theta_o \in \{ 5^\circ, 15^\circ, 25^\circ \} \), resulting in a set of 18 measurements per sample, 3 viewing directions times 3 frontlighting plus 3 backlighting directions.

The above combination of sample shape, camera, and illumination lends itself to accurate calibration. Additionally, and similar to arguments for BRDF acquisition [121], we argue that, to maximize angular information, the illumination should have broadband angular frequency content, and hence be as close as possible to a delta function. In the appendix, we provide a formalization of this argument. The utility of narrow, collimated beam sources for inverse scattering has been empirically demonstrated in the past, in different scattering acquisition systems [102, 45]. Finally, the use of both frontlighting and backlighting is motivated by the understanding that a backlighting beam produces measurements dominated by high-order scattering; such measurements are intuitively useful for determining the optical thickness of the material. Conversely, frontlighting results in measurements where low-order scattering is significant, and therefore is informative for the recovery of the material phase function.

**Multi-chromatic measurements.** Scattering parameters vary as functions of wavelength, and this spectral dependency can create perceptually important effects in appearance [37, 42].
To capture spectral variations we use monochromatic laser light at three RGB wavelengths, 488, 533, 635 nm, and solve the optimization problem of Equation (4.25) independently for each wavelength.

**Index of refraction.** To calibrate for the unknown material’s index of refraction, we use a set of additional measurements with backlighting such that $\theta_0 = \theta_b$ (corresponding to direct observation of the source in the absence of a medium). By measuring the shift in the location of the point-spread-function peak caused by refraction in these images, we can easily solve for the material index of refraction at each of the three wavelengths we use. We discuss this process in more detail in the appendix. We have found our measurement procedure to be adequately accurate for our purposes, but if necessary more accurate measurements of the material index of refraction can be obtained using a refractometer. Additionally, in experiments on synthetic data, we found that small perturbations of the index of refraction ($\pm 0.1$) did not affect recovered scattering parameters considerably.

### 4.4 Experiments

We now demonstrate and validate our approach for acquiring homogeneous scattering parameters. We begin with evaluations on synthetic data aimed at understanding the characteristics of our optimization problem. We then show results on two sets of measured materials. The first is a “validation set” of carefully-constructed nano-dispersions whose scattering parameters can be computed using Lorenz-Mie theory; this set provide a means for quantitative validation. The second set consists of everyday materials that are evaluated by their ability to produce accurate rendered images for novel geometries.
4.4.1 Capture and Computation Time

We first provide some quantitative information for the acquisition and inversion stages of our measurement pipeline. At the acquisition stage, as described in Section 4.3, for a single material we take measurements at three wavelengths and a set of 18 different scene configurations, for a total of 54 measurements. Each of these measurements is a high-dynamic range (HDR) image, composited from low-dynamic range images captured at 19 different exposures. In addition, for every material we measure, we capture a set of low-dynamic range calibration images. This process results in a total capture time of approximately 75 minutes per material. We provide more details about the calibration and high-dynamic range imaging procedures in the Appendix.

At the inversion stage, we solve the optimization problem of Section 4.2 on Amazon EC2 clusters of 100 nodes, with 32 computational cores and at least 20 GB of memory per node (required for the caching of intermediate light fields, as described in Section 4.2.2). We use the nodes to distribute the outer loop of Algorithm 2, that is, the gradient computation for each dictionary atom (for a dictionary of \( N = 200 \) atoms, each node is responsible for two atoms). The results are accumulated at a single master node, which then performs the gradient step of Algorithm 5, and the process is repeated for the number of iterations required until convergence is achieved. We found that processing one set of measurements requires approximately 200 iterations of the SGD algorithm. Overall, fitting one wavelength for a single material requires three to six hours, depending on the density of the material.

We use our own C/C++ implementation, which we have optimized through experiments on synthetic data. However, computation could be reduced by further fine-tuning the various parameters involved, such as dictionary, camera spatial resolution, number of iterations, number of samples per rendering, and so on.
4.4.2 Experiments with Synthetic Data

The optimization problem of Section 4.2 involves the inversion of a random walk process that includes multiple scattering events and is highly non-linear. Despite this, we almost always see in our experiments convergence to a solution that explains the measured data very well. This suggests that the error surface is fairly smooth. To provide more insight, we conduct a series of simulation experiments in which input image-sets are generated using small, artificial three-element dictionaries. Since the three mixing weights are constrained to a simplex, the set of phase functions spanned by three elements is a 2D space, allowing the entire cost surface to be visualized. For these experiments, we parameterize the 2D phase function space by the weights on the first two atoms \((\pi_1, \pi_2)\), and in each experiment we choose a “ground truth” phase function \((\pi_1^*, \pi_2^*)\) and compute for each \((\pi_1, \pi_2)\) the \(L^2\)-difference between input images rendered with that phase function \(\{I^n(\pi_1, \pi_2)\}_{n=1...N}\) and those rendered with the true one \(\{I^n(\pi_1^*, \pi_2^*)\}_{n=1...N}\).

Results from six representative experiments are shown in Figure 4.5. Each row shows three
Figure 4.6: Accuracy in the recovery of material parameters. The plots show recovery errors for albedo, extinction coefficient, and mean error for phase function, as a function of different values of albedo $a = \sigma_s/\sigma_t$ and mean free path $d = 1/\sigma_t$.

separate experiments in which the true phase function is the same while the optical density $\sigma_t$ differs. We find that the error surface has a clear minimum at the true value in all of these 2D experiments, and while an exact proof remains a subject for future research, the cost function appears to be very smooth and without spurious local minima, at least for these 2D problems.

In the next experiment with synthetic data, we compare accuracy on absorbing materials versus scattering materials, and on materials with varying optical densities. We consider a large set of artificial materials that are combinations of: (i) $\sigma_t$ values sampled logarithmically in the interval $[0.01, \ldots, 200 \text{ mm}^{-1}]$, for a total of 21 values; (ii) $\sigma_s$ values corresponding, for each $\sigma_t$, to 21 albedo values, linearly sampled between $a = 0$ (purely absorptive) and $a = 1$ (purely scattering); and (iii) a set of eight different phase functions spanning a wide range of shapes. For each artificial material we render synthetic images using geometry that matches our setup (Section 4.3) with a sample width of $w = 1 \text{ mm}$. Sensor noise is an important consideration for this analysis, so we simulate image noise using photon (Poisson) noise with the parameters reported for two different commercial DSLR cameras [56] (which is very large relative to the Monte Carlo rendering error). The noisy images are input to our optimization algorithm, and we measure error between the recovered parameters and the true ones. We use a tent dictionary with $M = 200$ atoms.

Figure 4.6 provides a summary of these experiments, by visualizing separately the relative
error between the estimated and true values of (left to right): albedo $\alpha = \sigma_s / \sigma_t$, extinction coefficient $\sigma_t$, and phase function $f_p(\theta)$. Each point in these tables corresponds to the percent error—averaged over the eight true phase function shapes—for distinct values of true albedo (horizontal axis) and extinction coefficient (vertical axis). These tables reveal which types of materials we can expect to measure accurately with our setup. Traveling from left to right in these tables makes a gradual transition from purely absorbing materials to purely scattering ones. Traveling from bottom to top moves through materials of increasing optical density, with the top being materials whose mean free path is two hundred times smaller than the sample width, $d_m = 1 / \sigma_t = w / 200$, and the bottom being materials whose mean free path is one hundred times larger than the sample width $d_m = 1 / \sigma_t = 100w$.

The first observation—based on the large, low-error regions in the center of the tables—is that estimation is accurate for a wide range of optical densities. This is a useful fact because it means the width of the glass cell need not be chosen with excessive care. We expect very accurate results as long as the sample width is within an order of magnitude of the material’s mean free path, and we expect graceful degradation when the width extends beyond this in either direction. For extremely optically-thin materials (lower rows in table), scattering events become very rare, and images are dominated by noise. For extremely optically-thick materials (top rows), the diffusion approximation [75] becomes applicable, and recovering both the phase function and the scattering coefficient becomes ill-posed. In practice, we simply choose the width for each material sample from a discrete set of available glass cells (1, 2.5, 5, 10 mm) so that they look reasonably translucent under natural light; see examples in Figure 4.7.

Errors induced by extreme optical thinness and thickness at the top and bottom of these tables should be interpreted differently. If a material is excessively thin at sample width $w$, it is relatively easy to instead use a glass cell that is larger. This is less true for materials that are excessively dense, however, since it is physically challenging to cast materials into glass cells that are too small ($w < 1$ mm). Thus, in cases of extreme optical thickness, our setup
will not provide material parameters that can accurately predict appearance on arbitrary geometries, but only for novel geometries at least as wide as the measured sample.

As expected, we also observe large errors in the estimated phase function when materials are extremely absorptive (left column of third table in Figure 4.6). These errors are somewhat of a computational artifact and have a limited impact on visual appearance. They occur because the appearance of these materials is dominated by attenuation due to absorption, so very little scattering is observed and there is little discernible information about the shape of the phase function. These errors do not impact our ability to predict material appearance, however, because the phase function makes little difference. Indeed, for purely absorptive materials (left-most column) there is no scattering at all, and the phase function can be defined arbitrarily without having any effect on appearance.
4.4.3 Validation Materials

It is common in graphics to evaluate measured scattering parameters by demonstrating renderings of visually plausible results. This is an important benchmark, but it does not directly assess the accuracy of the recovered physical parameters. Since our goal is to produce parameters that are faithful to the true mean free path lengths and phase functions in an absolute sense, being able to directly validate the scattering parameters is crucial.

To achieve comparison to “ground truth” parameters, we capture liquid materials whose exact physical structure are known, similar to materials that are used to calibrate instruments used in a variety of domains for particle sizing or estimating molecular weight [76, 119]. They are created by dispersing nano-scale spherical particles of known chemical composition into a homogeneous embedding medium of a different refractive index, using procedures that allow for very precise control of particle concentration, particle size distribution, and homogeneity. Given these parameters, Lorenz-Mie theory [14, 42] provides analytic expressions of the bulk material scattering parameters \( \{\sigma_t, \sigma_s, f_p(\theta)\} \) at any wavelength.

We measure nanodispersions of two types. First, we measure dispersions of polystyrene spheres in water that are almost monodisperse (single particle size) and are precise enough to be traceable to NIST Standard Reference Materials. We measure three such dispersions\(^2\), each having a 1% \((w/v)\) concentration of particles at a different particle radius: 200, 500, or 800 nm. Second, we measure a spherical polydispersion of aluminum oxide particles \((\text{Al}_2\text{O}_3)\) in water\(^3\), with an approximately known particle size distribution in the range 20 – 300 nm and mean radius of 30 nm. We use glass cells of width \(w = 1 \text{ mm}\) for all of these measurements, and instead of estimating the indices of refraction from image data, we use those predicted by Lorenz-Mie theory.

The results of our measurements are shown in Table 4.1. In all cases, the error in the

\(^2\)Nanobead NIST Traceable Particle Size Standards, Polysciences, Inc.

\(^3\)NanoArc Aluminum Oxide, Nanophase Technologies Corporation.
Table 4.1: Measurements of validation materials (controlled nano-dispersions). Values for $\sigma_s$ are reported in (mm$^{-1}$). Phase function error is given as $L_2$ difference normalized by the $L_2$-norm of the reference phase function. All four validation materials have negligible absorption, resulting in both the predicted and measured values for $\sigma_t$ to agree with those we report for $\sigma_s$ to the third decimal.

<table>
<thead>
<tr>
<th>dispersion</th>
<th>$\sigma_s$ predicted</th>
<th>$\sigma_t$ measured</th>
<th>$\sigma_t$ error (%)</th>
<th>$f_p(0)$ error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>G</td>
<td>B</td>
<td>R</td>
</tr>
<tr>
<td>polystyrene, 200 nm</td>
<td>17.220</td>
<td>28.363</td>
<td>36.317</td>
<td>17.078</td>
</tr>
<tr>
<td>polystyrene, 500 nm</td>
<td>59.082</td>
<td>79.557</td>
<td>88.626</td>
<td>58.431</td>
</tr>
<tr>
<td>polystyrene, 800 nm</td>
<td>65.757</td>
<td>70.438</td>
<td>68.589</td>
<td>66.976</td>
</tr>
<tr>
<td>Al$_2$O$_3$, 30 nm</td>
<td>47.341</td>
<td>93.389</td>
<td>129.870</td>
<td>48.536</td>
</tr>
</tbody>
</table>

recovered parameters is less than 5%. (Note that these materials are purely scattering, so $\sigma_t = \sigma_s$.) The largest error occurs for the aluminum oxide material, for which the particle size distribution is known much less precisely. Figure 4.8-left compares the green-channel phase functions recovered by our optimization (purple curves) to the ground truth phase functions predicted by Lorenz-Mie theory (dotted orange curves). We see that the matches are extremely close. As a reference, we compare both to Henyey-Greenstein phase functions; as the single parameter $g$ of an HG phase function is equal to its average cosine, we plot (green curves) the HG phase function that have $g$ values that are equal to the average cosine of the ground truth phase function. We note that their shapes deviate significantly from the ground-truth. This deviation is important for appearance, particularly for objects that have thin geometry with low-order scattering, where the phase function plays an important role visually. The middle columns demonstrate this by showing captured and fit (pseudocolored) images of the materials under frontal laser illumination at a new angle (which was not used in optimization). The rightmost column shows cross-sections of the image intensities. The deviations of the HG fits from ground-truth lead to discernible differences between the images.

These experiments highlight the fact that simple, single-parameter phase function models can be insufficient for modeling the appearance of scattering materials, and it justifies our choice to fit higher-dimensional phase function models.
Figure 4.8: Measurements of validation materials (controlled nano-dispersions). Left: For each material, we show for the green wavelength the theoretically predicted (dashed orange) and recovered (purple) phase functions, as well as the best Henyey-Greenstein (green) phase function fit. The recovered phase functions are in close agreement with the correct ones and the purple and orange curves tightly overlap. As another visualization, we show the images for a novel configuration: under frontal collimated laser illumination ($\theta_f = 25^\circ, \theta_o = 0^\circ$). We compare our phase function and the best Henyey-Greenstein fit (images are color-mapped for better visualization). The rightmost column shows a crosssection through the captured and re-rendered images for this configuration.

4.4.4 Other Materials

Next, we use our acquisition setup and optimization algorithm to measure several common materials. They can be grouped roughly into three categories:

- Highly scattering liquids of varying viscosities; including mustard, shampoo, hand cream, liquid designer clay, and different types of milk.
Table 4.2: Scattering parameters of materials measured using our proposed acquisition setup and inversion algorithm. Values for $c_s$ and $c_a$ are reported in (mm$^{-1}$). The average cosine of the measured phase functions is reported, while the entire phase functions are shown in Figure 4.9. Fitting and generalization errors are given as % $L_2$ difference normalized by the $L_2$-norm of the reference image, averaged across the fitting and novel captured images respectively.

By “eyeballing” each sample under natural light, we choose glass cell widths so that each sample looks reasonably translucent under ambient lighting. The results we report were captured using width $w = 1$ mm for materials in the first and third categories, except for glycerine soap; and $w = 10$ mm for the second category and glycerine soap. Photographs of samples in 1 mm, 2.5 mm, and 10 mm cells are shown in Figure 4.7. For each sample, we estimate the index of refraction as described in Section 4.3, and these range from values of 1.33 (for milk, reduced milk, milk soap, and the water soluble liquids) to 1.47 (for olive oil and glycerine).

The measured parameters are shown in Table 4.2. We quantitatively evaluate the quality of the recovered scattering parameters in two ways. First, we report the fitting error, which is the average $L_2$ image difference between input images and the corresponding images rendered with the recovered parameters, normalized by the $L_2$-norms of the input
images. Second, we compute a measure of generalization error by: i) using the recovered parameters to render laser-illumination images with different sample widths and lighting directions; and ii) comparing these simulated images to captured photographs in these same novel configurations. For the novel configurations, we use lighting angles $\theta_f, \theta_b \in \{10^\circ, 20^\circ, 30^\circ\}$ and glass cells with widths $w \in \{2.5, 5\text{ mm}\}$. The generalization error for each material is reported as the average relative $L^2$ image difference over the set of all novel configurations. As shown in the right two columns of Table 4.2, fitting errors are less than 4% and generalization errors below 5%.

Figure 4.9 shows the measured phase functions, each superimposed with an HG phase functions whose $g$-value is equal to the average cosine of the phase function we measure. Some of these phase functions are well approximated by the HG model but others, including hand cream, liquid clay, and mixed soap, are not. This set of tabulated phase functions is available at the project website.
4.4.5 Rendering of Measured Materials

As qualitative evaluation, Figure 4.10 shows an image rendered with our recovered material parameters under natural lighting. From left to right, are milk soap and glycerine soap (top and bottom, respectively), olive oil, blue curacao, and reduced milk. The soap geometry corresponds to scanned molded cubes made of the corresponding materials. We see that the recovered material parameters successfully reproduce the color variations that are critical to the translucent appearance of these materials. This is most notable in the glycerine soap, where blue wavelengths scatter first and cause a reddish glow in the middle of the object, but it is also visible on the left edge of the milk soap and the top-right corner of the milk. These effects are better observed in the bottom image of Figure 4.10, where all RGB vectors are normalized to unit norm to highlight the color variations in the various materials.

4.5 Discussion

This chapter has presented an optimization framework for solving the inverse scattering appearance matching problem (3.25) for homogeneous scattering materials. The framework handles multiple scattering in a physically accurate way, and does not require precise isolation of single scattering or imaging at the diffusive regime. This enables using the framework for the measurement of a broader set of materials, including turbid solids and liquids. The optimization also incorporates a large material dictionary and thereby avoids the restrictions of low-parameter phase function models. Our analysis and experiments show that we can recover accurate physical scattering parameters for a variety of materials.

While we proposed one possible scanning configuration, our optimization could be used to infer scattering parameters from images captured from a variety of scene geometries and incident light fields. The only requirement is that both lighting and geometry be precisely calibrated. Our setup combines the benefits of high-frequency angular lighting (for stable
Figure 4.10: Top: Rendering of materials in natural illumination using our acquired material parameter values. Bottom: Same image with RGB vectors normalized to unit norm, to visualize color variations.
optimization) and precise, stable calibration (for repeatability), but it limits measurements to three wavelengths and to solids that can be cast into glass cells of thickness within an order of magnitude of the mean free path. In principle, our optimization could be applied to images of more general solid objects, but this would require enhancing our setup to also recover the object shape and its surface microstructure (BSDF).

In addition, while we restricted ourselves to steady-state images, combinations of our optimization framework with more sophisticated imaging configurations could improve the optimization’s stability and convergence rate. In particular, the steady-state measurements we use have the form of Equation (4.18),

$$S(T, W) = D(I - K)_{-1} L_i = D \left( \sum_{n=1}^{\infty} K^n \right) L_i. \quad (4.32)$$

Instead of aggregating over applications of $K$ of all orders, we could imagine collecting different orders as separate measurements of the form

$$\{DK^n L_i, n = 0, \ldots, \infty \}. \quad (4.33)$$

These corresponds to the transient imaging measurements described in Section 3.5. In Chapter 5, we derive a different inverse scattering algorithm that uses such measurements of that form to do heterogeneous inverse scattering, whereas in Chapter 6, we describe an imaging system for collecting such measurements.
Chapter 5

Heterogeneous Inverse Scattering

In this chapter, we consider again the appearance matching objective of Equation 3.25, but we do away with the two assumptions that we used in Chapter 4: We assume we can controllably inject and measure light at the boundary ∂M of a medium M, using both steady-state and transient imaging. From these measurements, we seek to infer the heterogeneous scattering material parameters anywhere inside the medium.

The heterogeneous inverse scattering problem is considerably more challenging than the homogeneous one, for two reasons. The first is the explosion in the dimensionality of the problem: instead of a single set of material parameters k, we now need to infer a function over the volume, k(x), x ∈ M. The second is that, when we limit ourselves to non-invasive, non-destructive means of measurement, like imaging at the boundary, we do not have direct access to the internal parts of the material. The only measurements of the interior that we have are confounded by additional interactions with the boundary.

In order to overcome these challenges, we can take advantage of more general and numerous measurements. These can come either from considering many more geometry, viewpoint, and illumination conditions, or by using transient imaging in addition to steady-state imaging.
Towards this direction, in this chapter, we develop an algorithm for efficiently solving Equation 3.25 for heterogeneous materials, while using both steady-state and transient imaging measurements as input. As in Chapter 4, our algorithm is based on Monte-Carlo rendering and stochastic gradient descent, but uses a path-space formulation to be able to accommodate more general material spaces and measurement types. A more detailed comparison of the two algorithms is given in the appendix. After presenting our algorithm, we use simulations to evaluate the utility of different computational imaging configurations for heterogeneous inverse scattering problems.

5.1 Path-space Formulation of Volumetric Light Transport

Chapter 3 describes entries of the light transport matrix $T$ and measurements $S(T, W)$ as different accumulations of photon contributions based on their paths. This intuition has been formalized in computer graphics and is the foundation of path-based rendering algorithms. We can also use it to derive our inverse rendering algorithm. For notation, we define a path $\bar{x}$ as an ordered sequence of points in the medium $\mathcal{M}$,

$$\bar{x} = x_0 \to x_1 \to \ldots \to x_B,$$

(5.1)

for any finite integer $B > 1$, where $x_1$ and $x_B$ are on the boundary $\partial \mathcal{M}$, and all other points are in the medium $\mathcal{M}$. We denote the space of all such paths as $\mathcal{P}$. Using notation introduced previously, for each path segment, $\omega(x_b \to x_{b+1})$ is the direction of the line interval starting at $x_b$ and ending at $x_{b+1}$. Then, for each path $\bar{x}$, we denote by $o(\bar{x}) = x_0$ and $e(\bar{x}) = x_B$ its origin and end respectively, by $\omega_o(\bar{x}) = \omega(x_0 \to x_1)$ and $\omega_e(\bar{x}) = -\omega(x_{B-1} \to x_B)$ its starting and ending directions respectively, and by $\tau(\bar{x}) = \sum_{b=1}^{B} \|x_b - x_{b-1}\|_2$ its length.

Then, based on the path formulation of light transport [142, 115], every measurement of the
light transport matrix can be written as,

\[ S(T_p, W) = \int_{\mathcal{P}} W(\bar{x}) \bar{f}_k(\bar{x}) \, d\bar{x}, \]  

(5.2)

where we overload notation to make the sampling function of Equation (3.20) apply to paths, based on their endpoints and length,

\[ W(\bar{x}) \triangleq W(\{e(\bar{x}), \omega_e(\bar{x})\}, \{o(\bar{x}), \omega_o(\bar{x})\}, \tau(\bar{x})). \]  

(5.3)

The material-dependent throughput function \( \bar{f}_k \) determines the path’s radiance contribution,

\[ \bar{f}_k(\bar{x}) = \prod_{b=1}^{B-1} f_k(x_{b-1} \rightarrow x_b \rightarrow x_{b+1}), \]  

(5.4)

\[ f_k(x_{b-1} \rightarrow x_b \rightarrow x_{b+1}) = a(x_{b-1} \rightarrow x_b) \varrho(x_{b-1} \rightarrow x_b \rightarrow x_{b+1}), \]  

(5.5)

\[ a(x_{b-1} \rightarrow x_b) = a(x_{b-1}, x_b), \]  

(5.6)

\[ \varrho(x_{b-1} \rightarrow x_b \rightarrow x_{b+1}) = \begin{cases} f_s(x_{b-1} \rightarrow x_b \rightarrow x_{b+1}), & x_b \in \partial \mathcal{M}, \\ \varrho_s(x_b) f_p(x_{b-1}, \omega) (x_{b-1} \rightarrow x_b) \cdot \omega (x_b \rightarrow x_{b+1}), & \text{otherwise}. \end{cases} \]  

(5.7)

Equation (5.6) corresponds to volumetric attenuation along each segment of a path, defined earlier. Equation (5.7) corresponds to radiance transfer as direction changes at the end of a path segment, where scattering events occur. When \( x_b \in \partial \mathcal{M} \), the change in direction is due to a surface event at the medium boundary, and the amount of radiance transferred is determined by the material’s BSDF. Otherwise, the change in direction is due to a volume event, and radiance transfer is determined using the local phase function.

For a detailed derivation of the path-space formulation of light transport, we refer to [142, 115, 73]. Intuitively, it can be understood as follows: The volume rendering equation (3.14) provides a way for evaluating points of the light transport matrix \( T_p \), through an analytic expression for evaluations of the radiance at point-direction pairs inside \( \mathcal{M} \). However, this expression is a function of the radiance itself, evaluated at different point-direction pairs. The path-space formulation follows by recursively expanding this expression in...
Equation (3.20), until we reach a point on the inwards boundary $\Gamma_i$, where radiance is known.

### 5.2 Appearance Matching for Heterogeneous Inverse Scattering

We will now derive an efficient optimization algorithm for solving the appearance matching problem of Equation (3.25) for heterogeneous materials. We restate here the problem for convenience.

$$
\min_{n, k} \sum_{n=1}^{N} \frac{1}{2} (I^n - S(T_{\pi}, W^n))^2,
$$

where each $I^n$ is a calibrated measurement with known corresponding sampling function $W^n$, $\pi$ is an $M$-dimensional parameterization of the material parameters $k$, and $A$ is the space of admissible materials. The derivation of the algorithm mirrors that of Section 4.2 for the homogeneous case. However, instead of the operator formulation we used there, our derivation will use the path-space formulation of Section 5.1. A comparative evaluation of the two approaches can be found in the appendix.

#### 5.2.1 Differentiation of the light transport matrix

Using Equation (5.2), we will now formulate a path-space expression for the derivative of samplings of the pathlength-resolved light transport matrix with respect to any material parameter $\pi_m$,

$$
\frac{\partial S(T_{\pi}, W)}{\partial \pi_m} \bigg|_{\pi=\pi_o} = \int_{P} W(\bar{x}) \frac{\partial \hat{f}_k(\bar{x})}{\partial \pi_m} \bigg|_{\pi=\pi_o} d\bar{x},
$$

where we assume the necessary regularity conditions on $\hat{f}_k$ to allow exchanging the order of differentiation and integration. Given the product form of $\hat{f}_k$ in Equation (5.4), by applying
the chain rule we have that

$$\frac{\partial f_k (\bar{x})}{\partial \pi_m} \bigg|_{\pi = \pi_o} = \tilde{f}_k (\bar{x}) S_{k,m} (\bar{x}),$$  \hspace{1cm} (5.10)

$$S_{k,m} (\bar{x}) \triangleq \sum_{b=1}^{B-1} \bar{S}_{k,m} (x_{b-1} \rightarrow x_b \rightarrow x_{b+1}),$$ \hspace{1cm} (5.11)

$$S_{k,m} (x_{b-1} \rightarrow x_b \rightarrow x_{b+1}) \triangleq \frac{(\partial f_k (x_{b-1} \rightarrow x_b \rightarrow x_{b+1}) / \partial \pi_m)_{\pi = \pi_o}}{f_k (x_{b-1} \rightarrow x_b \rightarrow x_{b+1})}.$$ \hspace{1cm} (5.12)

In the statistics literature, $S_{k,m}$ is often called the score function of $f_k$ with respect to $\pi_m$.

Substituting Equation (5.10) into Equation (5.9), we obtain

$$\frac{\partial \mathcal{S} (\mathcal{T}_{\pi}, W)}{\partial \pi_m} \bigg|_{\pi = \pi_o} = \int_{\mathcal{P}} W (\bar{x}) \tilde{f}_k (\bar{x}) S_{k,m} (\bar{x}) \, d\bar{x}. \hspace{1cm} (5.13)$$

### 5.2.2 Monte Carlo integration

The integrals of Equation (5.2) and Equation (5.13) can be estimated using Monte Carlo integration: we first use any probability distribution $\mu$ on $\mathcal{P}$ to sample a set of paths $\{ \bar{x}_j, \ j = 1, \ldots, J \}$; then we form the respective unbiased estimates

$$I = \sum_{j=1}^{J} \frac{W (\bar{x}_j) \tilde{f}_k (\bar{x}_j)}{\mu (\bar{x}_j)}, \hspace{1cm} (5.14)$$

$$G_m = \sum_{j=1}^{J} \frac{W (\bar{x}_j) \tilde{f}_k (\bar{x}_j) S_{k,m} (\bar{x}_j)}{\mu (\bar{x}_j)}.$$ \hspace{1cm} (5.15)

The paths $\{ \bar{x}_j, \ j = 1, \ldots, J \}$ can be chosen using the efficient sampling strategies developed for physically accurate rendering. For instance, the volumetric particle tracing Algorithm 3 used in Chapter 4 corresponds to sampling paths by recursively importance sampling the terms $f_k$ of Equation 5.5, starting from the emittance $W_i$. Similar algorithms include volumetric path tracing, where sampling starts from the importance $W_o$; bidirectional path tracing, where we combine paths traced from both the importance and emittance; and Metropolis Light Transport, where paths are sampled using Markov Chain Monte Carlo methods. For a review, we refer to [142, 115]. Because of the need to render individual
elements of the light transport matrix (corresponding to the infinitesimal emittance and importance functions of Equations (3.23) and (3.24) respectively), for our application we use bidirectional path tracing. When rendering elements of the pathlength-resolved light transport matrix, we also use the local sampling modifications proposed by [73].

Finally, most rendering algorithms use the product form of Equation (5.4) to efficiently compute the term \( \bar{f}_k \) in Equation (5.14) recursively while tracing a path. The same can be done for the term \( \bar{S}_{k,m} \) in Equation (5.15), using the sum form of Equation (5.11). The expression (5.15) allows us to engineer any Monte Carlo rendering algorithm to evaluate, in addition to radiance, values of the gradient.

### 5.2.3 Stochastic Optimization

We now turn our attention to the appearance matching optimization problem of Equation (5.8). Denoting by \( E(\pi) \) its loss function, we have for its gradient

\[
\frac{\partial E}{\partial \pi_m} \bigg|_{\pi} = \sum_{n=1}^{N} \left( I^n - S(T_{\pi}, W^n) \right) \frac{\partial S(T_{\pi}, W^n)}{\partial \pi_m} \bigg|_{\pi}.
\]

Therefore, we can form an estimate \( g \) of this gradient using the stochastic estimates of Equations (5.14) and (5.15),

\[
g_m(\pi) = \sum_{n=1}^{N} \left( I^n - I^n \right) G^n_m.
\]

Similar to Section 5.2.3, the fact that we can efficiently compute stochastic estimates \( g \) of the gradient motivates us to use stochastic gradient descent algorithms to solve the appearance matching optimization problem of Equation (5.8). However, we need to make two considerations. First, in order for the gradient estimates \( g \) from Equation (5.17) to be unbiased, that is

\[
\mathbb{E} [g_m(\pi)] = \frac{\partial E}{\partial \pi_m} \bigg|_{\pi},
\]

the estimates \( I^n \) and \( G^n_k \) must be statistically independent. This can be achieved by rendering the two terms separately, that is, using independently selected sets of paths for each

\[102\]
rendering operation.

Second, the standard SGD iteration,

\[ \pi^{(t+1)} = \pi^{(t)} - \beta^{(t)} g^{(t)}, \]  

uses a common step size for all unknowns \( \pi \), typically chosen equal to \( \beta^{(t)} = \frac{\gamma}{\sqrt{t}} \). While theoretical convergence is robust to the proportionality constant \( \gamma \), using a common step-size for all parameters can be problematic when different parameters have gradients of very different magnitudes or if gradient vectors are very sparse: a parameter may be receiving zero gradients for several iterations, during which the step-size will continue decreasing; once a large-magnitude gradient is encountered, the parameter update will be very small due to the diminished \( \beta^{(t)} \). This second issue is typically encountered when solving the appearance matching problem of Equation (5.8) for heterogeneous scattering materials: the gradient estimate \( g_m \) for some material voxel will be zero if, during the rendering operations of Equations (5.15) and (5.14), no sampled paths travel through the voxel.

To ameliorate both the sensitivity and sparsity issues, a number of SGD variants have been developed that use separate, per-parameter step sizes \( \beta_{m}^{(t)} \), each decrementing adaptively based on the magnitudes of the per-parameter gradients in previous iterations. We experimented with the algorithms of [85, 154, 33], and chose to use ADADELTA, which we found empirically to perform the best among them and standard SGD for solving (5.8). We show the exact form of the iterations in Algorithm 6. In Section 5.4, we discuss initialization details.

### 5.3 Imaging Design

The inverse scattering algorithm of Section 5.2 can accommodate all the different types of imaging described in Section 3.5, by using different functions \( W \) in Equations (5.2) and (5.13). Given the range of imaging options available, we discuss theoretical results and empirical
Algorithm 6 ADADELTA.

Require: decay rate $\gamma < 1$, constant vector $\epsilon$.
1: $\pi^{(0)} \leftarrow \text{Initialize }$, $\sigma_{g}^{(0)} \leftarrow 0$, $\sigma_{\delta}^{(0)} \leftarrow \epsilon$.
2: while not converged do
3: $g^{(t)} \leftarrow \text{ComputeGradient } (\pi^{(t)})$.
4: $\sigma_{g,k}^{(t)} \leftarrow \gamma \sigma_{g,k}^{(t-1)} + (1 - \gamma) \left(\delta_{k}^{(t)}\right)^{2}$.
5: $\delta_{k}^{(t)} \leftarrow \frac{\sigma_{g,k}^{(t-1)}}{e_{\delta,k}^{(t)}}$.
6: $\pi^{(t+1)} \leftarrow \pi^{(t)} - \delta^{(t)}$.
7: $\sigma_{\delta,k}^{(t)} \leftarrow \gamma \sigma_{\delta,k}^{(t-1)} + (1 - \gamma) \left(\delta_{k}^{(t)}\right)^{2}$.
8: end while
9: return $\pi^{\text{opt}} = \frac{1}{T} \sum_{t=0}^{T} \pi^{(t)}$.

Figure 5.1: Inverse rendering algorithm. Left: We use the ADADELTA variant of stochastic gradient descent to minimize the appearance matching objective of Equation (5.8). Right: We use a modified Monte Carlo rendering algorithm to compute stochastic gradient estimates. When shading a path, we compute for each segment both the usual throughput terms $f_{m}$ and the score terms $S_{m,k}$. These are accumulated to determine the path’s contribution to the image $I$ and its derivative $G_{k}$ (Equations (5.4) and (5.11)). We repeat this process over many paths that are sampled to satisfy the sampling function $W$ corresponding to the input image measurement (Equation (5.3)).

observations that can be used to guide the design of imaging systems for inverse scattering applications.

5.3.1 Local Ambiguities

In steady-state imaging, there exist similarity relations which, under certain conditions, allow changing the scattering parameters at a point $x \in \mathcal{M}$ without changing the radiance at that point [150, 156]. We prove that these similarity relations also hold for pathlength-decomposed measurements.

Claim 3. Let $\{a_{n,l}(x, \tau), n > 0, -n \leq l \leq n\}$ be the coefficients of the spherical-harmonics expansion of the solution $L(x, \omega, \tau)$ of Equation (3.1) at some point $x \in \mathcal{M}$, and $\{f_{p,n}(x), n > 0\}$ the coefficients of the Legendre expansion of $f_{p}$ at that point. If there exists $N > 0$ such that $a_{n,l}(x, \tau) = 0$ for all $n > N$, then two materials $m, m^{*}$ will produce equal values $L(x, \omega, \tau)$ if, for
\[ n \leq 1 \leq N, \]

\[ \sigma_{\alpha}(x) = \sigma_{\alpha}^*(x), \quad (5.20) \]
\[ \sigma_{s}(x) \left(1 - f_{p,n}(x)\right) = \sigma_{s}^*(x) \left(1 - f_{p,n}^*(x)\right), \quad (5.21) \]

We prove this claim in the appendix. These similarity relations are local, as they describe ambiguities at one point \( x \in \mathcal{M} \). To reduce these ambiguities, Claim 3 suggests maximizing the angular frequency of \( L(x,\omega,\tau) \) throughout the medium. For a given shape of \( \mathcal{M} \), the only way to control this is through the incident illumination \( W_i(X_i) \), which provides a lower bound to the angular frequency of \( L(x,\omega,\tau) \) by way of spherical convolution with the phase function in the RTE (3.1) acting as a low-pass filter. The best light sources will have high angular and spatial frequencies, such as the ideal sources of Equation (3.23), or collimated area sources with high-frequency spatial profiles. These are analogous to the arguments we made in Section 4.3 for selecting point sources for the homogeneous inverse scattering problem.

### 5.3.2 Non-local Ambiguities

On top of local ambiguities, a heterogeneous medium can exhibit non-local ambiguities involving material parameters at different points \( x \). In the following, we argue that path-length decomposition combined with a specific input-output geometry can help reduce such non-local ambiguities. For this, we consider the scene in Figure 5.2(a): A cubic volume is discretized into \( h \)-sized voxels indexed by coordinates \([d, p]\). We call the coordinate \( d \) the depth of a voxel, and we call the set of voxels with the same \( d \) the layer at depth \( d \). We assume that material parameters \( m[d, p] \) are constant within each voxel, and that \( h \) is also the spatial and pathlength resolution at which we can image.

We use an ideal source (Equation (3.23)) co-located with an ideal sensor (Equation (3.24)). Then, the following claim states that, assuming all material parameters at layers \([h, 2h, \ldots, (n - 1)h]\)
Figure 5.2: (a) Configuration used for Claim 4. (b) Frontlighting allows decomposing the inverse scattering problem into sub-problems for depth-wise layers, where at each layer the material parameters of lower-depth layers are known. (c) Backlighting allows measuring deep layers at shorter pathlengths $\tau$, and thus higher SNR, than frontlighting.

are known from previous measurements, pathlength decomposed measurements at $\tau = 2nh$ and $\tau = (2n + 1)h$ provide linear equations in the parameters of layer $d = nh$.

Claim 4. Using pathlength decomposition, the configuration of Figure 5.2(a) provides measurements of the form

$$I_\tau = \begin{cases} 
Q_\tau + \sum_{p \in k(n)} \sigma_s [nh, p] \int_0^\pi f_p ([nh, p], \theta) R_{\tau,p}(\theta) \, d\theta, & \tau = 2nh, \\
S_\tau + \sum_{p \in k(n)} \sigma_t [nh, p] T_{\tau,p}, & \tau = (2n + 1)h,
\end{cases} \quad (5.22)$$

where $k(n)$ is the subset of voxels in layer $d = nh$ that are intersected by a source-centered circle of radius $nh$; $Q_\tau$ and $R_{\tau,p}(\theta)$ are functions of material parameters $\{m[d,p], d < nh\}$; and $S_\tau$ and $T_{\tau,p}$ of $\{m[d,p], d < nh; \sigma_s [nh, p]; f_p ([nh, p], \theta)\}$.

We prove this claim in the appendix. Using Figure 5.2(a), we can understand the first line of Equation (5.22) as follows. Measurements of the form $I_{2nh}$ are sums of contributions from paths of length $\tau = 2nh$. All these paths are contained inside a circle of radius $nh$ centered at the source/sensor location. We can split the paths into two categories: those that stay within layers $[h, 2h, \ldots, (n-1)h]$, and those that reach layer $d = nh$. Most paths fall into the first category, and their contributions are included in the term $Q$. A fraction of the paths are in the second category, and when they reach layer $d = nh$, their interaction is limited to a
single bounce somewhere within a particular subset of the voxels in that layer (denoted \(k(n)\) and shaded orange in Figure 5.2(a)). The parts of these paths that are contained in layers \(d < nh\) correspond to the term \(R\) in Equation (5.22), and their single-bounce interactions within layer \(d = nh\) lead to linear dependence on parameters \(\sigma_{sfp}(\theta)\). Thus, when the materials in layers \(d < nh\) are known, the terms \(Q\) and \(R\) can be computed; and we can build a system of linear equations in the parameters \(\sigma_{sfp}(\theta)\) for all voxels in layer \(d = nh\) by changing the source/sensor location and by shifting the sensor relative to the source.

Similarly, the second line of Equation (5.22) refers to measurements \(I_{(2n+1)h}\) of contributions from paths that have an odd number of steps, \(\tau = (2n + 1)h\). As explained in detail in the appendix, these paths may include up to two scattering events in layer \(d = nh\), but only one that involves \(\sigma_t[d,p]\). Therefore \(I_{(2n+1)h}\) will depend linearly on \(\sigma_t[d,p]\), and because the other parameters \(\sigma_s[d,p]\) and \(f_p([d,p],\theta)\) can be separately estimated from the even-step measurements, we can create a linear system of equations in \(\sigma_t\) for all voxels in layer \(d = nh\) by combining odd-step measurements from different source and sensor locations.

Overall, Claim 4 suggests a recursive, layer-wise procedure for inferring scattering parameters from pathlength-resolved measurements (Figure 5.2(b)): Assuming parameters at layers \(d < nh\) have been estimated from previous measurements, use measurements of the form of \(I_{2nh}\) and \(I_{(2n+1)h}\) to estimate parameters at layer \(d = nh\), through a linear system in \(\sigma_s[d,p]f_p([d,p],\theta)\) and another linear system in \(\sigma_t[d,p]\).

### 5.3.3 Noise Considerations

So far, we have shown that pathlength decomposed measurements in a frontlighting configuration, where there source and sensor are at the same side of the medium, helps reduce non-local ambiguities. To complete the picture, we must also consider the signal-to-noise ratio (SNR) of the measurements in Equation (5.22) as pathlength \(\tau\) increases. One observation is that their magnitude, and therefore SNR, decreases exponentially with \(\tau\) due to
volumetric attenuation (see Equation (5.6)). Another is that, from Figure 5.2(a), the ratio of the circle’s area contained in layers $d < nh$ versus that in layer $d = nh$ increases with $\tau$. This implies that the second terms in Equation (5.22), which contain all information about parameters at depth $d = nh$, become smaller relative to the terms $Q_\tau, S_\tau$, which are independent of those parameters.

As a result of these two factors, the information available for inferring material parameters becomes progressively worse at greater depths. One way to ameliorate this noise is to use a supplementary backlighting configuration (Figure 5.2(c)), where sources and sensors are on opposite sides of the medium. These backlighting measurements cannot be used in the recursive procedure described above, but they can provide cleaner measurements of deeper layers.

While these noise considerations also apply to the case of steady-state measurements from frontlighting and backlighting configurations, the preceding analysis of non-local ambiguities cannot be directly extended to that case. Previous results [8] indicate that, if the steady-state light transport matrix $T^s_{\pi}$ is known at all points $\Gamma_o$ and $\Gamma_i$, then it uniquely characterizes the material $k$ everywhere in $\mathcal{M}$. However, these results do not take into account noise, or deviations from ideal sources and sensors. In place of theoretical analysis, we evaluate the relative utility of steady-state versus pathlength-resolved measurements quantitatively. Specifically, in Section 5.5, we use the algorithm of Section 5.2 to perform inverse scattering experiments with simulated volumes and measurements of different types.

### 5.4 Initialization

Based on the discussion of Section 5.3, we discuss two initialization procedures for Algorithm 6.
5.4.1 Multi-resolution Initialization

In this first approach, we begin by assuming that the unknown volume is homogeneous, that is, that there is a single set of material parameters everywhere. Running Algorithm 6 produces a homogeneous material estimate. We then voxelize the volume at progressively higher material resolutions; at each resolution, we rerun Algorithm 6 to recover material parameters at that resolution, using the algorithm’s output from the previous resolution as initialization. We repeat this process until we reach the highest material voxel resolution we consider.

For the original, homogeneous run of Algorithm 6, we initialize the material parameters to \( g = 0 \) (uniform phase function), \( \sigma_a = 0 \), and \( \sigma_s = 10 \text{ mm}^{-1} \). We use these values to bias the path sampling algorithm towards exploring the entire material volume.

5.4.2 Greedy depth-wise initialization

When using pathlength-decomposed measurements with a frontlighting configuration, as in Figure 5.2(b), Claim 4 and the recursive inference argument discussed in Section 5.3 suggest a different way to do initialization. Specifically, starting at \( n = 1 \) and iteratively at every \( n \), we use Algorithm 6 to process only the measurements at pathlength \( \tau = nh \) and to optimize only for the material parameters at voxels of depth \( d = \tau/2 \) (up to the voxelization resolution). At every iteration, for voxels at lower depths, we fix the material parameters at the values recovered at previous iterations of this recursive procedure. At the end of the recursion, we can use the parameter estimates at every voxel recovered using the above procedure to initialize another run of Algorithm 6, which simultaneously processes all measurements (including measurements from configurations other than frontlighting), and simultaneously optimizes over all material parameters in the volume.

In practice, whenever possible we combine both initialization procedures: At every iteration
of the multi-scale procedure, we do material inference at the corresponding resolution using
the greedy depth-wise recursive procedure. At every iteration of either the multi-resolution
or the greedy depth-wise procedures, we run a few iterations of Algorithm 6, instead of
running it until full convergence. In our experiments, we have found that the combined use
of these two initialization procedures significantly accelerates convergence.

5.5 Experimental Evaluation

In this section, we use the inverse rendering framework of Section 5.2 to perform inverse
scattering experiments on simulated material data and measurements. For this experiments,
we implemented the inverse rendering framework on top of the Mitsuba physically based
renderer [68]. We extended the bidirectional path tracing algorithm to support spatial
probing and pathlength decomposition rendering, for both radiance and gradient estimation
(Equations (5.14) and (5.15), respectively). The stochastic optimization layer allows distribut-
ing the various rendering tasks involved in gradient computation (multiple parameters and
measurements) across a multi-CPU cluster. We ran all our experiments on 20-node Amazon
EC2 clusters, with 36 cores per node.

5.5.1 Comparison of Imaging Configurations

We perform inverse scattering experiments on synthetic volumes, to evaluate the perfor-
mance of different imaging configurations. Following Figure 5.2(a), we use a cubic medium
of size $10 \times 10 \times 10 \text{ mm}^3$ discretized at resolution 0.4 mm, resulting in 15625 voxels. To
reduce the space of possible imaging configurations, we assume that the material parameters
are characterized by a cross-section of the cube along the $\{d, p\}$ plane, remaining constant
across the third dimension, corresponding to $3 \times 625 = 1875$ unknowns. We use an imaging
resolution four times that of the material grid, corresponding to pixel size $0.1 \times 0.1 \text{ mm}^2$
and pathlength resolution 0.1 mm.
We fix the medium refractive index to $\eta = 1.3$ and assume a smooth dielectric BSDF at its boundary. We adopt the single-parameter Henyey-Greenstein model for the phase function [61]; therefore, each material voxel is associated with corresponding unknown values for $\sigma_a, \sigma_s,$ and the phase function parameter $g$ (equal to the phase function’s first moment). We constrain $\sigma_a, \sigma_s \in [1, 10]$ mm$^{-1}$, and $g \in [0, 0.6]$. We generate volumes by modeling each parameter inside the medium as a mixture of two Gaussians of random mean and variance.

We use these volumes to compare three imaging configurations: 1) Pathlength decomposition where, for each ideal source, we measure the radiance exiting the volume at the opposite direction and from the same position, as well as its two spatial neighbors. We take measurements with the ideal source placed at every pixel on the medium boundary, including in sidelighting and backlighting positions. 2) Steady-state imaging where, for every ideal source, we use an orthographic camera to measure radiance exiting from all pixels in one surface of the cube. As before, we take measurements with the ideal source placed at every pixel on the medium boundary, and at every position at three different orientations. 3) Similar to (1), but instead of spatial shifts, we take measurements at multiple source orientations. Each of these configurations produces 120000 measurements, or 64 measurements per unknown. When rendering simulated measurements, we add sensor noise using [56].

In Figure 5.3(a), we visualize the reconstructed parameters for one of the synthetic volumes. We observe that all three configurations are generally able to reconstruct all three spatially varying parameters, $\sigma_a, \sigma_s,$ and $g$, within a mean relative error 7% and maximum relative error 15%, concentrated around areas of high absorption. In Figure 5.3(b), we compare the RMS error in the estimation of each parameter by each configuration, averaged across five synthetic volumes. We see that, the configuration using only steady-state measurements has a lower RMS error. We expect that this is due to the very low SNR of pathlength-decomposition measurements corresponding to large pathlength values. We show in
Figure 5.3: Comparison of different imaging configurations. (a) Visualization of ground-truth and reconstructed material parameters as a function of location. (b) Mean relative reconstruction error for material parameters, averaged over multiple synthetic volumes.

Figure 5.3(b) the mean relative error obtained by a fourth configuration, created by replacing measurements in configuration (1) of magnitude comparable to the additive sensor noise, with steady-state measurements from configuration (2) corresponding to large camera-sensor distances. We see that this combination produces the lowest RMS error among all imaging configurations.

5.5.2 3D Reconstruction

In Figure 5.4, we show volumetric reconstructions of a dense heterogeneous smoke volume, with smoothly spatially-varying scattering parameters, and assuming index of refraction equal to 1. We generate our own parameters for the volume mesh provided by [68]. We use pathlength-resolved measurements in frontlighting, sidelighting, and backlighting configurations.

Figure 5.4(a) compares renderings of the smoke volume using the ground-truth and recovered material parameters under novel imaging configurations (not used as input to the inverse rendering algorithm). In Figure 5.4(b), we compare ground-truth and recovered
Figure 5.4: (a) Renderings of a smoke volume using ground-truth and recovered material parameters under novel viewpoint and illumination conditions. (b) Ground-truth and recovered parameters for a vertical cross-section through the smoke volume.

material parameters across a cross-section of the volume. Our algorithm accurately recovers all scattering parameters, with mean relative error 9.31% and maximum relative error 19.73%, and the recovered parameters can reproduce the appearance of the volume under new imaging conditions.

5.6 Discussion

We have presented an algorithm for solving the appearance matching problem (3.25) for heterogeneous materials. We have used this algorithm to perform a quantitative evaluation of various imaging techniques for the heterogeneous inverse scattering problem. Our
theoretical results provide formal justification for the use of pathlength decomposition in applications requiring volumetric reconstruction of complex materials. Additionally, our experimental results demonstrate that there are multiple imaging configurations, including both steady-state and pathlength decomposition measurements, that enable accurate recovery of heterogeneous scattering parameters. Our results can be used to guide the design of physical acquisition setups, where selecting from among or modifying the various configurations we evaluated can be done by taking into account other considerations, such as hardware availability, exposure time, and geometry constraints.
In Section 3.5, we discussed different ways to sample the Green’s function describing light transport in a scene. Using the path-space formulation of light transport described in Section 5.1, these can be viewed as different ways to decompose contributions of photons into measurements, based on characteristics of the paths they follow in a scene. These can be either the length of the paths, or their origin and endpoint. For this reason, we will refer to such imaging techniques as light transport decompositions.

There is a growing number of methods for measuring such light transport decompositions. These methods find applications in image-based rendering, image editing, and measuring scene shape in the presence of translucency and interreflections, visualizing light-in-flight, as well as “imaging around corners”.

Our qualitative results in Chapter 5 indicate that being able to capture different types of light transport decompositions can also be critical for enabling general, heterogeneous inverse scattering. However, as shown in Section 4.1, for these decompositions to be informative,
we need to be able to capture them at resolutions \( h \) smaller than the mean free path of the material we are imaging. Additionally, as shown in Section 5.3, the resolution at which we can perform these decompositions directly control the resolution at which we can recover heterogeneous material parameters \( k(x) \). For most materials of interest, these considerations imply that we need to be able to capture light transport decomposition at resolutions of micron-scale. This is three orders of magnitude than what is possible with previous methods for capturing light transport decompositions.

In this chapter, we introduce a new computational imaging system that uses optical interferometry to produce light transport decompositions at such resolutions. Our system uses optical configurations that are variations of the classical Michelson interferometer, and our analysis builds on techniques that have been used for optical coherence tomography (OCT). The use of interferometry allows our system to achieve pathlength resolutions as low as \( 10 \mu m \), as necessary for heterogeneous inverse scattering applications.

The chapter begins with background on the Michelson interferometer and the notions of spatial and temporal coherence length. We then introduce a mathematical model of interferometry in terms of a complex version of the pathlength-resolved Green’s function. We use it to show how sources with different coherence properties enable different kinds of light transport decompositions, differentiating light paths in terms of their endpoint locations, their optical lengths, or combinations of these two. We also characterize resolution and noise performance, and present a performance-optimized optical design that additionally allows resolving transport in terms of wavelength and polarization. Our prototype has three spectral channels, two polarization channels, a working volume of \( 2 \text{ cm} \times 2 \text{ cm} \times 1 \text{ cm} \), and spatial and pathlength resolutions that are both \( 10 \mu m \). We use this prototype to obtain micron-scale decompositions of light transport in scenes containing reflection, refraction, dispersion, scattering, and birefringence.
6.1 Related Work

Computer graphics and computer vision literature usually consider a discretized version of the steady-state Green’s function $T^s$. These correspond to a partition of the support of the camera’s importance function $W_o$ into $P$ area elements (“pixels”), and a similar partition of the support of the source’s emittance function into $L$ area elements. The continuous steady-state Green’s function $T^s$ is, then, analogously partitioned into a $P \times L$ matrix $T$, called the light transport matrix. Each entry $T_{pl}$ in this matrix represents the energy transferred from the $l$-th source element to the $p$-th camera element, through photons that follow paths starting at $l$ and ending at $p$. With reference to Figure 6.1, orange versus blue paths would correspond to entries $T_{pl}$ for different values of source location $l$. A similar discretization can be applied to pathlength and the pathlength-resolved Green’s function $T$, to obtain matrices $T^\tau$, for different values of $\tau$. Note that, while previously we used the Green’s function to describe light transport in a medium $M \subset \mathbb{R}^3$ filled with a scattering material, the radiative transfer equation and the rendering equation can be used to derive a Green’s function for any subset of $\mathbb{R}^3$, consisting of different scattering materials, surfaces, as well as vacuum. Throughout this chapter, we use the Green’s function and light transport matrix for such subsets, referring to them as “scenes”.

Using this discretization, the measurement equation of 3.20 can be written in discrete form as

$$i = \sum_{\tau} w(\tau) (M \odot T^\tau) l.$$  \hspace{1cm} (6.1)

This is a pathlength-resolved variant of the transport probing equation of O’Toole et al. [113]. In this equation, the operator $\odot$ represents pointwise multiplication (Hadamard product); $w(\tau)$ corresponds to the pathlength sampling function $W_\tau(\tau)$; $M$ is a $P \times L$ binary matrix with non-zero entries where the product of the emittance and importance functions $W_i(X_i)W_o(X_o)$ is non-zero; $l$ is an $L$-vector projected from the source and corresponding the non-zero values of $W_i(X_i)$; and $i$ is a $P$-vector corresponding to the image the camera.
Differentiating light paths. A pixel in a conventional image measures the sum of all lights paths arriving at that pixel. Entries of a light transport matrix distinguish paths with different starting locations (orange vs. blue), and entries of a “pathlength-resolved” light transport matrix additionally distinguish paths of different lengths (e.g., light vs. dark blue).

measures. Analogously to Section 3.5, conventional imaging corresponds to using a vector $w$ and matrix $M$ that are equal to 1 everywhere; whereas imaging methods for light transport decomposition collect measurements with different $M$ and $w(t)$. Below we discuss how the different types of decompositions described in Section 3.5 relate to the discretization of Equation (6.1), as well as review previous literature on implementing them. We will use Figure 6.1 to visualize paths captured by various approaches.

Spatial probing methods induce different matrices $M$ while keeping $w(t) = 1$ everywhere. With reference to Figure 6.1, these methods can be used to record contributions from only the blue or only the orange paths. By using different arrangements of a camera and a projector, and different sensor and illumination coding strategies, these methods capture individual elements of the transport matrix, or sparse or low-rank approximations to it [131, 116, 144, 7, 111, 113, 112].

Pathlength decomposition methods use a matrix $M = 1$ and controllable functions $w(t)$. Referring to Figure 6.1, this allows measuring the superposition of all paths (blue and orange) that have the same length. This has been dubbed transient imaging and can be used to produce light-in-flight visualizations of photons propagating in a scene. Implementations
include using a combination of pulsed laser and ultra-fast detector [143, 149, 148], or a combination of source and camera (time-of-flight sensor) that are synchronously modulated at radio-frequencies in time [58, 78, 110, 59]. In the latter case, the performance critically depends on the method used to solve the so-called “multi-path interference” problem [31, 59]. The equivalence of pathlength decomposition and light propagation has also been exploited for efficient rendering of light-in-flight [73].

Finally, there exist methods that provide simultaneous decompositions of space and pathlength. In the computer graphics literature, these fall into two categories. First are methods that decompose transport into a direct component (“one-bounce paths”) and an indirect or global component (“multi-bounce paths”) [106, 51, 124]. These imply a matrix $M$ that has only one non-zero entry per row, and a spatially-adaptive function $w(\tau)$ that selects for each such non-zero entry the shortest pathlength $\tau$ for which transport is non-zero. In these cases, the pathlength resolution $\Delta \tau$ determines how clean the resulting separations are, with sharper resolutions allowing less of the global component—such as that caused by sub-surface scattering in Figure 6.1—to bleed into the direct one. Second is the temporal frequency probing method of O’Toole et al. [110]. This allows decompositions with more general choices of matrix $M$ and function $w(\tau)$, for instance capturing transient images corresponding to fixed spatial probing patterns $M$.

**Interferometry.** Optical interferometry refers to a large set of imaging techniques that exploit interference between one or more electromagnetic fields [55]. Despite their popularity in other disciplines, interferometric techniques have rarely been used for computer graphics. An exception is Cossairt et al.’s [22] use of a Michelson interferometer for refocusing.

Many interferometric techniques can be interpreted using Equation (6.1). Abramson [1] used holographic techniques to create visualizations of light-in-flight similar to recent transient imaging; and the connection between interferometry and time-of-flight sensors has been discussed elsewhere [130, 93]. Here, we re-establish this relationship, and we expand it to
relate interferometry to other types of light path decompositions.

Among interferometric techniques, the one most closely related is optical coherence tomography (OCT) [63], which is used for range scanning and visualizing volume cross-sections at micron resolutions. There are many variants based on single or multiple shots (Fourier-domain vs. time-domain OCT) and single- or full-field illumination. In our paper, we adapt a full-field, time-domain OCT configuration for broader computational imaging. We re-interpret conventional OCT measurements using Equation (6.1), and then build on this to measure other light transport decompositions with different $M$ and $w(\tau)$ choices. We focus on time-domain OCT, as opposed to Fourier-domain, because it additionally allows decomposing transport according to wavelength and polarization.

**Comparison and trade-offs.** The main advantage of using interferometry to decompose transport is high pathlength resolution. Femtosecond lasers and time-of-flight sensors can achieve pathlength resolutions of about 600 $\mu$m and 10,000 $\mu$m, respectively. By comparison, we achieve pathlength resolutions of about 10 $\mu$m. Other interferometric techniques could be used to produce resolutions that are even higher, reaching sub-micron scales, for instance using very wide-band sources and post-capture processing that includes phase shifting. However, this would come at the expense of the ability to probe different spectral channels. Since we are interested in measuring wavelength-dependent effects, such as dispersion, birefringence, and scattering, we do not attempt to push pathlength resolution to those limits here.

This increased pathlength resolution is valuable in cases where it is necessary to obtain high-fidelity scans of small volumes, such as in laboratory material measurements or in scenes with natural microstructures. While macrophotography readily provides a way to image such scenes at spatial resolutions that approach the diffraction limit (a few microns), comparable pathlength resolutions cannot be achieved without the use of interferometry.

The main disadvantages of using interferometry to decompose transport are reductions
in field of view, depth of field, and speed of capture. Our working volume can only accommodate material samples or small objects, and our system is not appropriate for table-top or room-scale scenes. As we explain in Section 6.2, our method operates by slicing a volume in micron-sized depth intervals. This means that measuring a 1 cm deep volume requires 10,000 images and several hours, during which the scene must remain still. This is exacerbated by the fact that our optical setup is very sensitive to vibrations induced by environment sources, and even micron-scale vibrations can severely affect our measurements.

Another limitation of our approach relative to non-interferometric techniques is reduced flexibility in the spatial probing patterns $M$. We control these patterns by using mirrors of different shapes (Figure 6.6). This means that we can only induce patterns that correspond to manufacturable mirrors, and that changing the pattern $M$ requires substantial manual reconfiguration. This is in stark contrast to spatial probing methods [113, 112, 110] that allow complex patterns and simple reconfiguration in software.

The third limitation of our approach is shared by all methods for pathlength decomposition, namely, the reduced signal-to-noise ratio when multiple paths with widely varying energies arrive at the same pixel. In extreme cases, this makes it difficult or impossible to resolve low-energy paths, and unlike conventional HDR photography, this cannot be solved by capturing multiple exposures.

### 6.2 Optics Background

We first present the background necessary for understanding interferometry and our interpretation of it as decomposing transport. Our description of temporal coherence follows Goodman [47], and our description of spatial coherence follows Levin et al. [92].

Interferometry begins with the Michelson interferometer, shown in Figure 6.2, with a light
source of simultaneously small temporal coherence and spatial coherence. We will explain these notions and the operation of the Michelson interferometer intuitively, and then describe them analytically. The setup uses a beamsplitter to split the light wave emitted from the light source in two parts. One part is transmitted towards a reference mirror $M_r$ placed on a translation stage. The other part is reflected towards the target scene, which for our intuitive explanation is another mirror $M_s$, but can generally be any arbitrary scene. We refer to the sides of the setup containing the reference and target mirror as the reference arm and target arm, respectively. Both waves are then reflected, recombined at the beamsplitter, and imaged by a camera. We denote by $d_r, d_s$ the distance between the beamsplitter and mirrors $M_r$ and $M_s$, respectively.

The small temporal coherence of the light source means that, when $|d_r - d_s|$ is larger than an amount called the temporal coherence length of the source, the two waves do not interfere and the camera will measure an image equal to the sum of the two images of the mirrors, as shown in Figure 6.2.ii. If instead we use the translation stage to reposition $M_r$ to a position
so that $|d_s - d'_r|$ is smaller than the coherence length, the two waves will interfere and the camera will measure a fringe pattern, as shown in Figure 6.2.i.

When we replace mirror $M_s$ with a general scene, we can capture a sequence of images at different translations of the reference mirror $M_r$. By detecting interference patterns in the measured images, we can separate light paths of different pathlengths, up to a resolution equal to the temporal coherence length. We can additionally use sources with small spatial coherence to separate paths by location. This is because only light-paths that begin sufficiently close to each other on the source (i.e., are less than one spatial coherence length apart) and combine at the same sensor pixel will interfere. Note that this discussion is based on the coaxial interferometer design of Figure 6.2, which is the design we use exclusively in this paper. Oblique illumination is achievable, but it would require additional components as shown in Figure 6.2.iii.

This discussion highlights the importance of temporal and spatial coherence, which we will now describe analytically. They are wave phenomena, but scalar wave theory is sufficient to describe them, without having to employ the full complexity of electromagnetic theory. To simplify notation, and without loss of generality, we consider a two-dimensional world, with spatial points represented by their $(x, z)$ coordinates. Figure 6.3 shows the coordinate system.

### 6.2.1 Temporal Coherence

We begin by assuming that the source produces a plane wave propagating in the $z$ direction, so that the electromagnetic field is independent of $x$. Such a wave can be produced by transferring the output of a point source through a lens, as in Figure 6.8. If the wave is monochromatic at wavelength $\lambda$, the electromagnetic field at position $(x, z)$ and time $t$ can be described as

$$u(x, z, t) = a \cdot \exp \left(ik(-z\eta + ct)\right),$$  \hspace{1cm} (6.2)
where \( c \) is the speed of light in vacuum, \( \eta \) is the refractive index of the medium, \( k = 2\pi/\lambda \) is the wavenumber, and \( a \) is the complex amplitude of the source.

Real waves are not perfectly monochromatic, but are better described as superpositions of plane waves of different wavelengths,

\[
    u(x, z, t) = \int k a_k \exp(ik(-\eta z + ct)) \, dk, \quad (6.3)
\]

where the complex number \( a_k \) is the amplitude of each component. Typical real world sources can be modeled \cite{47} by assuming that \( a_k \) are random variables with phases sampled uniformly and powers \( |a_k|^2 \) sampled from a Gaussian with mean \( \bar{k} \) and standard deviation \( \Delta_k \),

\[
    |a_k|^2 \propto e^{-\frac{(k-\bar{k})^2}{2\Delta_k^2}}. \quad (6.4)
\]

The standard deviation \( \Delta_k \) is the *spectral bandwidth* of the source: smaller values of \( \Delta_k \) indicate a more monochromatic source.

Now suppose a sensor records at pixel \( x \) a measurement \( I(x) \) of the superposition of two copies of the wave \( u \) from Equation (6.3) that have traveled different distances \( z \) and \( z + \tau \).
The measurement is averaged over exposure time, producing

\[ I(x) = \langle |u(x, z, t) + u(x, z, t + \tau)|^2 \rangle_t \]
\[ = 2I_o(x) + 2 \Re \{ \text{corr}(x, \tau) \}, \quad (6.5) \]

where \( \langle \cdot \rangle_t \) denotes expectation over time, \( I_o(x) = \langle |u(x, z, t)|^2 \rangle_t \) is the intensity of the original wave, and

\[ \text{corr}(x, \tau) = \langle u(x, z, t)^* \cdot u(x, z + \tau, t) \rangle_t \]
\[ = 2I_o(x) + 2 \Re \{ \text{corr}(x, \tau) \}, \quad (6.6) \]

is the correlation of the two waves. If zero correlation exists between \( u \) and its shifted copy, the measurement will simply be equal to twice the intensity of the original wave, \( 2I_o(x) \). If positive correlation (constructive interference) or negative correlation (destructive interference) exists, the measurement will vary between 0 and \( 4I_o(x) \), depending on the relative shift \( \tau \) of the two copies. The maximum value is achieved when there is perfect correlation \( (\tau = 0) \). We will refer to the real part of the field correlation, equal to the difference \( I(x) - 2I_o(x) \), as the interference contrast. The temporal coherence length of the source, which we denote by \( L_c \), is the largest value of \( \tau \) for which significant correlation exists and therefore significant interference contrast is measured. To evaluate the temporal coherence length, we use the following claim.

**Claim 5.** The field correlation magnitude is a Gaussian with standard deviation inversely proportional to the spectral bandwidth \( \Delta k \),

\[ \text{corr}(x, \tau) = e^{i\bar{k}\tau} G_{\Delta k^{-1}}(\tau), \quad (6.8) \]

where

\[ G_{\Delta k^{-1}}(\tau) = e^{-\frac{1}{2}(\Delta k\tau)^2}. \quad (6.9) \]

The proof of the above claim is shown in the appendix. Here, we consider and experimental visualization of the result. Figure 6.4 shows the interference component we measured using the Michelson interferometer setup of Figure 6.2, plotted as a function of pathlength difference. As shown in the close-up plot, the profile looks like a fringe pattern, which varies as a function of the phase difference introduced by \( \tau \). When \( \bar{k}\tau = (2p + 1)\pi \) (p any
integer), we observe destructive interference (zero intensity); and when $\tilde{k}\tau = (2p)\pi$, we observe constructive interference (twice the signal intensity). As shown in the zoomed-out plot, the contrast of the fringe pattern decays and eventually becomes almost zero once the pathlength difference $\tau$ exceeds some amount. We can define this amount as the temporal coherence length of the source, which from Claim 5 is the standard deviation of the correlation window, $L_c = 1/\Delta_k$. We observe that $L_c$ is inversely proportional to spectral bandwidth. The larger the bandwidth of the source, the smaller the temporal coherence length, and therefore the higher the resolution at which we can discriminate between paths of different lengths.

6.2.2 Spatial Coherence

So far we have considered the case of a perfect plane wave propagating along the $z$ axis. Such a wave has perfect spatial correlation, in the sense that if we know the field value $u(x, z, t)$, we can also predict its value at any spatial shift $u(x + \xi, z, t)$. This is only realizable by using a lens to collimate the output of a source that has infinitesimal area. However, many real sources have larger effective areas. This induces another effect known as spatial incoherence [92], which also affects interference contrast.

Each point on an area source emits an independent wave, resulting cumulatively in a collection of independent plane waves $u_\theta(x, z, t)$ over a small angular range $\theta \in [-\Theta/2, \Theta/2]$. As in the previous section, we consider two copies of the cumulative wave, but in this case with one copy shifted not only in the $z$, but also in the $x$ direction. The intensity at a pixel on the sensor is given by integrating the intensities from Equation (6.5) for all independent waves,

$$I(x) = \int_{-\Theta/2}^{\Theta/2} \left( |u_\theta(x, z, t) + u_\theta(x + \xi, z + \tau, t)|^2 \right)_t \, d\theta. \quad (6.10)$$

This implies that the correlation is,

$$\text{corr}(x, \xi, \tau) = \int_{-\Theta/2}^{\Theta/2} \langle u_\theta(x, z, t)^* \cdot u_\theta(x + \xi, z + \tau, t) \rangle_t \, d\theta, \quad (6.11)$$
and similarly for the field intensity $I_0(x)$ and the interference contrast $I(x) - 2I_0(x)$. To evaluate how fast the interference contrast decays as a function of spatial shift $\xi$, we use the following claim.

**Claim 6.** For an area light source with an angular range $\theta \in [-\Theta/2, \Theta/2]$, the correlation decays as

$$\text{corr}(x, \xi, \tau) \approx W_{\Delta_c}(\xi)e^{i\hat{k}(\tau)}G_{\Delta_c^{-1}}(\tau),$$

(6.12)

where

$$W_{\Delta_c}(\xi) = \text{sinc} \left( \frac{\xi}{\Delta_c} \right), \quad \Delta_c = \frac{\lambda}{2\Theta}. \quad (6.13)$$

The proof of this claim is in the appendix. We see that, for area sources, the interference contrast decays very fast as the spatial shift $\xi$ increases, and no significant contrast is measured for values of $\xi$ larger than the quantity $\Delta_c$ of Equation (6.13). Analogously to the previous section, we define $\Delta_c$ as the spatial coherence length. The spatial coherence length is inversely proportional to the angular extent of the source. The wider the source, the smaller the spatial coherence length, and therefore the higher the resolution at which we can discriminate between paths with different points of origin on the light source.

### 6.2.3 Coherence Properties of Different Sources

There are many types of light sources with different coherence properties. As we will see in the next section (for instance, Figure 6.5), these properties determine the type of transport decomposition that our interferometer provides. We discuss various sources here, limiting our attention to those in the visible spectrum.

Laser sources are at one extreme end of coherence. These are high-power sources that can typically be modelled as ideal points (zero area), while also having very narrow, effectively monochromatic, spectral widths. As a result, their temporal and spatial coherence lengths can reach a few meters, implying poor spatial and pathlength resolution. For this reason, lasers are not appropriate for micron-scale light transport decompositions, but can be used
when micron-scale resolution is not necessary [1].

At the other extreme are light-emitting diode (LED) sources. These can be either “colored”, with a spectral bandwidth of a few tens of nanometers around their central wavelength, or very broadband, covering the entire visible spectrum. Because of their lower power density, LED sources have larger emitting areas, with widths ranging from several millimeters to a few centimeters. As a result, they have micron-scale temporal and spatial coherence lengths. Their exact values can be controlled by modulating the source output with an optical color filter, and placing an aperture between the source and the collimating lens: A narrower spectral filter means longer temporal coherence, and a smaller aperture means larger spatial coherence length. The increase in coherence comes at the expense of lower light output, and therefore longer exposure time.

Between the two extremes, there are two source categories. The first category includes superluminescent diode (SLD) sources, and supercontinuum laser sources. Like standard lasers, these are high-power sources that have infinitesimal effective area, but at the same time, they are polychromatic like standard LEDs. SLDs have bandwidths comparable to colored LEDs, and supercontinuum lasers span the entire visible spectrum. Therefore, these sources combine the meter-size spatial coherence lengths of lasers with the micron-scale temporal coherence lengths of LEDs.

Finally, the opposite combination corresponds to a monochromatic area source. Such a source can be created using a spatially dense bundle of multiple pinhole laser sources, or using specially manufactured, large-area laser semiconductors (diode bars and stacks). An alternative is gas-discharge lamps, which emit light only in very narrow spectral bands, corresponding to the spectral lines of the gas used (for instance, 589 nm for low-pressure sodium vapor lamps).

**Practical coherence values.** To illustrate how these calculations translate into practice, consider the coherence lengths of an LED source coupled with a color filter and aperture.
With respect to temporal coherence length, note that broadband sources and color filters are often described by the width $\Delta \lambda$ of the variation around the central wavelength $\bar{\lambda}$, rather than the variation $\Delta k$ around the wave number. Using simple Taylor expansion, we can relate this to the temporal coherence length as

$$L_c = 1/\Delta k \approx \frac{\bar{\lambda}^2}{2\pi \Delta \lambda}.$$  \hspace{1cm} (6.14)

Imagine a broadband LED, emitting in the entire visible range $[400, 700]$ nm. Ignoring any ultraviolet and infrared energy, we can set $\bar{\lambda} = 550$ nm and $\Delta \lambda = 150$ nm, resulting in coherence length $L_c \approx 0.3 \mu$m. In practice, this very fine resolution comes at the cost of very reduced interference contrast, due to chromatic aberration of optics. Alternatively, we can couple such a source with a color filter, or directly use a colored LED source. These alternatives give spectral widths in the range of $1 - 50$ nm, which correspond to temporal coherence lengths of a few microns. Additionally, as mentioned in Section 6.1, these narrower bands allow capturing measurements at multiple spectral channels. In Figure 6.4, we use the Michelson interferometer of Figure 6.2 to measure the temporal coherence lengths of such a combination, using three different color filters. Using a filter of bandwidth $25$ nm, we measure a temporal coherence length of about $10 \mu$m. This is the resolution we use in most of our experiments in the paper.

In terms of spatial coherence length, as indicated by Equation (6.13), the values we can achieve in practice are a function of the angular extent of the source. For visible wavelengths, assuming that we can collimate the source output to have an angular range of $4^\circ$, we have $\Delta_c \approx 8 \mu$m. With an angular extent of $1^\circ$ the coherence length increases to $\Delta_c \approx 32 \mu$m. In a natural environment where illumination arrives from multiple directions over the hemisphere, coherence length can approach the wavelength of light.
6.3 Light Path Decomposition

Equipped with analytical representations of temporal and spatial coherence, we are ready to introduce a mathematical model of interferometric decomposition of light transport. The core of this section is Equation (6.23), which relates discrete interference measurements to an underlying pathlength-resolved transport function—a continuous version of the pathlength-resolved transport matrix of Equation (6.1). Using this model, we re-interpret the output of time-domain, full-field OCT as decomposition by pathlength $\tau$ (transient imaging), and we describe combinations of sources and reference mirrors that provide other types of decompositions.

Throughout this section, we assume that we image with a camera focused at depth $z_o$, as shown in Figure 6.3. As we restrict the discussion to points on the $z = z_o$ plane, from this point forward we simplify notation by omitting the $z$ coordinate of fields. We denote by $u_{in,\rho}(x,t)$ the incoming field arriving at both the reference and target arms at time $t$ after being emitted from the source. This is a superposition of monochromatic waves, as in
Equation (F.7). We denote by $u_{s,\beta}(x, t)$ the field scattered by the scene at space point $(x, z_0)$ and time $t$. Similarly, we denote by $u_{r,\beta,\tau}(x, t)$ the reference field at point $(x, z_0)$ when the reference mirror is positioned at depth $z_0 + \tau/2$. We begin by deriving expressions that relate the reference and scattered fields to the input field.

**Reference field.** The reference field $u_{r,\beta,\tau}$ relates to $u_{in,\beta}$ through a spatial transformation $f(x)$ and a temporal shift,

$$u_{r,\beta,\tau}(x, t) = u_{in,\beta}(f(x), t - \tau/c). \quad (6.15)$$

In the standard Michelson interferometer of Figure 6.2, we have simply $f(x) = x$. We can produce more general transformations $f(x)$ with other mirror configurations (Figure 6.6). The form of $f(x)$ will determine the non-zero entries of the matrix $M$ from Equation (6.1).

**Scattered field.** The scattered field $u_{s,\beta}$ relates to the input field through a more general transformation, which is a function of the scene’s geometric, refractive, and material properties. To derive its form, we introduce some additional notation. We denote by $\{o_\gamma\}_{\gamma \in \mathbb{A}(x_1, x_2)}$ the set of all possible paths in the scene from point $(x_1, z_0)$ to point $(x_2, z_0)$ (see Figure 6.3 for the case $x_1 = x_2$), by $\ell(o_\gamma)$ the optical length of each path, and by $a(o_\gamma)$ the energy loss along the path—this loss is a function of the propagation distance and volumetric absorption along the path. Depending on the scene, such paths can be of many forms: direct paths traveling from the source to the object and then the camera, multiple reflections, multiple scattering, and so on. In the simple case that the scene is a perfect planar mirror, there are paths with non-zero absorption only if the start and end point are the same, $x_1 = x_2$, and for every $x_1$ only a single such path exists—the direct path from $(x_1, z_0)$ to $(x_1, z_0)$. For a path in air, where the refractive index $\eta \approx 1$, $\ell(o_\gamma)$ is equal to the geometric path length; in media with larger refractive indices, the path length also folds in the optical path delay. Finally, we denote by $\mathcal{T}(x_1, x_2, \tau)$ the complex scalar resulting from
the integration of all paths from \( x_1 \) to \( x_2 \) with optical length \( \tau \),

\[
\mathcal{T}(x_1, x_2, \tau) = \int_{\{o_\gamma | \ell(o_\gamma) = \tau\} \cap \mathcal{A}(x_1, x_2)} a(o_\gamma).
\] (6.16)

We call \( \mathcal{T}(x_1, x_2, \tau) \) the complex pathlength-resolved Green’s function. When the scene is a planar mirror, \( \mathcal{T}(x_1, x_2, \tau) = \delta(x_1 - x_2, \tau - d_s) \), where \( d_s \) is the depth of the scene mirror. By definition, we observe that the function \( \mathcal{T}(x_1, x_2, \tau) \) is the complex-valued continuous version of the pathlength-resolved light transport matrix, \( \mathcal{T}^* \), introduced in Section 6.1.

In parts of the following discussion, we will be employing the light transport matrix discretization for visualization and intuition, with the understanding that we continue to use the complex continuous version \( \mathcal{T}(x_1, x_2, \tau) \).

We can now use the above functions to derive the scattered field \( u_{s,\theta}(x, z_0) \). This is the superposition of contributions along all possible paths, with a time delay proportional to the path length,

\[
u_{s,\theta}(x_1, \tau) = \int_{\{o_\gamma \} \cap \mathcal{A}(x_1, x_2)} a(o_\gamma) \int_{\mathcal{A}(x_1, x_2)} \alpha(x + \xi, \tau - \frac{\ell(o_\gamma)}{c}) \, d\xi
\]

\[
= \int_{\xi} \int_{\xi} \mathcal{T}(x + \xi, x, \xi) u_{in,\theta}(x + \xi, \tau - \frac{\ell(o_\gamma)}{c}) \, d\xi \, d\xi.
\] (6.17)

The above equation relates the field to the function \( \mathcal{T}(x_1, x_2, \tau) \) for the scene. The term \( u_{in,\theta}(x + \xi, \tau - \frac{\ell(o_\gamma)}{c}) \) is complex-valued and includes the phase difference between the integrated components, due to the different lengths of paths in each component.

**Interference of reference and scattered fields.** Having derived expressions for the reference and scattered fields, we now compute the interference contrast that can be measured by a camera. Ignoring for now diffraction blur, a camera focused at depth \( z_0 \) will measure the superposition of the two fields,

\[
I_{i,\tau}(x) = \left\langle \left| u_{r,\theta,\tau}(x, t) + u_{s,\theta}(x, t) \right|^2 \right\rangle_t \, d\theta
\]

\[
= I_s(x) + I_{r,\tau}(x) + 2 \text{Re} \left\{ \int_{\theta} \left\langle u_{r,\theta,\tau}(x, t)^* \cdot u_{s,\theta}(x, t) \right\rangle_t \, d\theta \right\},
\] (6.18)
where by \( I_s(x), I_{r,\tau}(x) \) we denote the intensities of the scattered and reference fields, respectively,

\[
I_s(x) = \int_{\theta} \left\langle |u_{s,\theta}(x,t)|^2 \right\rangle_t \, d\theta, \quad I_{r,\tau}(x) = \int_{\theta} \left\langle |u_{r,\theta,\tau}(x,t)|^2 \right\rangle_t \, d\theta.
\tag{6.20}
\]

The intensity components \( I_s \) and \( I_{r,\tau} \) correspond to the images the camera would measure if we blocked the mirror or scene arms of the setup, respectively, and measured the other. By subtracting these components, we remain with the interference contrast, equal to the real part of the correlation of the reference and scattered fields,

\[
C_{\tau}(x) = \int_{\theta} (u_{r,\theta,\tau}(x,t)^* \cdot u_{s,\theta}(x,t))_t \, d\theta.
\tag{6.21}
\]

\( C_{\tau}(x) \) is a correlation signal, analogous to \( \text{corr}(x, \xi, \tau) \) in Equation (6.11). We denote it by a distinct symbol because of its importance in our application. Using the coherence properties for correlation signals from Section 6.2, we can prove the following claim.

**Claim 7.** The correlation of the reference and scattered fields is

\[
C_{\tau}(x) \propto \int_\xi W_{\Delta \xi}(\xi) \int_\xi \tilde{T}(f(x) + \xi, x, \xi) e^{ik(\xi-\tau)} G_{\Delta \xi^{-1}}(\xi - \tau) \, d\xi \, d\xi.
\tag{6.22}
\]

The proof of this claim is shown in the appendix. The above claim shows that the correlation \( C_{\tau}(x) \) is equal to the complex pathlength-resolved light transport function \( \tilde{T}(f(x), x, \tau) \) from entrance point \((f(x), z_0)\) to exit point \((x, z_0)\), blurred in space with a sinc of width proportional to the spatial coherence length, and blurred in the pathlength dimension with a Gaussian of width proportional to the temporal coherence length of the source. By recording the intensity \( I_{i,\tau}(x) \) of Equation (6.18) and subtracting the individual intensities of the source and reference fields, we have

\[
T^m(f(x), x, \tau) \triangleq |I_{i,\tau}(x) - (I_s(x) + I_{r,\tau}(x))|^2 \tag{6.23}
\]

\[
= |2 \text{Re}(C_{\tau}(x))|^2
\]

\[
= \left| 2 \text{Re} \left( \tilde{T}(f(x), x, \tau) * W_{\Delta \xi} * (e^{ik\tau} G_{\Delta \xi^{-1}}(\tau)) \right) \right|^2.
\]

This equation is our main result. It shows that using interference, we can measure blurred
samples of the complex pathlength-resolved light transport function. The spatial resolution is controlled by the size of the spatial blur kernel, which is in turn controlled by the illumination angle of the light source. Similarly, the pathlength resolution is controlled by the size of the blur kernel in the pathlength dimension, which in turn is controlled by the spectral bandwidth of the light source. The exact correspondence \((f(x), x)\) between entrance and exit points where we sample the light transport function depends on the transformation implemented at the reference arm of the setup. Finally, the pathlength where we sample the light transport function depends on the depth of the reference arm.

We note here that the above derivation applies to the complex Green’s function \(\tilde{T}\) instead of the Green’s function \(T\) derived from light transport theory. The complex variant is often referred to in wave optics as the transmission matrix, and its relationship to the transport version has been studied extensively in the related literature [114]. Review this subject in detail is beyond the scope of this thesis, but we mention that in scenes containing scattering materials with very small mean free path, or opaque materials with very diffuse BRDF, the measurements of the form (6.23) are proportional to those we would obtain from a decomposition of the radiometric Green’s function \(\tilde{T}\). For the case of scattering materials, we discuss this relationship in more detail in the appendix.

### 6.3.1 Decomposition Types

Based on our model in Equation (6.23), the interferometer of Figure 6.2 can be equipped with various sources and mirror configurations to produce different types of light transport decompositions. The following discussion is summarized in Figures 6.5 and 6.6.

**Pathlength decomposition.** The first case is equivalent to transient imaging, and corresponds to setting \(M = 1\) and \(w(\tau) = \delta(\tau - \tau_o)\) in Equation (6.1). For the former, we require a source with very large, essentially infinite, spatial coherence length, meaning that the spatial blur kernel \(W_\Delta\) in Equation (6.23) is equal to 1 everywhere. For the latter, we need
By equipping the Michelson interferometer with sources having different coherence properties, we can capture different light path decompositions. (a) Pathlength decomposition separates all paths in a scene \( (M = 1) \), regardless of their endpoint locations, in terms of their optical length \( t \). It separates direct paths (blue), reflections (purple), retroreflections (orange), and all scattering (greens). (b) Spatial probing separates paths with certain start point locations while ignoring their lengths (by effectively measuring summations over pathlength). For a diagonal \( M \), it separates direct, retroreflection, and backscattering (dark green) paths. (c) Finally, it is also possible to combine the previous two cases and separate paths in terms of both pathlength and endpoint locations.

Figure 6.5: By equipping the Michelson interferometer with sources having different coherence properties, we can capture different light path decompositions. (a) Pathlength decomposition separates all paths in a scene \( (M = 1) \), regardless of their endpoint locations, in terms of their optical length \( t \). It separates direct paths (blue), reflections (purple), retroreflections (orange), and all scattering (greens). (b) Spatial probing separates paths with certain start point locations while ignoring their lengths (by effectively measuring summations over pathlength). For a diagonal \( M \), it separates direct, retroreflection, and backscattering (dark green) paths. (c) Finally, it is also possible to combine the previous two cases and separate paths in terms of both pathlength and endpoint locations.

A source with a very short temporal coherence length, making the pathlength blur kernel \( G_{\Delta_t} \) in Equation (6.23) very narrow. As discussed in Section 6.2.3, this is achieved using an SLD or a supercontinuum laser. Capturing an entire transient sequence corresponds to sweeping over \( \tau_0 \) values, which can be done by densely scanning the reference arm to different positions. This hardware combination and capture procedure is exactly equivalent to conventional full-field, time-domain OCT.

**Spatial probing.** Referring again to Equation (6.1), another form of decomposition corresponds to setting, \( w(\tau) = 1 \) everywhere, and using a spatial probing pattern \( M \) with non-zero entries only for the paths we want to preserve. This requires a source with a
near-infinite temporal coherence length and a very small spatial coherence length. From Section 6.2.3, such a source can be produced using a laser bundle or gas-discharge lamps. Additionally, we require a reference mirror configuration whose shape induces the spatial transformation $f(x)$ corresponding to the desired pattern $M$. Using a planar mirror as in the standard Michelson interferometer corresponds to a matrix $M$ that is non-zero only in its main diagonal. We discuss alternatives later in the section. Due to the infinite temporal coherence length, this case does not require a scan of the reference arm. We mention spatial probing here for completeness, but note that, for this type of decomposition, interferometry does not offer any advantages compared to the other, easier to implement and use, techniques discussed in Section 6.1. For this reason, we did not implement this case, and only used spatial probing combined with pathlength decomposition, as discussed in the next paragraph.

**Pathlength decomposition and spatial probing.** The last case decomposes light transport simultaneously in terms of pathlength and spatial correspondences. As suggested by the previous two cases, we can do this using a source with temporal and spatial coherence lengths that are both small. An LED source is the appropriate choice as described in Section 6.2.3. As is needed for pure pathlength decomposition, capture requires a dense scan of the reference arm to many different positions.

**Creating probing patterns.** In the last two decomposition types, replacing the planar mirror $M_r$ in the reference arm of the Michelson interferometer with different mirror shapes induces different correspondences $f(x)$, that can be used to probe off-diagonal elements of the light transport matrix. For instance, the right-angle pair of mirrors in Figure 6.6(b) induces a vertical flip, or $f(x) = -x$. This corresponds to measuring the anti-diagonal of the light transport matrix, which includes two-bounce paths as shown in Figure 6.6(b). Shifting such a pair of mirrors vertically in the figure measures arbitrary anti-diagonals, as in Figure 6.6(c); and combinations of parallel mirrors and orthogonal mirror pairs sample different blocks of the light transport matrix in different ways, as shown in Figure 6.6(d). Note that the probing
Figure 6.6: Reference mirror configurations in a Michelson interferometer induce different spatial probing patterns. Assuming co-axial source and camera: (a) A planar mirror probes the main diagonal of the light transport matrix, corresponding to direct, retroreflection, and backscattering paths. (b) A right-angle mirror pair probes the anti-diagonal, which includes 2n-bounce reflection paths. (c) Offsetting a right-angle pair allows probing away from the anti-diagonal. (d) Different combinations of these basic configurations probe blocks of the light transport matrix in different ways.

pattern $M$ produced by these mirror configurations will be different if we change our setup so that the camera and source are no longer co-axial.

### 6.4 Pipeline and Design Considerations

Having presented our setup and methodology, we now discuss the computational post-processing and other practical considerations for optimizing the performance of our setup for graphics applications. Throughout this section, we consider a toy scene consisting of a tilted diffuse plane, as shown in Figure 6.7(a).
Capture and computational post-processing. Measuring the scene and mirror images $I_s$ and $I_{r,\tau}$ separately is impractical, as it doubles acquisition time. In practice, since the reference arm contains a mirror configuration, the intensity $I_{r,\tau}$ does not vary much with $\tau$ and can be estimated by smoothing the $\tau$ dimension. For this reason, we perform a single scan of the scene, capturing a set of frames corresponding to $I_{i,\tau}(x)$ for various values of $\tau$, as shown in Figure 6.7(a)-(b). Scanning a volume of 1 cm using this process typically requires capturing 10000 images (one image per micron-translation of the reference arm), a process which can take up to several hours for scenes requiring long exposure times. As we mention in Section 6.5, for scenes that require multiple spectral channels or high-dynamic range, we need to repeat the scanning process several times, resulting in hundreds of thousands of images.

Following the capture section, in post-processing we approximate the interference-free component as

$$\hat{I}_s(x) + \hat{I}_r(x) \approx \text{median}_\tau(I_{i,\tau}(x)). \quad (6.24)$$

Furthermore, in real scenes the interference signal $T^m$ of Equation (6.23) involves a high-variation pseudo-random speckle noise. To eliminate these speckle artifacts, we blur our measurements over a small spatial window with a filter $g(x)$. Our final estimate of the optical pathlength decomposition function is, then, computed as

$$\hat{T}^m(f(x), x, \tau) = T^m(f(x), x, \tau) * g(x)$$

$$= |I_{i,\tau}(x) - \hat{I}_s(x) - \hat{I}_r(x)|^2 * g(x). \quad (6.25)$$

This blurring is one of the factors that result in a reduction in the spatial resolution of our measurements. We discuss speckle noise in more detail in the appendix. In the remaining of this section, we discuss other factors, as well as practices to maximize the signal-to-noise ratio of the measurements (6.25).

Contrast. Assume we want to capture the pathlength component which corresponds to
optical length $\tau$ and that the reference mirror is positioned at the corresponding depth. We will define the contrast as the spatially averaged magnitude of the interference (correlation) component, over the intensity components,

$$\text{Contrast} = \frac{\sqrt{\mathbb{E}_x[T^m(f(x), x, \tau)]}}{\mathbb{E}_x[I_s(x) + I_{r,\tau}(x)]}. \quad (6.26)$$

To avoid saturation, we set the exposure time so that the averaged intensities without interference fall in the middle of the sensor’s dynamic range, $\mathbb{E}_x[I_s(x) + I_{r,\tau}(x)] \approx 0.5$. Then, Equation (6.26) is simply the averaged magnitude of the interference contrast.
Figure 6.8: Optimizing the Michelson interferometer of Figure 6.2. The schematic to the left describes our implementation. We consider the effect of the following imaging parameters: source aperture area $A_s$, camera aperture area $A_c$, and pixel pitch $p$. The optimized settings correspond to camera F-number of $f/32$, source aperture so that the source’s angular extent is approximately half that of the camera, and pixel pitch of 4 $\mu$m. Different configurations are produced by modifying only one parameter. For each configuration, the middle graph shows the measured contrast for the tilted diffuse plane scene of Figure 6.7; and the right graph compares contrast and total exposure time.

A short calculation shows that the contrast in Equation (6.26) is maximized when the energy of the reference field $I_{r,T}(x)$ is equal to the energy of the scattered field at the optical length $T_m(f(x), x, \tau)$. Because in diffuse scenes only a small portion of the energy in the target arm returns to the camera, to achieve the above ratio, we need to attenuate the very high amplitude of the reference arm. In most scenes, the intensity at a single point is the result of contributions from multiple paths, making it impossible to a-priori match the energy of the pathlength $\tau$ component only. Instead, we attenuate the reference arm so that its intensity matches the total intensity. As discussed in Section 6.5, we do this using either cross-polarization, or neutral density filters.

**Optics blur.** The ideal analysis of the Section 6.3 does not take into account blur introduced by the optics of the interferometry setup, including diffraction and sensor blur. These blur processes can reduce interference contrast, unless the various aperture and magnification parameters are chosen appropriately. We discuss these effects below, and in greater detail in the appendix.

We consider as before a camera focused at $z_0$, as shown in Figure 6.3. From diffraction theory [46], we have that instead of the field $u_\theta(x, t)$ arriving at point $(x, z_0)$, the camera
measures the field blurred by the camera diffraction blur kernel, \( W_{\Delta \phi} \ast u_\theta(x, t) \), where the diffraction blur width \( \Delta_\phi = \bar{\lambda} / (2\Phi) \) is inversely proportional to the acceptance angle \( \Phi \) of the camera.

The intensity measured by the camera is the intensity of the blurred field, averaged over the area of a pixel on the sensor,

\[
I(x) = \Pi_{\Delta x} \ast \int_{-\Theta/2}^{\Theta/2} \langle |W_{\Delta \phi} \ast u_\theta(x, t)|^2 \rangle_t \, d\theta,
\]

where \( \Pi_{\Delta x} \) is the pixel-sized rectangular function.

As a result of the above two processes, measurements from Equation (6.23) take the form,

\[
T^m(f(x), x, \tau) = |2 \text{Re}(\beta \ast C_\tau(x))|^2
\]

\[
= \left| 2 \text{Re} \left( \beta \ast W_{\Delta \phi} \ast \bar{T}(f(x), x, \tau) \ast (e^{i\bar{k}G_{\Delta c^{-1}}(\tau)}) \right) \right|^2,
\]

where \( \beta = \Pi_{\Delta x} \ast W_{\Delta \phi} \).

It is important to observe that the additional blur \( \beta \) acts on the complex valued correlation component, rather than on powers (intensities). Therefore, if the correlation signal \( C_\tau(x) \) contains spatial features whose frequency is higher than the widths \( \Delta_x, \Delta_\phi \) of the blur kernels, the averaged power of measurement signal is reduced. From Equation 6.22, we know that the signal \( C_\tau(x) \) has limited spatial resolution \( \Delta_c \), as a result of the limited spatial coherence of the light source. Therefore, to achieve good contrast, we need to select the imaging parameters such that the optics blur remains below the signal resolution, that is, \( \Delta_x < \Delta_c \) and \( \Delta_\phi < \Delta_c \).

We note that requesting \( \Delta_\phi < \Delta_c \) is equivalent to requesting that the angular extent of the source is smaller than the acceptance angle of the camera, \( \Theta < \Phi \). In general, using a wide source is desirable, as it increases the overall power of the illumination and therefore reduces exposure time. However, increasing the source angular extent without adjusting the imaging parameters accordingly will also reduce the contrast of the correlation signal. To
relate these directly to imaging parameters, we denote by \( Y \) the numerical aperture of the camera lens. A short calculation shows that a camera imaging at magnification \( m \) accepts light up to a maximal angle \( \Phi = \frac{Y}{1 - 1/m} \).

In Figure 6.8, we quantify the above observations by measuring the detection contrast for the tilted diffuse plane scene of Figure 6.7 under different imaging configurations. We vary the size of the source and camera aperture, and the size of the camera pixel—which is analogous to changing the camera magnification. We observe that setting imaging parameters sub-optimally can result in loss of more than half the detection contrast. On the other hand, optimizing detection contrast comes at the cost of increased exposure time.

### 6.5 Implementation and Experiments

Figure 6.8 shows a schematic of our optical setup. We select sources and mirrors corresponding to configurations (a) and (c) of Figure 6.5, and (a) and (b) of Figure 6.6, respectively. The exact configuration depends on the application, as described below.

Our schematic has two additions compared to Figure 6.2. First, we modulate the output of the source source using color filters of bandwidth 25 nm (e), producing a temporal coherence length of about 10 \( \mu \)m (Figure 6.4). To enable RGB imaging, we use a filter wheel to automatically change between filters centered at different wavelengths: 625 nm (red), 525 nm (green), and 450 nm (blue).

Additionally, we use three polarizers: at the source (d), camera (j), and reference arm (g). The polarizers on the source and camera are rotated to be either parallel or crossed with each other, depending on which polarization component of the scene’s transport is to be measured. The polarizer on the reference arm (g) is rotated relative to the other two in order to attenuate the reference beam by different amounts. This enables HDR imaging: we scan exposure brackets where, for each setting of the camera’s exposure time, the reference
polarizer (g) is adjusted so that the reference beam’s intensity is at roughly one-quarter of the sensor’s dynamic range (for a total intensity at roughly half the dynamic range when imaged together with properly exposed parts of the target scene). When we need to measure both polarization components, we remove polarizers (d) and (j), and replace (g) with a sequence of neutral density filters.

We discuss our implementation in detail in the appendix. In the rest of this section, we use this assembly to analyze transport in various scenes and for various applications. We first discuss experiments using an LED source (Figure 6.5(c)).

**Depth scanning.** Following Section 6.3.1, we use an LED source and a planar reference mirror to measure the pathlength decomposition of the diagonal component \( T^m(x, x, \tau) \). As our camera and source are co-axial, we then obtain the depth \( D(x) \) at pixel \( x \) as

\[
D(x) = \min \{ \tau : T^m(x, x, \tau) \geq 0 \}.
\]  

Figure 6.9 shows scans of a few different objects. Each scan measures a volume of 15 mm at a step size of 1 \( \mu \)m, and requires roughly eight hours of capture time. The effective depth

---

![Figure 6.9: Scanning depth. For each scene from top to bottom: an unprocessed HDR frame; scanned depth, visualized using pseudo-color and a periodic change in luminance for a contouring effect; and rendering of scanned 3D mesh. All scans are at a depth/pathlength resolution of 10 \( \mu \)m. Scenes from left to right: US quarter, gnocchi, soap carving, plastic toy cup, dry pasta, and gummy bear.](image)
Figure 6.10: Depth scans of complex scenes. From left to right: an unprocessed HDR frame; scanned depth, visualized as in Figure 6.9; and rendering of 3D mesh. Top row: Chess knight between a diffuse wall to the right and a mirror to the left. The mirror-reflection is interpreted as a real object at a greater depth. Bottom row: Gummy bear between two diffuse walls. The scene geometry is recovered accurately, despite the conflation of paths from the diffuse walls behind the near-transparent gummy bear.

resolution is 10 μm. We can measure depth in a number of challenging situations, such as when materials exhibit substantial scattering (soap), semi-transparency (gummy bear), and strong caustics and specular or diffuse interreflections (cup and pasta). The very fine pathlength resolution of our setup translates directly into sharp depth resolution, which allows capturing details such as the hair texture on the coin and the fine texture on the gnocchi surface.

Figure 6.10 shows depth maps obtained for two more scenes. The stage step size and effective resolution are the same as before, but the volume and capture time increase to roughly 25 mm and 12 hours, respectively, to capture the longer reflection paths. The first scene includes sharp specular interreflections, with a single chess piece between an angled diffuse wall (right) and a mirror (left). In the acquired depth map, the mirror reflections are interpreted as real objects at a greater depth. Despite the indirect reflection by the mirror, the
depth map of the chess piece is accurately recovered. The second scene shows a translucent gummy bear between two diffuse walls. Compared to the rightmost column of Figure 6.9, where the gummy bear is surrounded by air, here the combination of a diffuse wall behind a near-transparent gummy bear induces many distinct paths of similar intensity at each camera pixel. Despite this conflation, the scene geometry is still recovered accurately.

**Separation of direct and global components.** We can use the same source and mirror combination as before to also do direct-global separation. Since the source and camera are coaxial, the direct component at each pixel is the energy of the shortest path that contributes to the diagonal of the light transport matrix. Thus,

\[
I_{\text{direct}}(x) = T^m(x, x, D(x)),
\] (6.30)

where \(D(x)\) is the depth from Equation (6.29). The global component sums the contributions from all other paths and, given an image \(I(x)\) captured using a conventional camera, can be obtained as

\[
I_{\text{global}}(x) = I(x) - T^m(x, x, D(x)).
\] (6.31)

We note that spatial probing alone is not sufficient for computing the exact direct component, and it is necessary to also do pathlength decomposition. With reference to Figure 6.1, spatial probing (diagonal of the light transport matrix) captures the various blue paths, which include direct paths and retroreflections, but also backscattering. Using pathlength decomposition to select only the shortest path additionally removes the backscattering contributions.

Figure 6.11 shows such a decomposition for a close-up of a strawberry. We produced this separation from 12 scans of the strawberry (4 exposures × 3 spectral channels). Each scan was performed at a step size of 10 μm, meaning that we slightly undersampled in the pathlength domain relative to the source’s pathlength resolution of 10 μm. Additionally, we used 8 × 8 pixel binning on the camera to speed-up frame streaming, and large source and camera apertures to decrease exposure time. Even though these imaging settings
are suboptimal (see discussion of Figure 6.8), they were necessary to reduce the total scanning time to a little below two hours, thereby accommodating the limited shelf life of the strawberry.

We observe that the direct component is essentially achromatic and captures features characteristic of direct illumination, such as the specular highlights and high-frequency surface structure. The indirect component in this image is primarily volumetric scattering, and we see that it accounts for the red color of the strawberry and a large part of the coloration of the strawberry seeds.

**Scattering.** As discussed in Section 6.3.1, we can also obtain a decomposition of scattering paths that begin and end at the same locations on the surface of a scattering volume. To demonstrate this, Figure 6.12 shows the results from scanning a scene with materials of different scattering properties, such as a metal, a diffused wall and a highly-scattering zirconia coating. Figure 6.12(c) shows a processed frame from the scan, where the intensity tails at the zirconia coating and diffuse wall regions are produced by the strong backscattering of those materials. In Figure 6.12(d), we see the pathlength slice $T^m((x, y_0), (x, y_0), \tau)$. (We use pairs $(x, y)$ to indicate the 2D pixel coordinates). At pixels corresponding to zirconia coating and the diffuse wall, the interference remains noticeable over a long range of pathlengths.
Figure 6.12: Scanning a scene with materials of different scattering properties using an LED source. (a) A photograph of the scene: A turbine blade with a highly diffusive zirconia coating is placed in front of a spectralon slab. (b) An unprocessed frame. (c) A processed frame, showing as an image the value of the measurement $T_m((x,y_0),(x,y),\tau_o)$ for some value of $\tau_o$ and all pixels $(x,y)$. (d) The function $T_m((x,y_0),(x,y),\tau)$ for pixels $(x,y_o)$ on the horizontal crosssection shown in (b), and for all values $\tau$.

Conversely, we see negligible scattering in the uncoated metal pixels.

In Figure 6.13(a), we show scattering profiles $T_m((x_o,y_0),(x_o,y_o),\tau)$, as a function of path-length. These were measured by scanning uniform slabs of various translucent materials: whole milk, milk soap, wax, and the zirconia coating mentioned previously. For materials that are optically dense and strongly forward scattering, such as whole milk, the measured intensity drops very sharply. On the other hand, for the diffuse zirconia coating, we can measure non-zero interference for pathlengths as long as 3 mm inside the medium.

In Figure 6.13(b), we additionally show the scattering profiles measured in the zirconia material for different polarization channels. Specifically, we measure the scattering profiles
Figure 6.13: Pathlength decomposition of scattering paths starting and ending at a single point on the surface of a scattering volume. All measurements use an LED source. Left: Measured scattering profiles $T^m((x_o, y_o), (x_o, y_o), \tau)$ at a single pixel $(x_o, y_o)$ as a function of $\tau$, for various materials. Right: Measured profiles for the parallel and cross-polarized scattering component of zirconia.

when the polarizers (d) and (j) in Figure 6.8 are rotated so that their axes are parallel or orthogonal (“crossed”). We observe at short pathlengths, the component corresponding to parallel polarization is much stronger than that corresponding to crossed polarization. However, as the pathlength increases, the parallel- and cross-polarized components become equal. This is in agreement with the well-known observation that low-order scattering preserves polarization, but high-order scattering becomes randomly polarized. For reference, we also plot the scattering profile for this material from Figure 6.13(a), where there is no decomposition of polarization.

**Dispersion and birefringence.** As discussed in Section 6.3, interferometry produces light path decompositions in terms of *optical* instead of geometric pathlength. Two light paths that have the same geometric length can have different optical length if they travel through media with different refractive indices. This could be either because they travel through different optical materials, or because of phenomena such as dispersion and birefringence, where the refractive index of a material changes as a function of the wavelength and polarization of light. We visualize these three phenomena using our setup with an LED source and a planar reference mirror.
In Figure 6.14, we scan a scene consisting of a glass container partially filled with water, placed in front of a mirror. Paths from the light source to the mirror and back to the camera all have the same geometric length. However, they have different optical lengths, depending on whether they travel through only air, through the empty part of the glass container, or through the part filled with water. Additionally, optical lengths vary as a function of wavelength, because of the dispersion of glass and water. This produces a “rainbow” when we visualize the measurements $T^m((x_o, y), (x_o, y), t)$ as a function of pathlength, even though such a rainbow is not visible in a static image. The path delay between different spectral bands increases for paths going through both water and glass (roughly 20 μm), compared to paths going through glass only (roughly 10 μm), due to the additional dispersion of water.

In Figure 6.15, we visualize birefringence by scanning a scene consisting of a plastic protractor placed in front of a mirror. The mechanical stresses in different parts of the plastic material result in the refractive index of the material varying spatially as a function of polarization and wavelength, a behavior known as photoelasticity. This creates familiar color patterns when the protractor is viewed under standard polarized light without any interference (Figure 6.15(b)). The refractive index variations additionally mean that paths travelling through the protractor to the mirror and back to the camera will have different spatially varying pathlength. This is shown in Figure 6.15(d), where we scanned and visualized the contributions from different pathlengths at different parts of the protractor. The spatial variations of the measured optical pathlengths are smaller than 50 μm, and they are roughly aligned with the RGB color patterns that appear under standard polarized light.

**Using different probing patterns.** As discussed in Section 6.3.1, by using different reference mirror configurations, we can probe different components of the light transport matrix. In Figure 6.16, we compare measurements of the same scene with two of these configurations. The scene is a realization of the two-dimensional mirror-diffuser 90° corner we have been using as example throughout Section 6.3.1 (Figures 6.5 and 6.6), with the addition of a thin
Figure 6.14: Visualizing dispersion in the pathlength domain. (a) A schematic of the scene, consisting of a glass container partially filled with water and placed in front of a mirror. As light travels through the glass and water, different wavelengths are delayed by different amounts due to dispersion. (b) A photograph of the scene we scan with an LED source. We consider the close-up shown in the box. (c) Frames acquired at a wavelength of 525 nm, and at three different pathlengths. Spatial fringes indicates interference of paths that travel through, from left to right, air only, air and glass, and air, glass, and water. (d) Visualization of the interference measured across a vertical cross-section of the frames, as a function of optical pathlength and for three different wavelengths. We observe that paths travelling through glass or through both glass and water are delayed more compared to paths travelling through air only. Additionally, the delay is a function of wavelength, resulting in the appearance of a “rainbow” as a function of optical pathlength.

slab of glass bisecting the angle between them. We do all measurements using an LED source (Figure 6.5(c)) with our three spectral filters.

We first use a planar mirror in the reference arm, which allows measuring the pathlength decomposition of the primary diagonal of the scene’s light transport matrix (Figure 6.6(a)). The left part of Figure 6.16(b) shows the measurements for pixels on a horizontal cross-section through the scene, as a function of pathlength. The slanted white stripe in the diffuser region of these measurements corresponds to direct paths (blue in the schematic of Figure 6.16(a)), whose pathlengths increase steadily toward deeper parts of the diffuse wall.
Figure 6.15: Visualizing photoelasticity. (a) We use an LED source to scan a scene consisting of a plastic protractor in front of a mirror. We consider the close-up shown in the box. (b) A standard RGB frame under polarized light with no interference or scanning. The color patterns on the protractor correspond to wavelength and polarization-dependent variations of refractive index due to photoelasticity. (c) Low-resolution (1 mm) measurement of optical pathlength at every pixel, shown color-coded. (d) Zooming on the protractor, top row shows (left to right) the red, green and blue channels from (b); bottom row shows optical pathlength measurements for the same three channels and at a 10 μm pathlength resolution.

Similarly, the slanted rainbow stipe in the mirror region corresponds to longer, retroreflection paths (orange in the schematic) that are additionally affected by the spectral dispersion of the glass slab. The right of Figure 6.16(b) shows 2D image frames from the corresponding transient sequence at two different pathlengths. When played as a transient video, one sees the white stipe move right to left along the diffuser wall, and then sometime later, the rainbow stripe move right to left along the mirror wall.

Next, we place a right-angle mirror pair on the reference arm to measure the anti-diagonal component (Figure 6.6(b)). This component includes paths whose starting and ending locations are symmetric around the glass slab, such as mirror-glass-diffuser and diffuser-glass-mirror paths (purple in the schematic of Figure 6.16(a)). All such paths in this configuration have the same length, producing a horizontal line in the pixel-pathlength
Figure 6.16: Probing different light transport matrix components. (a) We use an LED source to scan at three wavelengths a scene consisting of a thin glass slab placed between a mirror and a diffuser. A frame is shown at the bottom-right. The scene includes three types of paths: direct paths (blue), mirror-diffuser-mirror retroreflection paths (orange), and two-bounce reflection paths (purple). (b) Measurements of the primary diagonal of the light transport matrix, corresponding to direct and retroreflection paths. The left visualizes the measurements $T^m((x, y_0), (x, y_0), \tau)$ over pathlength $\tau$ for a horizontal line of pixels (a pixel-pathlength slice). The right shows image frames $T^m((x, y), (x, y), \tau_0)$ for all pixels $(x, y)$ at two pathlength values. (c) Measurements of the anti-diagonal of the light transport matrix, corresponding to paths with start and end locations symmetric around the glass slab. These are two-bounce reflection paths that, for each wavelength, all have the same optical length. This produces horizontal stripes in the pixel-pathlength slice $T^m((x, y_0), (-x, y_0), \tau)$ (left), and a short sequence of all-bright frames $T^m((x, y), (-x, y), \tau_0)$ in the transient video (right, shown for the green wavelength). The rainbow patterns in (b) and (c) correspond to optical pathlength differences of $15 - 30 \mu m$ between different wavelengths, due to dispersion in the glass slab.

visualization of Figure 6.16(c). This implies that all frames of the corresponding transient sequence are zero, except for a single frame at the common length $\tau_0$ of the reflection paths, which is non-zero at all pixels. This frame is shown in the right part of Figure 6.16(c).
The spread of the rainbow patterns that are induced by dispersion in the glass slab are approximately 15 µm for reflection and 30 µm for retroreflection paths. The frame in Figure 6.16(c) corresponds to the green horizontal line in x-τ space. When played as a video, the transient sequence has analogous frames for the other two colors.

Comparing different pathlength decompositions. We now consider using an SLD source (Figure 6.5(c)). As discussed in Section 6.3, allows measuring pathlength decompositions of the entire light transport matrix, instead of only specific components of the matrix as in the LED case. In Figure 6.17, we use the plastic toy cup scene from Figure 6.9 to compare the two cases.

Figure 6.17(c) shows a pixel-pathlength slice for the dotted vertical line of pixels through the visual field shown in (a). This was measured using an LED source (Figure 6.5(c)) and corresponds to a pathlength decomposition of the light transport diagonal. It includes only direct reflection and backscattering paths—paths that, after scattering, exit the object from the same point at which they entered. Because interreflection and most scattering paths are removed, this decomposition facilitates measuring the 3D shape of the object through Equation 6.29, as shown in Figure 6.9.

Figure 6.17(b) shows the analogous pixel-pathlength slice measured using an SLD source (Figure 6.5(a)). These measurements correspond to a pathlength decomposition of the entire light transport for this scene. Comparing 2D image frames from the two decompositions at the same pathlength (orange insets), we observe that in the SLD case, a much larger part of the frame has high intensity. This is due to the large number of additional scattering paths, other than backscattering, included in the measurements. Moreover, non-zero measurements are present at much longer pathlengths, corresponding to subsequent caustics and specular interreflections.

Light-in-flight visualizations. As discussed in Section 6.3.1, the pathlength decompositions measured by our setup combined with an SLD source, when composed into a video, are
Figure 6.17: Capturing different pathlength decompositions. (a) We scan a scene consisting of a plastic toy cup, and show pixel-pathlength slices for a vertical cross-section of pixels through the visual field. (b) When we capture a pathlength decomposition of the entire light transport matrix for this scene using an SLD, the pixel-pathlength slice includes direct paths, scattering, caustics, and specular interreflections. These are also apparent in the 2D image frames of the transient sequence (orange and blue insets). (c) When we measure the pathlength decomposition only of the primary diagonal of the light transport matrix using an LED, the pixel-pathlength slice has only direct and backscattering paths. Equivalent to light-in-flight visualizations produced by conventional transient imaging. Figure 6.18 shows representative frames from such visualizations for two miniature scenes.
Figure 6.18: Light-in-flight visualizations for two different scenes using an SLD source. For each row, the leftmost column shows a conventional photo of the scene, and columns two to four show representative frames. From top to bottom, the scenes are: chess knight between diffuse wall and mirror; zirconia layer on ground glass plate between diffuse walls.

An additional example is shown in Figure 6.17(a), for the toy cup scene. The field of view in each scene is roughly $2 \times 2 \text{ cm}^2$, and the sequences are captured at a pathlength resolution of $10 \mu\text{m}$, though for visualization they have been downsampled to $20 \mu\text{m}$. This corresponds to a temporal resolution of $70 \text{ fs}$ in vacuum, or 15 trillion frames per second. The project website includes full sequences for these and additional scenes.

The scenes showcase different transport phenomena. The chess knight scene shows qualitatively different paths, including three types of reflections. In the zirconia scene, the central zirconia coating remains bright for several frames, due to the high amount of subsurface scattering. There is also transmission from the rear diffuse walls, through the ground glass that surrounds the central zirconia coating, and back toward the camera. Finally, at the vertex of the diffuse walls, we can more intense diffuse interreflections. The high pathlength resolution of our setup allows us to capture light-in-flight sequences with very crisp and well-localized wavefronts.
6.6 Discussion

We have presented a computational imaging system that decomposes light transport based on spatial layout, pathlength, wavelength, and polarization. The system uses interferometry, and produces decompositions at a pathlength resolution of 10 μm. We have analyzed the various trade-offs in optimizing the system’s performance, and used it to scan scenes with various optical phenomena.

The high-fidelity decompositions produced by the system may be useful for a variety of direct and inverse problems in computer graphics and computer vision. In addition to the heterogeneous inverse scattering problem we have discussed in Chapter 5, we are intrigued by the challenges of developing inverse rendering algorithms for shape, reflectance, illumination, or occluded imaging that take advantage of the information available in these high-resolution light path decompositions.
Chapter 7

Conclusions and Future Research

Throughout this thesis, we have established a comprehensive framework for solving the inverse scattering problem. Our framework formulates inverse scattering as an appearance matching problem: First, use an imaging system to collect calibrated (known geometry and illumination) measurements of the unknown medium. Second, identify the medium by searching for the material parameters that, when used to render synthetic images of the medium, can match the measurements.

Our framework falls within the larger paradigm of computational imaging. Rather than treat the measurement and inference stages independently, we seek to co-design them in a way that facilitates solving the objective function of the appearance matching problem. In particular, our contributions can be summarized as follows.

In Chapter 2, we identified the two main limitations shared by all previous inverse scattering algorithm: First, their inapplicability in cases where mid-order scattering is significant. Second, the complete lack of algorithms that can recover high-resolution measurements of phase functions of arbitrary shape. In addition, we provided twofold motivation for developing an inverse scattering framework that can tackle these issues. We first showed phenomenologically that prior low-dimensional phase function models cannot accurately
model many common materials. We then demonstrated that variations in the shape of the phase function and mid-order scattering can produce perceptually significant changes in translucent appearance.

In Chapter 3, we provided a rigorous discussion of the radiative transfer framework for describing light propagation inside scattering media. We used this framework to mathematically define different types of images in terms of radiometric quantities inside a scattering medium. We also used it to provide a formal definition of inverse scattering as an appearance matching optimization problem.

In Chapter 4, we presented an inverse rendering algorithm for solving the inverse scattering appearance matching problem in homogeneous media. The algorithm is based on a combination of stochastic optimization, Monte Carlo rendering, and an operator-theoretic formulation of radiative transfer. Taking as input steady-state measurements of a homogeneous medium, it can accurately recover all of its material parameters, including high-resolution phase function shape, while accounting for all orders of scattering, low, mid, and high. We demonstrated these characteristics by measuring validation materials with ground-truth parameters, as well as a diverse set of common materials, many of which had never been measured before.

In Chapter 5, we extended our inverse rendering framework, to allow using measurements from more general types of imaging as input, and returning spatially-varying material parameters for general heterogeneous media as output. Key to this extension was re-deriving our algorithm from the path-based formulation of radiative transfer underlying all physically-accurate rendering algorithms. Using this theoretical and algorithmic framework, we also provided theoretical and empirical results that can be used to guide the design of imaging systems for inverse scattering applications.

Finally, in Chapter 6, we devised, designed, and implemented a computational imaging system that uses interferometry to implement different types of imaging, corresponding to
different ways to decompose light paths (in terms of length, spatial endpoints, and so on). The use of interferometry allows capturing these measurements at micron-resolution, three orders of magnitude than what was previously possible. These high-resolution, light path decomposition measurements are realizations of the measurements suggested by the results of Chapter 5 as optimal for inverse scattering.

7.1 Future Directions

Despite the generality of the inverse scattering framework developed in this thesis, both in terms of the image measurements it can process and the types of material parameters it can accurately infer, there are still several aspects of the inverse scattering problem that remain unsolved and provide for exciting future research directions.

First, the generality and accuracy of our inverse scattering framework comes at a high cost. On the imaging side, all of the imaging systems we discussed require sophisticated optical hardware, and long acquisition times. On the computation side, our stochastic optimization frameworks require very long runtimes, even deployed on clusters of thousands of computing cores. These can be prohibitive for many real-world applications, where there is limited access to imaging hardware and acquisition and post-processing time. It would be interesting to develop inverse scattering systems that limit generality and accuracy in return for reduced time and simpler imaging. For very optically thin or thick media, there is already a mature literature of such methods, in the form of single-scattering and diffusion algorithms, respectively. For turbid media, the discussion of Section 2.2 and the findings of [44] suggest that simple imaging configurations, such as point-spread functions and geometric edges, provide rich information for all scattering material parameters. Combining such configurations with simple feature extraction and data-driven inference algorithms could lead to the development of scatterometers with low hardware complexity and computational requirements.
Second, and related to the above, the computational imaging framework we presented is an active acquisition system: It requires that we have full control of illumination, camera configuration, and potentially even material shape; and that we can perform highly-accurate calibration of these aspects of the scene. In this thesis, we have not explored how sensitive our appearance matching framework is to calibration errors. Systematically evaluating these sensitivities would provide useful knowledge for engineers of inverse scattering acquisition systems. Moreover, there has been very little exploration of the extent to which inverse scattering can be performed passively, by assuming very little control and knowledge of the scene. This lack of progress can be attributed to the very high complexity of light transport in scattering materials. Instead, computer vision research over the past three decades has focused on the reconstruction of other scene parameters, such as shape and (Lambertian) reflectance, under an opaque world assumption that greatly simplifies the physics of light transport. As computer vision algorithms are becoming more and more adept at solving these scene reconstruction tasks, making further progress will require relaxing this assumption and accounting for scattering. This can help improve the reconstruction fidelity for other scene parameters such as shape [28, 65], or even allow for passive scatterometry from a few images taken “in the wild”. As before, enabling these capabilities will require combining physics-based modeling and inverse rendering algorithms, with statistical modeling and learning algorithms tailored to translucency. The latter can be based on both phenomenological observations such as those in Chapter 2 and [44], and data-driven approaches.

Going beyond inverse scattering, a full (optical) characterization of a real material requires determining many characteristics beyond its scattering parameters. For instance, in this thesis we briefly considered its reflectance (BSDF) and index of refraction, which we assumed are both known; this assumption is unrealistic for many applications. Like scattering parameters, reflectance and index of refraction can be spatially-varying. Additionally, all these material parameters typically vary with wavelength and polarization, including non-linear interactions (fluorescence, birefringence). As discussed in the introduction,
inferring each of the scattering, reflectance, and refractive properties in isolation corresponds to the inverse problems of scatterometry, reflectometry, and refractometry, respectively. Even though there is a significant body of work towards solving each of these problems (including this thesis for scatterometry) on its own, recovering all three types of parameters simultaneously can only be done under severely limiting assumptions (homogeneity, thin media, and so on). In principle, the appearance matching formulation of Equation 3.25 can be extended to solve this joint, inverse material rendering problem, simply by adding more minimization unknowns. Likewise, the combination of stochastic optimization and Monte Carlo rendering we used for inverse scattering can theoretically also be used to perform this minimization (at least to the extent that we can continue using light transport and geometric optics to model light propagation). In practice, to extend our framework to such fully general inverse material rendering problems would require addressing challenges similar to those encountered in the transition from homogeneous (Chapter 4) to heterogeneous (Chapter 5) inverse scattering. Namely, we would first need to utilize different types of imaging that provide sufficient information for determining the much larger number of, often ambiguous, unknown material parameters. This information could be obtained by considering measurements involving polarization and spectral diversity, as well as sophisticated geometric configurations for the sources, sensors, and samples. Then, we would need to scale up our optimization algorithm, to handle the dimensionality explosion in the number of both inputs (measurements) and outputs (material parameters).

Finally, we mention that the Monte Carlo gradient rendering procedure of Chapter 5 is a computational framework that, when combined with gradient-based optimization, can be used in a plethora of inverse problems from image measurements, beyond those involving recovery of material parameters. These could include shape and depth, location and orientation of scene objects or the camera, and illumination. This procedure has two primary advantages compared to alternatives for evaluating derivatives of images with respect to scene parameters. The first advantage is generality: As shown in Section 5.1, it can be applied to measurements captured with very general imaging models that describe a wide array
of real sensing systems. Moreover, the throughput function can be differentiated for all possible scene parameters; even when analytic differentiation is not possible, its gradient can be evaluated exactly using automatic differentiation [48]. The second advantage is physical accuracy: It automatically accounts for all global illumination effects such as interreflections, occlusions, or, as demonstrated in this thesis, scattering; and it does so for all possible scenes. The main disadvantage of the Monte Carlo gradient rendering procedure is its high computational cost. This is especially problematic in cases where the paths with non-zero contributions to the image gradient are a lower-dimensional manifold of the ambient path space; these include differentiation with respect to surface geometry. The above advantages and disadvantages are completely analogous to those of Monte Carlo techniques for forward problems (forward rendering). For the inverse case, the increased computational cost is mitigated by the fact that stochastic optimization allows taking advantage of even very noisy gradient evaluation. As in the forward case, making Monte Carlo gradient rendering methods practical will require devising sampling techniques that produce paths that contribute strongly to the image gradient. This motivates a line of research where existing path sampling techniques used for rendering images, such as Metropolis Light Transport, are adapted for computing image gradients.
Appendix A

Comparison of Inverse Rendering Frameworks

In this appendix, we compare the inverse rendering frameworks introduced in Chapters 4 and 5.

A.1 Operator and Path Formulations

A first difference between the frameworks relates to the mathematical machinery involved. The starting point for the framework of Chapter 4 is an operator-theoretic expression for radiance inside scattering volumes, whereas the derivation of the framework of Chapter 5 starts from a path-integral expression. This difference is, however, superficial: it is possible to express both frameworks using path-integral and operator-theoretic formulations, as shown by the following two claims.

Claim 8. For any step-size $h$, we define an $n$-step path $\bar{x}$ as an ordered sequence of points in the medium $\mathcal{M}$,

$$\bar{x} = x_0 \xrightarrow{h} x_1 \xrightarrow{h} \ldots \xrightarrow{h} x_n,$$  \hspace{1cm} (A.1)
where $\overrightarrow{h}$ indicates that each segment of the path has length $h$,

$$\|x_j - x_{j-1}\|_2 = h, \quad \forall j \in \{1, \ldots, n\}. \quad (A.2)$$

We denote the space of all such paths as $\mathcal{T}_n$. If $h \ll 1 / \sigma_l(x)$ for all $x \in \mathcal{M}$, then any entry of the light transport matrix $T_m$ at pathlength $\tau = nh$ can be written as

$$T_m(X_o, X_i, nh) = \int_{\mathcal{T}_n} \Delta(\vec{x}, X_o, X_i) s_m(\vec{x}) \, d\vec{x}, \quad (A.3)$$

where

$$\Delta(\vec{x}, X_o, X_i) = \delta(x(X_i) - o(\vec{x})) \cdot \delta(\omega(X_i) - \omega_o(\vec{x}))$$

$$\delta(x(X_o) - e(\vec{x})) \cdot \delta(\omega(X_o) - \omega_e(\vec{x})). \quad (A.4)$$

and the term $s_m$ determines the path’s radiance contribution,

$$s_m(\vec{x}) = \prod_{j=1}^{n-1} s_m \left( x_{j-1} \xrightarrow{h} x_j \xrightarrow{h} x_{j+1} \right), \quad (A.5)$$

$$s_m \left( x_{j-1} \xrightarrow{h} x_j \xrightarrow{h} x_{j+1} \right) =$$

$$\begin{cases} 
1 - h \sigma_l(x_{j-1}) + h \sigma_s(x_{j-1}) f_p(x_{j-1}, 0), & \omega(x_{j-1} \xrightarrow{h} x_j) = \omega(x_j \xrightarrow{h} x_{j+1}), \\
\sigma_s(x_{j-1}) f_p(x_{j-1}, \omega(x_{j-1} \xrightarrow{h} x_j)) \cdot \omega(x_j \xrightarrow{h} x_{j+1}), & \text{otherwise}. 
\end{cases} \quad (A.6)$$

Claim 9. We define the input term

$$L_{V,j}(x, \omega) \triangleq \exp(-\sigma_l r_{\partial\mathcal{M}}(x, \omega)) L(x - r_{\partial\mathcal{M}}(x, \omega) \omega, \omega) + \int_{0}^{r_{\partial\mathcal{M}}(x, \omega)} \exp(-\sigma_l r') Q(x - r' \omega, \omega) \, dr',$$

the single-scattering operator (sometimes called single-bounce operator)

$$B_k(L)(x, \omega) \triangleq \sigma_s \int_{0}^{r_{\partial\mathcal{M}}(x, \omega)} \exp(-\sigma_l r')$$

$$\int_{S^2} p(\psi^T \omega) L(x - r' \omega, \psi) \, d\mu(\psi) \, dr'.$$
and the corresponding volume rendering operator

\[ V_k \triangleq (I - B_k)^{-1} = \sum_{j=0}^{\infty} B_k^j, \]  

(A.9)

where in Equation (A.9) we again used the Neumann series expansion for the second equality. Then, we can use these operators to rewrite the volume rendering equation (3.15) in operator form as

\[ L = L_{V,i} + B_k L \Rightarrow \]
\[ L = (I - B_k)^{-1} L_{V,i} \Rightarrow \]
\[ L = V_k L_{V,i} \]  

(A.10)

Claim 8 can be derived by recursively expanding radiance \( L \) in Equation (4.9), analogous to the derivation of Equation (5.2). The integration is performed over paths consisting of segments of equal length \( h \). At the end of each such segment, a photon will undergo a propagation event and will: 1) continue traveling in the same direction (probability \( 1 - h\sigma_t + h\sigma_s f_p(0) \)); or 2) scatter towards a new direction determined by the local phase function (probability \( h\sigma_s \)). The length of a path is determined completely by the number of such segments, \( \tau(\vec{x}) = nh \). The derivation of Claim 9 can be found in [5, 72].

A.2 Pathlength and Bounce Separations

The more fundamental difference between the two formulations relates to the types of paths corresponding to consecutive applications of the respective operator. In the case of the operator \( K \), the term \( K^n L_i^z \) is the accumulation of contributions from photons that have travelled paths that all share the same length \( nh \). On the other hand, the term \( B^n L_i^z \) corresponds to contributions from photons that have travelled paths that share the same number \( n \) of scattering events (“bounces”). For this reason, we will refer to the framework of Chapter 4 as the pathlength-based framework, and to that of Chapter 5 as the bounce-based framework.
The above characteristic of the pathlength-based framework allows greatly simplifying the expression of pathlength decomposed measurements. As discussed in Section 4.1, entries of the light transport matrix and measurements at \( \tau = nh \) can be expressed in the simple form \( D K^n L_i \); that is, the pathlength \( \tau \) (up to discretization) directly determines a unique power of operator \( K \).

This can also be seen from Equation A.3: the integration is over paths consisting of \( n \) segments of fixed length \( n \); and the path sampling function \( D \) is independent of pathlength and need only reject paths based on their endpoints. By comparison, in Equation (5.2) of the bounce-based formulation, the integration is performed over paths of arbitrary length, and consequently \( W \) is a function of pathlength.

As implied by the name, the bounce-based framework simplifies expressions relating to a hypothetical bounce-decomposed imaging, where each measured image corresponds to contributions from photons that have undergone a specific number of bounces. In turn, this number of bounces would directly determine a unique power of operator \( B \).

### A.3 Practical Considerations

When deciding between the two frameworks for inverse scattering applications, there are two main differences that need to be taken into account.

First, the operator \( K \) of the pathlength-based framework depends linearly on both \( \sigma_t \) and the product \( \sigma_s \cdot f_p \). This enables the dictionary parameterization of Section 4.2.1, where a single-set set of weights can be used to represent all material parameters. In particular, this property manifests in Claim 1, which greatly simplifies the operator-theoretic gradient expressions. The operator \( B \), on the other hand, depends on \( \sigma_t \) through the exponential of the volumetric attenuation term. As a result, such a dictionary representation is only possible for the product \( \sigma_s \cdot f_p \), and cannot be used to express \( \sigma_t \).
Second the number of terms in the product expression of $s_m$ for the pathlength-based formulation can be considerably larger than that in the expression of $\tilde{f}_m$ for the bounce-based formulations respectively. This is because the pathlength-based formulations breaks a continuous segment between scattering events into multiple smaller segments of equal length $h$. The practical consequence of this is that sampling and shading paths for the pathlength-based operation is computationally much more expensive than for the bounce-based operation.

The above two considerations inform our choice to use the pathlength-based framework for homogeneous, and the bounce-based framework for heterogeneous inverse scattering. In the homogeneous case, where we optimize over a small number of parameters, the linear dependence of operator $K$ allows us to use compact and very expressive dictionary parameterizations. On the other hand, in the heterogeneous case, the very large number of unknowns makes this computationally prohibitive, and therefore we use the bounce-based formulation.
Appendix B

Appendix to Chapter 2

In this appendix, we discuss several details related to Chapter 2. We group these results into sections corresponding to the sections of the chapter.

B.1 Characterizing Translucent Appearance

In Figure B.1, we show an example image of the scene “Lucy + Campus” rendered with phase function of average cosine value $\bar{C} > 0.8$ that we removed from the set of phase functions we use in all our experiments.

B.2 Computational Experiment

B.2.1 Embeddings Using the $L_1$-norm Metric

In Figure B.2, we show the two dimensional embeddings produced using the $L_1$-norm metric for all the nine scenes. In Tables B.1-B.3, we report numerical values for each of the
Figure B.1: Image of the scene “Lucy + Campus” rendered with phase functions of average cosine value $C > 0.8$.

three embedding similarity measures and all the scene pairs, for embeddings produced with the $L_1$-norm metric.

B.2.2 Embeddings Using the $L_2$-norm Metric

In Figure B.3, we show the two dimensional embeddings produced using the $L_2$-norm metric for all of the nine scenes. In Tables B.4-B.6, we report numerical values for each of the three embedding similarity measures and all the scene pairs, for embeddings produced with the $L_2$-norm metric.

B.2.3 Embedding Using Mean-Opinion-Score Metric

In Figure B.4, we show the two-dimensional embeddings produced when using the mean-opinion-score metric by [94] for all of the nine scenes, using only the 40 phase functions selected by the clustering approach we describe in Section 6 of the main paper. We use
Figure B.2: Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the $L_1$-norm image metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

the implementation that the authors have made publicly available\(^1\). In Figure B.5, we compare with the corresponding embeddings produced using the cubic root metric. We observe that the geometric structure of the embeddings produced by the two metrics is very similar. For corresponding embeddings produced by the two different metrics, using a linear regression hypothesis test, we found that, for both coordinates, the hypothesis of linear relation between them is statistically significant at the 99% confidence level.

\(^1\)http://hdrvdp.sourceforge.net/
<table>
<thead>
<tr>
<th>(1) “Lucy + Campus”</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.021</td>
<td>0.035</td>
<td>0.020</td>
<td>0.006</td>
<td>0.018</td>
<td>0.116</td>
<td>0.138</td>
<td>0.415</td>
</tr>
<tr>
<td>“Lucy + Dragon”</td>
<td>0</td>
<td>0.006</td>
<td>0.010</td>
<td>0.043</td>
<td>0.066</td>
<td>0.191</td>
<td>0.169</td>
<td>0.540</td>
</tr>
<tr>
<td>“Lucy + Buddha”</td>
<td>0</td>
<td>0.010</td>
<td>0.064</td>
<td>0.096</td>
<td>0.214</td>
<td>0.185</td>
<td>0.588</td>
<td></td>
</tr>
<tr>
<td>“Lucy + Candle”</td>
<td>0</td>
<td>0.045</td>
<td>0.072</td>
<td>0.171</td>
<td>0.152</td>
<td>0.538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“optically thin material”</td>
<td>0</td>
<td>0.006</td>
<td>0.104</td>
<td>0.151</td>
<td>0.357</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“low absorption material”</td>
<td>0</td>
<td>0.106</td>
<td>0.167</td>
<td>0.314</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Eucalyptus”</td>
<td>0</td>
<td>0.046</td>
<td>0.189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + St. Peter’s”</td>
<td>0</td>
<td>0.321</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Dining room”</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.1: Pairwise Procrustes distances for the embeddings of Figure B.2.

<table>
<thead>
<tr>
<th>(1) “Lucy + Campus”</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.987</td>
<td>0.971</td>
<td>0.984</td>
<td>0.998</td>
<td>0.992</td>
<td>0.994</td>
<td>0.974</td>
<td>0.971</td>
</tr>
<tr>
<td>“Lucy + Dragon”</td>
<td>1</td>
<td>0.994</td>
<td>0.982</td>
<td>0.985</td>
<td>0.966</td>
<td>0.994</td>
<td>0.989</td>
<td>0.953</td>
</tr>
<tr>
<td>“Lucy + Buddha”</td>
<td>1</td>
<td>0.980</td>
<td>0.966</td>
<td>0.937</td>
<td>0.980</td>
<td>0.984</td>
<td>0.916</td>
<td></td>
</tr>
<tr>
<td>“Lucy + Candle”</td>
<td>1</td>
<td>0.976</td>
<td>0.957</td>
<td>0.977</td>
<td>0.963</td>
<td>0.919</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“optically thin material”</td>
<td>1</td>
<td>0.994</td>
<td>0.992</td>
<td>0.970</td>
<td>0.976</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“low absorption material”</td>
<td>1</td>
<td>0.983</td>
<td>0.954</td>
<td>0.987</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Eucalyptus”</td>
<td>1</td>
<td>0.992</td>
<td>0.974</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + St. Peter’s”</td>
<td>1</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Dining room”</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.2: Pairwise Pearson’s correlation coefficient for the vertical coordinate of the embeddings of Figure B.2.

<table>
<thead>
<tr>
<th>(1) “Lucy + Campus”</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.997</td>
<td>0.995</td>
<td>0.995</td>
<td>0.997</td>
<td>0.993</td>
<td>0.994</td>
<td>0.955</td>
<td>0.866</td>
</tr>
<tr>
<td>“Lucy + Dragon”</td>
<td>1</td>
<td>0.998</td>
<td>0.998</td>
<td>0.990</td>
<td>0.983</td>
<td>0.993</td>
<td>0.973</td>
<td>0.831</td>
</tr>
<tr>
<td>“Lucy + Buddha”</td>
<td>1</td>
<td>0.999</td>
<td>0.985</td>
<td>0.976</td>
<td>0.996</td>
<td>0.975</td>
<td>0.812</td>
<td></td>
</tr>
<tr>
<td>“Lucy + Candle”</td>
<td>1</td>
<td>0.985</td>
<td>0.976</td>
<td>0.995</td>
<td>0.977</td>
<td>0.811</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“optically thin material”</td>
<td>1</td>
<td>0.998</td>
<td>0.987</td>
<td>0.933</td>
<td>0.898</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“low absorption material”</td>
<td>1</td>
<td>0.978</td>
<td>0.919</td>
<td>0.915</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Eucalyptus”</td>
<td>1</td>
<td>0.966</td>
<td>0.831</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + St. Peter’s”</td>
<td>1</td>
<td>0.695</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Lucy + Dining room”</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.3: Pairwise Pearson’s correlation coefficient for the horizontal coordinate of the embeddings of Figure B.2.
Figure B.3: Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the $L_2$-norm image metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) “Lucy + Campus”</td>
<td>0.006</td>
<td>0.017</td>
<td>0.024</td>
<td>0.005</td>
<td>0.012</td>
<td>0.103</td>
<td>0.101</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td>(2) “Lucy + Dragon”</td>
<td>0.005</td>
<td>0.009</td>
<td>0.020</td>
<td>0.029</td>
<td>0.146</td>
<td>0.124</td>
<td>0.604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) “Lucy + Buddha”</td>
<td>0.002</td>
<td>0.039</td>
<td>0.054</td>
<td>0.171</td>
<td>0.136</td>
<td>0.648</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) “Lucy + Candle”</td>
<td>0.048</td>
<td>0.064</td>
<td>0.188</td>
<td>0.147</td>
<td>0.669</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) “optically thin material”</td>
<td>0.004</td>
<td>0.084</td>
<td>0.104</td>
<td>0.473</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) “low absorption material”</td>
<td>0.091</td>
<td>0.117</td>
<td>0.457</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) “Lucy + Eucalyptus”</td>
<td>0.033</td>
<td>0.264</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) “Lucy + St. Peter’s”</td>
<td>0.382</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) “Lucy + Dining room”</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.4: Pairwise Procrustes distances for the embeddings of Figure B.3.

B.3 Psychophysical Experiments

B.3.1 Clustering for Stimuli Selection

In Figures B.6-B.8, we show all members of three representative clusters, out of the forty clusters produced by the clustering algorithm used for the stimulus image selection. We
show these examples to demonstrate that the clusters includes materials that appear very similar, and can be represented well by an exemplar in a psychophysical experiment.

### B.3.2 Visualization of Embeddings

In Figures B.9-B.12, we show cropped versions of all 40 stimuli images used in each of the two psychophysical experiments, overlayed on the corresponding computational and perceptual embeddings.
Figure B.4: Two-dimensional embeddings of images rendered with the 40 phase functions selected by the clustering algorithm, produced using the mean-opinion-score image metric. The embedded dots, one per rendered image for a total of 40 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them. To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

B.3.3 Bootstrapping Analysis

As a way of assessing the stability of these embeddings, we performed a bootstrapping analysis: we created 5 perturbations of the “Lucy + Campus” data, by selecting for each of those a random subset of 7200 triplets (24% of the total number, using the 75.94% consistency we report in the main paper), and flipping the corresponding user-supplied ratings; we then used each of the perturbed data sets to learn a two-dimensional embedding, in the same way as described above. Figure B.13 compares the embedding using the original data with the embeddings learned from perturbed data. We observe that, despite the differences, the overall structure and ordering of the images in the embedding remain consistent.
Figure B.5: Comparison of two-dimensional embeddings produced by the cubic root and mean-opinion-score metrics, for images rendered with the 40 phase functions selected by the clustering algorithm. The embedded dots, one per rendered image for a total of 40 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them.

B.4 Analysis

B.4.1 Generalized Procrustes Analysis

We provide some discussion of the planar Procrustes analysis techniques that we use in Section 2.3. For more details, we refer to [32].
We will use the term configuration to refer to a set of landmark points in $\mathbb{R}^2$, denoted by

$$\mathcal{E} = \{x_n \in \mathbb{R}^2, n = 1, \ldots, N\}. \quad \text{(B.1)}$$

In the following, all configurations we consider will have the same number of points $N$, and there will be a correspondence between points in different configurations with the same index $n$. In the context of Sections 2.3 and 2.4, each two-dimensional embedding is a configuration, with one embedded point per rendered image for a total of $N = 753$ images; corresponding points across different embeddings represent images of different scenes rendered with the same phase function.

Given two configurations $\mathcal{E}_1$ and $\mathcal{E}_2$, we can align $\mathcal{E}_2$ with $\mathcal{E}_1$ by finding a translation, rotation, and uniform scaling transform such that corresponding points in the two configurations optimally match, under some matching criterion. When the matching criterion is the $l_2$
distance between points, this corresponds to solving the optimization problem

\[
\min_{s \in \mathbb{R}, T \in \text{SO}(2), c \in \mathbb{R}^2} \sum_{n=1}^{N} \|x_{1,n} - (sT x_{2,n} + c)\|^2_2,
\]  

(B.2)

over all scale factors \( s \), rotation matrices \( T \), and translation vectors \( c \). This alignment problem is called the full Procrustes superimposition of \( \mathcal{E}_2 \) onto \( \mathcal{E}_1 \), and has a closed-form solution that we denote by \( P(\mathcal{E}_1, \mathcal{E}_2) \). Note that \( P \) is not symmetric in its two arguments, that is, the transformation for aligning \( \mathcal{E}_2 \) onto \( \mathcal{E}_1 \) is different from the transformation for aligning \( \mathcal{E}_1 \) onto \( \mathcal{E}_2 \).

Given any configuration \( \mathcal{E} \), we can normalize it through a uniform scaling and translation such that

\[
\frac{1}{N} \sum_{n=1}^{N} \|x_n\|^2_2 = 1, \quad \sum_{n=1}^{N} x_n = 0.
\]  

(B.3)

For two configurations \( \mathcal{E}_1 \) and \( \mathcal{E}_2 \), one way to compare their similarity is first to normalize them, and then compute the minimum of the objective function of (B.2). The square root

**Figure B.7:** Member images of representative cluster 2.
Figure B.8: Member images of representative cluster 3.

of this minimum value, which we denote by $d_P (E_1, E_2)$, is called the full Procrustes distance between $E_1$ and $E_2$. Due to the normalization step, this is a scale-independent measure of shape similarity. Furthermore, $d_P (E_1, E_2)$ is symmetric in its two arguments, unlike the corresponding full Procrustes superimpositions transformations

Given a set of configurations $\{E_1, \ldots, E_K\}$, we can define their full Procrustes mean as the solution of the optimization problem

$$
\min_{\mathcal{E}} \frac{1}{K} \sum_{k=1}^{K} d_P^2 (\mathcal{E}, E_k). \tag{B.4}
$$

This problem also has a closed-form solution. The square root of the minimum of the objective function of (B.4) is called the root-mean-square of full Procrustes distances for this set of embeddings. The full Procrustes mean configuration and root-mean-square of full Procrustes distances are measures of central tendency and variability for sets of shapes analogous to sample arithmetic mean and standard deviation for sets of real numbers. The
procedure of solving the optimization problem (B.4) is called generalized Procrustes analysis.

We use generalized Procrustes analysis to compute the full Procrustes mean of the embeddings for the nine scenes produced using the cubic root metric (Figure 4 of the main paper). The mean embedding is shown in Figure B.14, where it is also compared with the embedding for “Lucy + Campus”. In Tables B.7 and B.8, we also report the results of the same correlation analysis as in Section 7 of the main paper for parameterizing the two dimensions of the embedding, as applied to the mean embedding.

Figure B.9: Visualization of computational embedding for sidelighting experiment.
Figure B.10: Visualization of perceptual embedding for sidelighting experiment.

Figure B.11: Visualization of computational embedding for backlighting experiment.
Figure B.12: Visualization of perceptual embedding for backlighting experiment.

Figure B.13: Bootstrapping analysis: comparison of (a) embedding learned from data provided from the subjects; and (b)-(f) embeddings learned from five random perturbations of the data.

B.4.2 Phase Function Distance Metric

In Figure B.15, we show the two-dimensional weight function learned in the metric learning experiments of Section 7.2 of the main paper. The function is shown as a matrix, whose
Figure B.14: Comparison between (a) the two-dimensional embedding for the scene “Lucy + campus”, produced using the cubic root metric; and (b) the full Procrustes mean of all the embeddings of Figure 4 of the main paper, as produced by generalized Procrustes analysis.

<table>
<thead>
<tr>
<th>Scene</th>
<th>Correlation with $C$</th>
<th>Correlation with $C^2$</th>
<th>Best power $a \in [1.5, 2.5]$</th>
<th>Correlation with $C^0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Procrustes mean</td>
<td>0.9619</td>
<td>0.9942</td>
<td>2.06</td>
<td>0.9943</td>
</tr>
</tbody>
</table>

Table B.7: Parameterization of the vertical dimension of the full Procrustes mean of the embeddings of Figure 4 of the main paper. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.)

<table>
<thead>
<tr>
<th>Scene</th>
<th>Correlation with $1/\sqrt{1 - M_C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Procrustes mean</td>
<td>0.9109</td>
</tr>
</tbody>
</table>

Table B.8: Parameterization of the horizontal dimension of the full Procrustes mean of the embeddings of Figure 4 of the main paper. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.)

dimensions correspond to the axis $\theta_1$ and $\theta_2$. The diagonal of this matrix is shown in the polar plot of Figure 10(a) of the main paper.
Figure B.15: Two-dimensional learned weight function for comparing phase functions, visualized as a matrix. Zooming in shows the high values of the diagonal, which is plotted separately in Figure 10(a) of the main paper.
Appendix C

Appendix to Chapter 3

C.1 A Note on Units

In Equation (3.10), both $L_i$ and $L^s_i$ have units of radiance, and the Dirac delta function has units of time inverted. As a consequence, the two sides of the equation as written are not equal dimensionally. To address this issue, it is necessary to rewrite this equation as

$$L_i (x, \omega, t) = L^s_i (x, \omega) \ T \delta (t), \ (x, \omega) \in \Gamma_i,$$  \hspace{1cm} (C.1)

where $T$ is a constant with units of time. Then, Equations (3.11) and (3.12) become, respectively,

$$L^s (x, \omega) = \frac{1}{T} \int_0^\infty L (x, \omega, t) \ dt,$$  \hspace{1cm} (C.2)

$$T^s (X_o, X_i) = \frac{1}{T} \int_0^\infty T (X_o, X_i, t) \ dt.$$  \hspace{1cm} (C.3)

Similarly for the pathlength versions, the corresponding equations need to include a constant $cT$, where $c$ the speed of light in the medium.

The above issue arises because the definition of the Dirac delta distribution implies that it is not an adimensional quantity, but instead it always has the same dimensions as the inverse
of its arguments. The constant $T$ we use above to resolve this issue can be any arbitrary constant, including one of unit value.

Throughout this thesis, including in the proof in the following section, we assume that Dirac delta distributions are multiplied by a constant of the correct dimensions, as necessary for equations to make sense dimensionally. To avoid confusion and keep notation simple, we do not explicitly show these constants, and instead we assume that they are subsumed in the Dirac delta term.

**C.2 Relationship Between Time-resolved and Stationary Green’s Function**

We prove the following claim relating solutions of the time-dependent and the stationary versions of the radiative transfer equation (RTE).

**Claim 10.** Let $L(x, \omega, t)$ be the solution to the time-dependent RTE

\[
\frac{1}{c} \frac{\partial L(x, \omega, t)}{\partial t} + \omega \cdot \nabla L(x, \omega, t) = -\sigma_t(x) L(x, \omega, t) \\
+ \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) L(x, \psi, t) \, d\psi,
\]

subject to boundary conditions

\[
L(x, \omega, t) = g(x, \omega) \delta(t), \quad (x, \omega) \in \Gamma_i, \quad \text{(C.5)}
\]

\[
L(x, \omega, t) = 0, \quad (x, \omega) \in (\mathcal{M} \times S^2) \setminus \Gamma_i, t = 0, \quad \text{(C.6)}
\]

\[
L(x, \omega, t) = 0, \quad (x, \omega) \in (\mathcal{M} \times S^2), t < 0. \quad \text{(C.7)}
\]

Let $L^s(x, \omega)$ be the solution to the stationary RTE

\[
\omega \cdot \nabla L^s(x, \omega) = -\sigma_t(x) L^s(x, \omega) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) L^s(x, \psi) \, d\psi,
\]

\[
L^s(x, \omega) = g(x, \omega), \quad (x, \omega) \in \Gamma_i, \quad \text{(C.8)}
\]
subject to boundary conditions

\[ L^s (x, \omega, t) = g(x, \omega), \quad (x, \omega) \in \Gamma_i. \quad (C.9) \]

Then,

\[ L^s (x, \omega) = \int_0^\infty L(x, \omega, t) \, dt. \quad (C.10) \]

**Proof.** Let \( \tilde{L}(x, \omega, s) = \mathcal{L}\{L(x, \omega, t)\} \) be the Laplace transform of \( L(x, \omega, t) \) with respect to time. Applying the Laplace transform to Equation (C.4), it becomes

\[ \frac{s}{c} \tilde{L}(x, \omega, s) + \omega \cdot \nabla \tilde{L}(x, \omega, s) = -\sigma_t(x) \tilde{L}(x, \omega, s) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) \tilde{L}(x, \psi, s) \, d\psi. \quad (C.11) \]

If we define

\[ \tilde{\sigma}_a(x, s) = \sigma_a(x) + \frac{s}{c}, \quad (C.12) \]
\[ \tilde{\sigma}_t(x, s) = \sigma_s(x) + \tilde{\sigma}_a(x, s), \quad (C.13) \]

then we can rearrange the terms in Equation (C.11) to bring it in a form analogous to the stationary RTE (C.8),

\[ \omega \cdot \nabla \tilde{L}(x, \omega, s) = -\tilde{\sigma}_t(x, s) \tilde{L}(x, \omega, s) + \sigma_s(x) \int_{S^2} f_p(x, \omega \cdot \psi) \tilde{L}(x, \psi, s) \, d\psi. \quad (C.14) \]

As discussed in Section 3.3, the radiative transfer equation can be equivalently rewritten in the form of the volume rendering equation, which we show again here for convenience.

\[ L^s(x, \omega) = a(x, x_{\partial M}(x, \omega)) g(x_{\partial M}(x, \omega), \omega) \]
\[ + \int_{\partial M(x, \omega)} a(x, x - r' \omega) \sigma_s(x - r' \omega) \]
\[ \int_{S^2} f_p(x - r' \omega, \omega \cdot \psi) L^s(x - r' \omega, \psi) \, d\psi \, dr'. \quad (C.15) \]
where for each point \( x \in S \) and direction \( \omega \in S^2 \), we denote
\[
\begin{align*}
 r_{\partial M} & (x, \omega) = \min \{ r : x - r\omega \in \partial M \}, \\
x_{\partial M} & (x, \omega) = x - r_{\partial M} (x, \omega) \omega,
\end{align*}
\] (C.16)

and where \( a(x, y) \) is volumetric attenuation along the line segment connecting \( x, y \in M \),
\[
a(x, y) = \exp \left( -\int_0^{||x-y||} \sigma_t (x - r\omega (y \rightarrow x)) \, dr \right). \] (C.18)

We can similarly rewrite Equation (C.14) into the following equivalent integral form,
\[
\bar{L}(x, \omega, s) = \tilde{a}(x, x_{\partial M} (x, \omega), s) \tilde{L}(x_{\partial M} (x, \omega), \omega, s) \\
+ \int_0^{r_{\partial M}(x, \omega)} \tilde{a}(x, x - r'\omega, s) \sigma_s (x - r'\omega) \\
\int_{S^2} f_p (x - r'\omega, \omega \cdot \psi) \tilde{L}(x - r'\omega, \psi, s) \, d\psi \, dr'. \] (C.19)

In Equation (C.19), we have
\[
\tilde{a}(x, y, s) = \exp \left( -\int_0^{||x-y||} \tilde{\sigma}_t (x - r\omega (y \rightarrow x), s) \, dr \right), \] (C.20)

and using Equations (C.13) and (C.18),
\[
\tilde{a}(x, y, s) = \exp \left( -\frac{s}{c} ||x-y|| \right) a(x, y). \] (C.21)

Using Equation (C.21), we can rewrite (C.19) as
\[
\bar{L}(x, \omega, s) = \exp \left( -\frac{s}{c} r_{\partial M} (x, \omega) \right) a(x, x_{\partial M} (x, \omega)) \tilde{L}(x_{\partial M} (x, \omega), \omega, s) \\
+ \int_0^{r_{\partial M}(x, \omega)} \exp \left( -\frac{s}{c} r' \right) a(x, x - r'\omega) \sigma_s (x - r'\omega) \\
\int_{S^2} f_p (x - r'\omega, \omega \cdot \psi) \tilde{L}(x - r'\omega, \psi, s) \, d\psi \, dr'. \] (C.22)

From the properties of the Laplace transform, we have that for any function \( f(t) \),
\[
\mathcal{L}^{-1} \left\{ \exp \left( -\frac{s}{c} r \right) f(s) \right\} = f\left( t - \frac{r}{c} \right). \] (C.23)

Applying the inverse Laplace transform to Equation (C.22), and using the property of
Equation (C.23) and the boundary condition (C.5), we have

\[
L (x, \omega, t) = a (x, x_{\partial M} (x, \omega)) g (x_{\partial M} (x, \omega), \omega) \delta \left( t - \frac{r_{\partial M} (x, \omega)}{c} \right) \\
+ \int_{0}^{r_{\partial M} (x, \omega)} a (x, x - r' \omega) \sigma_s (x - r' \omega) \\
\int_{S^2} f_p (x - r' \omega, \omega \cdot \psi) L \left( x - r' \omega, \psi, t - \frac{r'}{c} \right) \, d\psi \, dr'.
\]  
\text{(C.24)}

Finally, let

\[
G (x, \omega) = \int_{0}^{\infty} L (x, \omega, t) \, dt.
\]  
\text{(C.25)}

Using the boundary condition of Equation (C.7), we have

\[
\int_{0}^{\infty} L \left( x, \omega, t - \frac{r}{c} \right) \, dt = G (x, \omega), r \geq 0.
\]  
\text{(C.26)}

Then, integrating both sides of Equation (C.24) over time and using equation (C.26), we have

\[
G (x, \omega) = a (x, x_{\partial M} (x, \omega)) g (x_{\partial M} (x, \omega), \omega) \\
+ \int_{0}^{r_{\partial M} (x, \omega)} a (x, x - r' \omega) \sigma_s (x - r' \omega) \\
\int_{S^2} f_p (x - r' \omega, \omega \cdot \psi) G \left( x - r' \omega, \psi \right) \, d\psi \, dr'.
\]  
\text{(C.27)}

Comparing Equation (C.27) with the stationary volume rendering equation (C.15), we see that they are identical. Therefore, we can identify \( L^s (x, \omega) \) with \( G (x, \omega) \).
Appendix D

Appendix to Chapter 4

D.1 Operator-theoretic Framework for Volumetric Light Transport

In the first part of this appendix, we discuss in more detail technical aspects of the operator-theoretic framework for volumetric light transport and its inversion we presented in Sections 4.1 and 4.2. Similar or equivalent results to those we present here have been derived in previous literature on operator-theoretic approaches to light transport [11, 10], for example the boundedness and invertibility of the rendering operator when applied to functions satisfying Hölder conditions [89]. Here, we re-derive many of these results, as necessary for the completeness of the operator-theoretic formulation we use.

D.1.1 Background

We begin by introducing some definitions and results from the fields of functional analysis, operator theory, and analysis on the sphere, as necessary for our discussion. We draw these from [80, 153, 77, 88, 49]. For detailed expositions to these fields, we refer to any of the dedicated mathematical textbooks.
\textbf{\(L^p\) spaces.} We will be concerned with the \(L^p \left( S \times S^2 \right)\), \(p \in [1, \infty)\) spaces of functions, that is, the set of all functions \(f : S \times S^2 \rightarrow \mathbb{R}\) satisfying
\[
\left( \int_S \int_{S^2} |f(x, \omega)|^p \, d\mu(\omega) \, dx \right)^{\frac{1}{p}} < \infty, \tag{D.1}
\]
where \(S\) a convex subset of the Euclidean space \(\mathbb{R}^3\), \(S^2\) is the two-dimensional unit sphere, and \(\mu\) is the usual spherical Lebesgue measure on \(S^2\). In particular, we will use the spaces \(L^1 \left( S \times S^2 \right)\) and \(L^2 \left( S \times S^2 \right)\), for \(p = 1, 2\) respectively. Equipped with the corresponding functional norms,
\[
\|f\|_{L^1(S \times S^2)} \triangleq \int_S \int_{S^2} |f(x, \omega)| \, d\mu(\omega) \, dx, \tag{D.2}
\]
\[
\|f\|_{L^2(S \times S^2)} \triangleq \left( \int_S \int_{S^2} |f(x, \omega)| \, d\mu(\omega) \, dx \right)^{\frac{1}{2}}. \tag{D.3}
\]
for \(L^1 \left( S \times S^2 \right)\) and \(L^2 \left( S \times S^2 \right)\) respectively, each of these spaces is a Banach space. Further equipped with the inner product, for all \(f, g \in L^2 \left( S \times S^2 \right)\),
\[
\langle f, g \rangle_{L^2(S \times S^2)} \triangleq \int_S \int_{S^2} f(x, \omega) \, g(x, \omega) \, d\mu(\omega) \, dx, \tag{D.4}
\]
the space \(L^2 \left( S \times S^2 \right)\) is a Hilbert space. We will also use the Banach spaces \(L^1 \left( S^2 \right)\) and \(L^1 \left( S \right)\), and the Hilbert spaces \(L^2 \left( S^2 \right)\) and \(L^2 \left( S \right)\), defined similarly.

For the Hilbert space \(L^2 \left( S^2 \right)\), the following theorem holds.

\textbf{Theorem 1.} For any function \(f \in L^2 \left( S^2 \right)\), consider the series \(\sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m y_l^m\), where
\[
f_l^m = \langle f, y_l^m \rangle_{L^2(S^2)}, \tag{D.5}
\]
and \(\{y_l^m, l = 0, \ldots, \infty, m = -l, \ldots, l\}\) is the set of (real-valued) spherical harmonics. Then, this series converges in the sense of \(\|\cdot\|_{L^2(S^2)}\) and its limit is \(f\),
\[
f = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m y_l^m. \tag{D.6}
\]
The series of Equation (D.6) is known as the \textit{spherical harmonics expansion} of \(f\). The set of
spherical harmonics is an orthogonal basis for $L^2(S^2)$, analogous to the Fourier basis for square-integrable functions on a real interval.

**Bounded operators and operator norm.** An operator $\mathcal{F} : \mathcal{H} \mapsto \mathcal{H}$, where $\mathcal{H}$ some Banach space, is called bounded if there exists $C \in \mathbb{R}$, $C > 0$ such that, for all $f \in \mathcal{H}$,

$$\| \mathcal{F} (f) \|_{\mathcal{H}} < C \| f \|_{\mathcal{H}}.$$  \hspace{1cm} (D.7)

We denote by $\mathcal{B} (\mathcal{H})$ the set of all such operators, and by $\mathcal{I}$ the unity operator in this set. For each bounded operator $\mathcal{F}$, the smallest $C$ in Equation (D.7) is called its operator norm induced by the Banach space norm $\| \cdot \|_{\mathcal{H}}$, or simply operator norm,

$$\| \mathcal{F} \|_{\text{Op}} \triangleq \sup \left\{ \frac{\| \mathcal{F} (f) \|_{\mathcal{H}}}{\| f \|_{\mathcal{H}}} \mid f \in \mathcal{H}, f \neq 0 \right\}. \hspace{1cm} (D.8)$$

The set $\mathcal{B} (\mathcal{H})$ equipped with the norm $\| \cdot \|_{\text{Op}}$ is a vector space, with the sum, scalar product, and product operations defined as

$$(\mathcal{F} + \mathcal{G}) (f) = \mathcal{F} (f) + \mathcal{G} (f), \hspace{1cm} (D.9)$$

$$(\mathcal{F}) (af) = a \mathcal{F} (f), \hspace{1cm} (D.10)$$

$$(\mathcal{F} \mathcal{G}) (f) = \mathcal{F} (\mathcal{G} (f)), \hspace{1cm} (D.11)$$

where $\mathcal{F}, \mathcal{G} \in \mathcal{B} (\mathcal{H})$ and $a \in \mathbb{R}$. Then, the operator norm satisfies in particular the following properties with regards to these operations

$$\| \mathcal{F} + \mathcal{G} \|_{\text{Op}} \leq \| \mathcal{F} \|_{\text{Op}} + \| \mathcal{G} \|_{\text{Op}}, \hspace{1cm} (D.12)$$

$$\| a \mathcal{F} \|_{\text{Op}} = |a| \| \mathcal{F} \|_{\text{Op}}, \hspace{1cm} (D.13)$$

$$\| \mathcal{F} \mathcal{G} \|_{\text{Op}} \leq \| \mathcal{F} \|_{\text{Op}} \| \mathcal{G} \|_{\text{Op}}. \hspace{1cm} (D.14)$$

Finally, for any $\mathcal{F} \in \mathcal{B} (\mathcal{H})$, we denote by $B (\mathcal{F}, r)$ the sphere of radius $r > 0$ and center $\mathcal{F}$ in $\mathcal{B} (\mathcal{H})$,

$$B (\mathcal{F}, r) = \left\{ \mathcal{G} \in \mathcal{B} (\mathcal{H}), \| \mathcal{F} - \mathcal{G} \|_{\text{Op}} < r \right\}. \hspace{1cm} (D.15)$$
Spectrum, resolvent, and spectral radius. For any \( \mathcal{F} \in \mathcal{B} (\mathcal{H}) \), its spectrum \( \sigma (\mathcal{F}) \) is defined as the set of all \( \lambda \in \mathbb{R} \) for which the operator \( (\lambda I - \mathcal{F}) \) does not have a bounded inverse, where the inverse of an operator is defined in the usual way. Then, for any \( \lambda \notin \sigma (\mathcal{F}) \), the resolvent operator \( R (\mathcal{F}, \lambda) \) is defined as

\[
R (\mathcal{F}, \lambda) \triangleq (\lambda I - \mathcal{F})^{-1}.
\]  

We will in particular be interested in the unit resolvent operator,

\[
R (\mathcal{F}) \triangleq R (\mathcal{F}, 1) = (I - \mathcal{F})^{-1}.
\]  

In the following, when we say that an operator is “invertible”, we implicitly mean boundedly invertible, that is, that the inverse is bounded. Similarly, when we say that its resolvent (for some \( \lambda \)) exists, we implicitly mean exists and is bounded.

For any \( \mathcal{F} \in \mathcal{B} (\mathcal{H}) \), its spectral radius \( r (\mathcal{F}) \) is defined as

\[
r (\mathcal{F}) \triangleq \lim_{n \to \infty} \| \mathcal{F}^n \|^\frac{1}{n} \text{Op}.
\]  

The definition of the spectral radius is independent of the operator norm used in Equation (D.18). We will use the following two properties of the spectral radius.

**Theorem 2.** For any \( \mathcal{F} \in \mathcal{B} (\mathcal{H}) \),

\[
r (\mathcal{F}) \leq \| \mathcal{F} \| \text{Op},
\]  

\[
r (\mathcal{F}) = \sup \{ |\lambda| : \lambda \in \sigma (\mathcal{F}) \}.
\]  

We will also use the following theorem.

**Theorem 3.** For any \( \mathcal{F} \in \mathcal{B} (\mathcal{H}) \) with \( r (\mathcal{F}) < 1 \), the series \( \sum_{n=0}^{\infty} \mathcal{F}^n \) converges absolutely and its limit is equal to the unit resolvent of \( \mathcal{F} \),

\[
\sum_{j=0}^{\infty} \mathcal{F}^j = R (\mathcal{F}) = (I - \mathcal{F})^{-1}.
\]
with equality defined appropriately.

The series in Equation (D.21) is often referred to as the Neumann series. That the unit resolvent operator at the right hand side part of Equation (D.21) exists and is bounded is a consequence of the assumption that \( r (\mathcal{F}) < 1 \), Equation (D.20), and the definition of spectrum.

**Integral operators and isotropic kernels.** Given any function \( K : S^2 \times S^2 \rightarrow \mathbb{R} \), we can define the corresponding integral operator \( \mathcal{C}_K \) on a Banach space \( \mathcal{H} \) of functions on \( S^2 \) by

\[
\mathcal{C}_K (f) (\omega) \equiv \int_{S^2} K (\omega, \psi) f (\psi) \, d\mu (\psi), \quad \forall f \in \mathcal{H}.
\]  

Equation (D.22)

The function \( K \) is called the kernel of the integral operator. When the domain \( \mathcal{H} \) of \( \mathcal{C}_K \) is one of the \( \mathcal{L}^p (S^2) \) spaces, the following theorem applies.

**Theorem 4.** If for a kernel \( K : S^2 \times S^2 \rightarrow \mathbb{R} \) there exists a constant \( C > 0 \) such that

\[
\sup_{\omega \in S^2} \int_{S^2} |K (\omega, \psi)| \, d\mu (\psi) < C,
\]

\[
\sup_{\varphi \in S^2} \int_{S^2} |K (\omega, \varphi)| \, d\mu (\omega) < C,
\]

then for any \( f \in \mathcal{L}^p (S^2) \), \( p \in [1, \infty) \),

\[
\| \mathcal{C}_K (f) \|_{\mathcal{L}^p (S^2)} \leq C \| f \|_{\mathcal{L}^p (S^2)}.
\]

Equation (D.25) is also known as (one form of) Young’s inequality. Equation (D.25), has the following obvious consequences.

**Lemma 1.** An operator \( \mathcal{C}_K \) satisfying the conditions of Theorem 4 with domain any of the spaces \( \mathcal{L}^p (S^2) \), \( p \in [1, \infty) \), has codomain \( \mathcal{L}^p (S^2) \) (\( \mathcal{C}_K : \mathcal{L}^p (S^2) \rightarrow \mathcal{L}^p (S^2) \)), is in \( B (\mathcal{L}^p (S^2)) \) (is bounded), and its corresponding norm is \( \| \mathcal{C}_K \|_{Op} \leq C \).

A kernel \( K \) is called isotropic when it is a function of only the spherical distance between \( \omega \).
and \( \psi \); equivalently, this means that \( K \) can be written in the form

\[
K(\omega, \psi) = \kappa(\omega^T \psi), \quad \forall (\omega, \psi) \in S^2 \times S^2, \tag{D.26}
\]

for some function \( \kappa : [-1, 1] \to \mathbb{R} \) that is called the profile of \( K \). The following theorem applies to integral operators corresponding to such kernels.

**Theorem 5.** Let \( K \) be an isotropic kernel for which its profile \( \kappa \) is integrable,

\[
\int_{-1}^{1} |\kappa(t)| \, dt < \infty. \tag{D.27}
\]

Then, for the corresponding integral operator \( C_K \) and any square-integrable function \( f \in L^2(S^2) \), the following holds

\[
C_K(f)(\omega) = \sum_{l=0}^{\infty} C_l \kappa_l \sum_{m=-l}^{l} f_l^m y_l^m(\omega). \tag{D.28}
\]

where

\[
\kappa_l = \int_{-1}^{1} \kappa(t) P_l(t) \, dt, \quad l \in \{0, \ldots, \infty\} \tag{D.29}
\]

are the coefficients of the expansion of \( \kappa \) in terms of Legendre polynomials \( \{P_l(t), t \in [-1, 1], l = 0, \ldots, \infty\} \); \( f_l^m, l \in \{0, \ldots, \infty\}, m \in \{-l, \ldots, l\} \) are the coefficients of the spherical harmonics expansion of \( f \) (Equation (D.6)); and \( \{C_l, l = 0, \ldots, \infty\} \) are constants independent of \( \kappa \) and \( f \).

Theorem 5 is known as the *Funk-Hecke theorem*. The integral operator \( C_K \) for an isotropic kernel \( K \) is often described as an (isotropic) *spherical convolution*, and Theorem 5 is seen as the analogous of the convolution theorem for functions on \( \mathbb{R} \). Theorem 5 is well known in computer graphics, where it has been used, for instance, for the problems of surface light transport and normal map filtering [133, 121, 53, 9]. In the remaining of this section, we use the kernel \( K \) and its profile \( \kappa \) interchangeably, and write directly \( K(\omega^T \psi) \) for convenience.

### D.1.2 Volume Light Transport Operators

We continue to study the various operators we defined in Section 4.1. In our analysis, we assume that light fields \( L \) are functions in either the Banach space \( L^1(S \times S^2) \) or the Hilbert
space $\mathcal{L}^2(S \times S^2)$, as is usual in functional analytic treatments of light transport [5, 142],
and establish results for both cases.

We first introduce some notation. We define the following linear operators in terms of the
material parameters $k = \{\sigma_t, \sigma_s, f_p(\theta)\}$,

$$C_{f_p}(f)(\omega) \triangleq \int_{S^2} f_p(\omega^T \psi) f(\psi) \, d\mu(\psi), \quad (D.30)$$

$$A_k(f)(\omega) \triangleq \left( \left( 1 - h \sigma_t \right) I + h \sigma_s C_{f_p} \right)(f)(\omega), \quad (D.31)$$

$$T_{h\omega}(g)(x) \triangleq g(x - h\omega), \quad (D.32)$$

where $C_{f_p}$ and $A_k$ operate on functions defined on $S^2$, and $T_{h\omega}$ operates on functions defined
on the spatial domain $\mathbb{R}^3$. Then, the single-step propagation operator of Equation (4.16) can
be expressed as their composition

$$K_k(L)(x, \omega) = A_k T_{h\omega}(L)(x, \omega) \quad (D.33)$$

**Phase function and its integral operator.** The phase function can, in its most general form,
be any function $f_p : S^2 \times S^2 \mapsto \mathbb{R}^+$ that satisfies appropriate normalization and reciprocity
conditions [70]. Typically though, the phase function is restricted to be isotropic, that is,
invariant to rotations of the incident direction and cylindrically symmetric. Under this
assumption, the phase function is an isotropic kernel,

$$f_p(\omega, \psi) = f_p(\omega \cdot \psi) \quad (D.34)$$

with some profile $f_p : [-1,1] \mapsto \mathbb{R}^+$ that is restricted to take only positive values. Any
function of this form obviously satisfies reciprocity in terms of their two arguments; and
therefore, to be a valid phase function, it only needs additionally to satisfy the normalization
constraint

$$\int_{S^2} f_p(\omega \cdot \psi) \, d\mu(\psi) = 1, \forall \omega \in S^2. \quad (D.35)$$

This condition can be written equivalently in the following two forms (see, for instance,
Lemma 1.3.1, page 9 of [49]),

\[ 2\pi \int_{-1}^{1} f_p(t) \, dt = 1, \quad (D.36) \]

\[ 2\pi \int_{0}^{\pi} f_p(\cos(\theta)) \sin(\theta) \, d\theta = 1, \quad (D.37) \]

where in Equation (D.37) we used \( \theta = \arccos(t) \). In an abuse of notation, \( f_p(\theta) \) is often used instead of \( f_p(\cos(\theta)) \); we also use this convention, which leads us to the normalization condition we showed in Equation (3.19).

For a material with phase function \( f_p(\theta) \), we consider first the operator \( C_p \) of Equation (D.30), as applied to functions in \( L^1(S^2) \) or \( L^2(S^2) \). As can be seen immediately using the terminology we established in Section D.1.1, this is an integral operator of the form of Equation (D.22) for the kernel corresponding to the material phase function. Furthermore, from the normalization condition of Equation (D.35), it is obvious that this kernel satisfies the conditions (D.23)-(D.24) for Theorem 4, with \( C = 1 \). Therefore, from Lemma 1, the operator \( C_p \) is in both \( B(L^1(S^2)) \) and \( B(L^2(S^2)) \) and has norm

\[ \|C_p\|_{\text{Op}} \leq 1 \quad (D.38) \]

**Boundedness of single-step propagation operator.** For a material \( k = \{c_\ell, c_s, f_p(\theta)\} \), we consider now the angular part of its single-step propagation operator, \( A_k \) defined in Equation (D.31), as applied to functions in \( L^1(S^2) \) or \( L^2(S^2) \). From Equations (D.12)-(D.13), we have

\[ \|A_k\|_{\text{Op}} \leq (1 - h\sigma_t) \|I\|_{\text{Op}} + h\sigma_s \|C_p\|_{\text{Op}} \]

\[ \leq (1 - h\sigma_t) + h\sigma_s \]

\[ = 1 - h\sigma_a \quad (D.39) \]

where in Equation (D.39) we used the fact that the operator norm of the unity operator is trivially equal to one and the bound (D.38) we derived, and in Equation (D.40) we used
Both Equations (D.38) and (D.40) are derived for the case of the \( \mathcal{L}^1 (S^2) \) and \( \mathcal{L}^2 (S^2) \) domains (functions of direction only). However, they can be easily extended to the case of the \( \mathcal{L}^1 (S \times S^2) \) and \( \mathcal{L}^2 (S \times S^2) \) (functions of both space and direction) that light fields reside in by a simple application of Fubini’s theorem.

Finally, for the full single-step propagation operator \( \mathcal{K}_k \) defined in Equation (4.11), as applied to functions in \( \mathcal{L}^1 (S \times S^2) \) or \( \mathcal{L}^2 (S \times S^2) \), using Equation (D.14) we have

\[
\| \mathcal{K}_k \|_{\text{Op}} \leq \| A_k \|_{\text{Op}} \| T_{h \omega} \|_{\text{Op}} \leq 1 - h \sigma_a, \tag{D.41}
\]

where we have used the fact that \( \| T_{h \omega} \|_{\text{Op}} = 1 \). Therefore, we have proved that the single scattering operator \( \mathcal{K}_k \) for material \( k \) is in both \( \mathcal{B} ( \mathcal{L}^1 (S \times S^2) ) \) and \( \mathcal{B} ( \mathcal{L}^2 (S \times S^2) ) \).

**Existence of rendering operator.** We now turn our attention to the rendering operator \( \mathcal{R}_k \) defined in Equation (4.16), and we observe that it is equal by definition to the unit resolvent \( \mathcal{R} (\mathcal{K}_k) \), as defined in Equation (D.17), of the single scattering operator.

From Equation (D.19) and the bound of Equation (D.41) we derived earlier, we have that

\[
r (\mathcal{K}_k) \leq 1 - h \sigma_a, \tag{D.42}
\]

For any absorbing material \( k, \sigma_a > 0 \), the above inequality implies that \( r (\mathcal{K}_k) < 1 \). Consequently, from Theorem 3, we have that the rendering operator \( \mathcal{R}_k \) for that material exists, is bounded (is in both \( \mathcal{B} ( \mathcal{L}^1 (S \times S^2) ) \) and \( \mathcal{B} ( \mathcal{L}^2 (S \times S^2) ) \)), and can be represented in terms of the Neumann series, as in the second equality in Equation (4.16).

**D.1.3 Dictionary representation**

In this section, we prove the linearity of the single-step propagation operator with respect to convex combinations of materials. For convenience, we first restate the material mixing
equations. We assume we are given a dictionary of admissible homogeneous materials 
\( D = \{k_m, m = 1, \ldots, M\} \), where 
\( k_m = \{\sigma_{t,m}, \sigma_{s,m}, f_{p,m} (\theta)\} \), and their corresponding single-
step propagation operators \( \{K_{k_m}\} \). Then, for any weight vector \( \pi \) in the \( M \)-dimensional simplex \( \Delta^M \),

\[
\pi = [\pi_m] \in \mathbb{R}^M, \quad \pi_m \geq 0, m = 1, \ldots, M, \quad \sum_{m=1}^{M} \pi_m = 1,
\]  

(D.43)

we can represent a novel mixture material \( k(\pi) = \{\sigma_t(\pi), \sigma_s(\pi), f_p(\theta, \pi)\} \), in terms of the dictionary atoms as

\[
\sigma_t(\pi) \triangleq \sum_{m=1}^{M} \pi_m \sigma_{t,m},
\]  

(D.44)

\[
\sigma_s(\pi) \triangleq \sum_{m=1}^{M} \pi_m \sigma_{s,m},
\]  

(D.45)

\[
f_p(\theta, \pi) \triangleq \frac{\sum_{m=1}^{M} \pi_m \sigma_{s,m} f_{p,m}(\theta)}{\sum_{m=1}^{M} \pi_m \sigma_{s,m}}.
\]  

(D.46)

We restate here for convenience the linear mixing lemma.

**Lemma 2.** For any vector \( \pi \in \Delta^M \), the single-step propagation operator \( K(\pi) \) for the mixed material \( k(\pi) \) defined by Equations (D.44)-(D.46), is a convex combination of the single-step propagation operators of the individual atoms with the same mixing weights,

\[
K(\pi) = \sum_{m=1}^{M} \pi_m K_{k_m}.
\]  

(D.47)

**Proof.** Denoting \( f(x, \psi) = L(x - h\omega, \psi) \), and using Equation (4.11),

\[
\sum_{m=1}^{M} \pi_m K_{k_m} (L) \overset{(4.11)}{=} \left( \sum_{m=1}^{M} \pi_m - h \left( \sum_{m=1}^{M} \pi_m \sigma_{t,m} \right) \right) f
\]

\[
+ h \int_{\mathbb{R}^2} \left( \sum_{m=1}^{M} \pi_m \sigma_{s,m} f_{p,m} \left( \omega^T \psi \right) \right) f(x, \psi) \, d\mu(\psi).
\]  

(D.48)

For the expression of Equation (D.48) to be an operator of the form of Equation (4.11), we
express the integral as

$$
\left( \sum_{m=1}^{M} \pi_m \sigma_{s,m} \right) \int_{\mathbb{S}^2} \frac{\sum_{m=1}^{M} \pi_m \sigma_{s,m} f_{p,m}(\omega^T \psi)}{\sum_{m=1}^{M} \pi_m \sigma_{s,m}} \hat{f}_{\psi}(\pi(\theta)) \, f(x, \psi) \, d\mu(\psi). \quad (D.49)
$$

From Equations (D.48) and (D.49), we get exactly the single-step propagation operator of the material \( k(\pi) \) in Equation (D.47).

By using Lemma 2 and Equations (4.22)-(4.24) and (D.47), we can see that the previous results for \( \mathcal{R} \) directly extend to the case of a dictionary of materials. Specifically, as long as the bulk absorption coefficient of the mixture material, given by \( \sigma_a(\pi) = \sigma_t(\pi) - \sigma_s(\pi) \), is non-zero, then the rendering operator \( \mathcal{R}(\pi) \) exist and is bounded. This can be achieved either by having a purely absorbing (\( \sigma_t > 0 \) and \( \sigma_s = 0 \)) material in the dictionary that is assigned a non-zero mixture weight, or by using a dictionary where all of the atoms have non-zero absorption.

### D.1.4 Differentiation of Rendering Operator

In Section 4.2, we derived the expression of Equation (4.28) for the gradient of the objective function \( E(\pi) \) of the appearance matching optimization problem of Equation (4.25). In the derivation, we used Claim 2 to differentiate the rendering operator \( \mathcal{R}(\pi) \) with respect to the components of the weight vector \( \pi \). In this appendix, we provide the precise technical formulation of the lemma and its proof, as well as show the full derivation of Equation (4.28).

We will use two lemmata. The first follows from Theorem 3.11, page 210 of [80].

**Lemma 3.** For a Banach space \( \mathcal{H} \), if for a bounded operator \( \mathcal{F} \in \mathcal{B}(\mathcal{H}) \) its unit resolvent \( R(\mathcal{F}) \) exists, then there exists some \( r > 0 \) such that, for all operators \( \mathcal{G} \) in the sphere \( B(\mathcal{F}, r) \), their unit resolvent \( R(\mathcal{G}) \) exists. Furthermore, the unit resolvent \( R(\mathcal{G}) \), as a function of \( \mathcal{G} \), is continuous in \( B(\mathcal{F}, r) \), with continuity defined in the sense of \( \| \cdot \|_{Op} \).
The second lemma follows from the second resolvent equation (see for instance page 65 of [77]).

**Lemma 4.** For a Banach space $\mathcal{H}$, if for two bounded operators $\mathcal{F}, \mathcal{G} \in \mathcal{B}(\mathcal{H})$ their unit resolvents $R(\mathcal{F})$ and $R(\mathcal{G})$ exist, then

$$R(\mathcal{F}) - R(\mathcal{G}) = R(\mathcal{F})(\mathcal{F} - \mathcal{G})R(\mathcal{G}).$$

(D.50)

We now restate and prove Lemma 2.

**Lemma 5.** For a Banach space $\mathcal{H}$, let the operator $\mathcal{F}(\pi)$ for any weight vector $\pi \in \mathbb{R}^m$ be defined as $\mathcal{F}(\pi) \triangleq \sum_{m=1}^{M} \pi_m \mathcal{F}_m$, where $\{\mathcal{F}_m \in \mathcal{B}(\mathcal{H}), m = 1, \ldots, M\}$ is a set of bounded linear operators. Then, if its unit resolvent $R(\mathcal{F}(\pi))$ exists, the following differentiation rule holds

$$\frac{\partial}{\partial \pi_m} R(\mathcal{F}(\pi)) = R(\mathcal{F}(\pi)) \mathcal{F}_m R(\mathcal{F}(\pi)).$$

(D.51)

**Proof.** If we denote by $\{e_m \in \mathbb{R}^M, n = 1, \ldots, M\}$ the set of standard basis vectors in $\mathbb{R}^M$, from the definition of $\mathcal{F}(\pi)$ we have for every $\varepsilon \in \mathbb{R}$,

$$\mathcal{F}(\pi + \varepsilon e_m) = \mathcal{F}(\pi) + \varepsilon \mathcal{F}_m \Rightarrow$$

$$\|\mathcal{F}(\pi + \varepsilon e_m) - \mathcal{F}(\pi)\|_{\text{Op}} = |\varepsilon| \|\mathcal{F}_m\|_{\text{Op}},$$

(D.53)

where in (D.53) we used Equation (D.13). As $\mathcal{F}(\pi), R(\mathcal{F}(\pi)) \in \mathcal{B}(\mathcal{H})$, the conditions of Lemma 3 are satisfied. If $r$ is as in that lemma, and given that $\|\mathcal{F}_m\|_{\text{Op}}$ is finite from the boundedness assumption, then for all $|\varepsilon|$ such that

$$|\varepsilon| < \frac{r}{\|\mathcal{F}_m\|_{\text{Op}}},$$

(D.54)

it follows from Equation (D.53) that $\mathcal{F}(\pi + e_m) \in B(\mathcal{F}(\pi), r)$, and from Lemma 3 that the unit resolvent $R(\mathcal{F}(\pi + e_m))$ exists. Also, for any such $\varepsilon$, the conditions of Lemma 4
are satisfied for operators $F(\pi)$ and $F(\pi + \epsilon e_m)$, therefore we have

$$
\Delta_m(\epsilon) \triangleq \frac{R(F(\pi + \epsilon e_m)) - R(F(\pi))}{\epsilon} \\
\overset{(D.50)}{=} \frac{R(F(\pi + \epsilon e_m))(F(\pi + \epsilon e_m) - F(\pi)) R(F(\pi))}{\epsilon} \\
\overset{(D.52)}{=} \frac{R(F(\pi + \epsilon e_m))(\epsilon F_{km}) R(F(\pi))}{\epsilon} \\
= R(F(\pi + \epsilon e_m)) F_m R(F(\pi)). \quad (D.55)
$$

When $\epsilon \to 0$, from Equation (D.53) we have that

$$
\lim_{\epsilon \to 0} F(\pi + \epsilon e_m) = F(\pi), \quad (D.56)
$$

where the convergence is in the sense of $\| \cdot \|_{Op}$. As we established earlier, for all $|\epsilon|$ small enough to satisfy Equation (D.54), Lemma 3 applies for $F(\pi)$ and $F(\pi + \epsilon e_m)$ is in the ball defined by the lemma. Then, using now the continuity statement of the lemma with Equation (D.56), it follows that

$$
\lim_{\epsilon \to 0} R(F(\pi + \epsilon e_m)) = R(F(\pi))^{-1}, \quad (D.57)
$$

and therefore, combining with Equation (D.55),

$$
\lim_{\epsilon \to 0} \Delta(\epsilon) = R(F(\pi)) F_{km} R(F(\pi)), \quad (D.58)
$$

as desired.

We consider now the dictionary single-step propagation and rendering operators, $K(\pi)$ and $R(\pi) = R(K(\pi)) = (I - K(\pi))^{-1}$, respectively. As we established at the end of the previous subsection, under the condition that the equivalent bulk material $k(\pi)$ has non-zero absorption, both of these operators are in both $B(L^1(S \times S^2))$ and $B(L^2(S \times S^2))$. Therefore, the conditions of Lemma 5 are satisfied. Using it, we can derive Equation (4.28)
as follows

\[
\frac{\partial E}{\partial \pi_m} \bigg|_{\pi} = \sum_{n=1}^{N} \frac{\partial}{\partial \pi_m} \left( \mathcal{D}^n (\mathcal{I} - \mathcal{K} (\pi))^{-1} L_{i,n}^s - I^m \right)^2
\]

\[
= \sum_{n=1}^{N} 2 \left( \mathcal{D}^n (\mathcal{I} - \mathcal{K} (\pi))^{-1} L_{i,n}^s - I^m \right) \mathcal{D}^n \left( \frac{\partial}{\partial \pi_m} (\mathcal{I} - \mathcal{K} (\pi))^{-1} L_{i,n}^s \right)
\]

\[= \sum_{n=1}^{N} 2 \left( \mathcal{D}^n (\mathcal{I} - \mathcal{K} (\pi))^{-1} L_{i,n}^s - I^m \right) \mathcal{D}^n \left( (\mathcal{I} - \mathcal{K} (\pi))^{-1} \mathcal{K}_m (\mathcal{I} - \mathcal{K} (\pi))^{-1} L_{i,n}^s \right),
\]

where in using Equation (D.51), we expanded the unit resolvent to its explicit expression.

### D.1.5 Scattering Parameter Recovery and Illumination

We start this section by stating the following lemma.

**Lemma 6.** For a Banach space $\mathcal{H}$, if for an operator $\mathcal{F}$ its unit resolvent $R (\mathcal{F})$ exists, then

\[
R (\mathcal{F}) = \mathcal{I} + R (\mathcal{F}) \mathcal{F}.
\]

Its validity can be seen trivially by using the definition of unit resolvent from Equation (D.17) in Equation (D.60) (this lemma is analogous to a simplified version of the matrix inversion lemma known for finite dimensional matrices).

For the remaining of this section, we constrain our analysis to light fields in the space $\mathcal{L}^2 (S \times S^2)$. We also assume that we have a material $k$ satisfying the absorption condition we derived in Section D.1.2, so that both the single-step propagation and rendering operators, $\mathcal{K}_k$ and $\mathcal{R}_k$ respectively, are in $\mathcal{B} (\mathcal{L}^2 (S \times S^2))$. As established by Equation (D.33), $\mathcal{K}_k$ can be written as the tensor product

\[
\mathcal{K}_k (L) (x, \omega) = \mathcal{T}_{hw} A_k (L) (x, \omega),
\]

of the direction-dependent and independent of the material bulk parameters operator $\mathcal{T}_{hw}$ applying only on the spatial coordinate of $L$, and the dependent on the material parameters
operator $A_k$ applying only on the angular coordinate of $L$. Using this decomposition and Equation (D.60), we obtain for the rendering operator $R_k$

$$R_k(L) (x, \omega) \triangleq (I - K_k)^{-1} (L) (x, \omega)$$

$$= I (L) (x, \omega) + R_k T_{hw} A_k (L) (x, \omega). \quad (D.62)$$

We now turn our attention to the angular domain operator $A_k$. This operator is equal to the weighted sum of the unity operator and the integral operator $C_{fr}$, both applying on the space $L^2(S^2)$. From the discussion at the beginning of Section D.1.2, we see immediately that the operator $C_{fr}$ satisfies all of the conditions of Theorem 5, and therefore we can write

$$C_{fr} (f) (\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_l f_{p,l} f_l^m y_l^m (\omega) \ a.e., \quad (D.63)$$

where $\{p_l, l = 0, \ldots, \infty\}$ is the expansion of the phase function profile in terms of Legendre polynomials. Furthermore, from Theorem 1, we can write

$$I (f) (\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m y_l^m (\omega) \ a.e.. \quad (D.64)$$

Using Equations (D.63) and (D.64) with the definition of $A_k(\pi)$ in Equation (D.31), we obtain

$$A_k (f) (\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \omega_l f_l^m y_l^m (\omega) \ a.e. \quad (D.65)$$

where

$$\omega_l = 1 - h \sigma_l + h \sigma_s f_{p,l}, \ l = 0, \ldots, \infty. \quad (D.66)$$

By combining Equations (D.65) and (D.62), we obtain for the rendering operator $R_k$

$$R_k (L) (x, \omega) = I (L) (x, \omega) + R_k T_{hw} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} \omega_l L (x)_l^m y_l^m (\omega) \ a.e.,$$

where for each $x \in S$, $\{ L (x)_l^m \}$ is the spherical harmonics expansion of $L(x, \omega), \omega \in S^2$.

The final result of the above derivation, Equation (D.67), shows that the rendering operation $R_k$ at a first stage performs a frequency decomposition and filtering of any input light field $L$. 

203
with the filtering weights being related to the material parameters through Equation (D.66). It is also easy to show that this result applies directly to the case of a material dictionary and corresponding single-step propagation and rendering operators, $\mathcal{K}(\pi)$ and $\mathcal{R}(\pi)$ respectively. Indeed, we see immediately from Equation (D.66) that the frequency terms $w_l$ of the decomposition in Equation (D.67) are linear mixtures with weights $\pi$ of the corresponding frequency terms for each the individual dictionary materials.

Equation (D.67) provides us with the following insight: the light sources we use in an inverse volume rendering setup should be designed in a way such that the corresponding input light field $L^i_s(\mathbf{x}, \omega)$ is “broadband” in the angular domain, so that as few of the frequency terms $L(x)m$ in Equation (D.67) as possible are zero (or, intuitively, as fewer “columns” of the rendering operator are cancelled out). This implies that the optimal lighting conditions for inverse volume rendering are functions which are Dirac delta in the angular domain, that is, collimated light sources.

### D.2 Acquisition Setup

In the second part of this appendix, we provide details about our implementation of the proposed acquisition setup, as shown in Figure 4.4, as well as the necessary calibration procedures.

#### D.2.1 Implementation

**Imaging system.** We use a CRi Nuance MSI-FX hyperspectral camera. The camera uses an integrated liquid crystal tunable filter that can be tuned to narrow wavelength bands of full-width-half-maximum of approximately 20 nm, and centered at any wavelength $\lambda \in [420 \text{nm} \ 720 \text{nm}]$. The camera is equipped with a $1392 \times 1040$ pixel 12-bit grayscale sensor.
The camera is equipped with a Nikon 105mm f/2.8D AF Micro-Nikkor macro lens, which we focus at its closest working distance for maximum magnification. The lens is mounted through an F-to-C mount that includes an apochromatic optical element. Overall, this configuration produces an effective reproduction ratio of 1 : 1.60, field of view of 14.25 mm × 10.75 mm, and subtended angle of 4°.

A large depth of field is necessary in order for a sufficient part of the glass cells to remain in focus when imaging them at large viewing angles, and across all wavelengths (chromatic aberration means that the in-focus part of the field of view shifts as wavelength changes). For this reason, we use an f/16 aperture setting in all of our measurements. We avoided using smaller apertures, as that resulted in very noticeable diffraction artifacts.

**Light generation and control.** We use three separate continuous-wave laser sources for each of the three wavelengths we use for imaging. For the blue wavelength, 488 nm, we use a laser diode pigtailed to a fiber cable. For the green, 532 nm, and red, 635 nm, wavelengths, we use separate laser modules that output collimated beams, which we then couple into fiber cables using two micropositioner-and-lens modules. We use single-mode, non-polarization preserving 460-HP fiber cables in all of our setup.

The fiber-coupled outputs of our three laser sources are combined into multichromatic white light coupled into a single fiber using an off-the-shelf RGB combiner designed for the three wavelengths of our sources. The output of the RGB combiner is connected through a fiber cable to the input of a 1 × 2 computer-controlled optical MEMS switch. The outputs of the switch are connected through fiber cables to two collimator modules. For collimation, we use reflective collimators, each employing a 90° off-axis parabolic mirror element to achieve achromatic collimation.

The entire optical circuit is shown in Figure D.1. Our use of coherent laser sources is motivated by the need to simultaneously achieve: 1) thin and very-well collimated beams that we can easily calibrate and simulate; and 2) high output power to enable us to measure
optically thick materials. In our setup, we produce beams of diameter 1 mm, divergence of 0.05°, and power of approximately 1 mW at each of the three wavelengths we use.

**Mechanical control.** We use two computer-controlled rotation stages in our acquisition setup, as shown in Figure 4.4. The bottom stage is used to rotate both the material sample and collimator configuration relative to the camera, thus controlling the viewing direction \( \theta_o \). It has an angular resolution of 0.001°, and produces rotations repeatable within 0.02°. It is important to note that repeatability of the bottom rotation stage is not critical: as discussed in the following subsection, its position is geometrically calibrated to within subpixel error after each rotation.

The top rotation stage is used to rotate the two collimators relative to the material sample,
thus controlling the lighting directions \( \theta_f, \theta_b \). This stage has the same resolution of 0.001 °, but a much higher repeatability, producing rotations repeatable within 0.003 °, which we have measured to correspond to subpixel error. Repeatability of this stage is important, as the positions of the collimators are geometrically calibrated only once relative to a fixed plane in the 3D scene, and their locations relative to that plane are afterwards assumed known.

Finally, the platform the material sample is placed on is mounted on a computer-controlled XY-linear translation stage, that is used only during the calibration of the collimator positions, as described in the following. Those produce translations of the sample relative to the rest of the setup at a linear resolution of 0.3 "m and repeatable within 1.5 "m. The repeatability of this stage is also not critical: as discussed in the following subsection, its position is geometrically calibrated to within subpixel error after every translation.

\textbf{Geometry.} To create material samples with very precise geometry, we cast them into rectangular glass spectrophotometer cells. The cells are manufactured from B270 borosilicate crown optical glass, using a fully fused method. This process results in glass cells with the following specifications: surfaces parallel to less than 0.05 °; surface flatness less than 4 Newtonfringes; dimension tolerances less than 10 "m.

\textbf{D.2.2 Calibration}

\textbf{Geometric calibration.} For the geometric calibration of our capture setup, we use two planar targets, one opaque and one a thin transparency, having a square grid of circular control points of diameter 0.25 mm printed on them with micrometer accuracy. The calibration procedure then is performed in three stages. First, we capture a set of images of the opaque target, at different orientations and positions. Then, we fix the opaque target on the center stage of the setup, where the sample is normally located. We use a pair of mechanized
translation stages to perform translations of the target, parallel to its surface normal. For each such translation, we capture images of the target, first illuminated by ambient light, and then by the two light sources at each of their three rotations (top rotation stage) relative to the center stage. Finally, we replace the opaque target with the cell containing the material sample, with the transparent target adjacent to the cell surface facing the camera. Each time the bottom rotation stage is moved, we capture an image of the sample under ambient light with the target present, as well as images of the sample without the target under laser illumination (which constitute our measurements).

The set of target images are then processed using the algorithm of [60]. This algorithm allows us to recover the intrinsic parameters of the camera, as well as the extrinsic parameters (relative locations of the planar target in the camera-coordinate system) for each target image. From the images of the opaque target under laser illumination and for multiple parallel translations, we can fit a straight line that corresponds to the direction of the corresponding light source. Finally, from the images of the transparent target, we can calibrate the location of the frontal surface of the cell relative to the camera and the rest of the setup.

We performed the calibration of the sources and sample every time we captured a new material. The source direction, controlled by the top (high-repeatability) rotation stage controlling the orientation of the sources relative to the sample, were calibrated once per capture session, for a fixed sample position. The location of the sample relative to the entire setup was calibrated each time the bottom stage was rotated.

**Index of refraction.** To measure the index of refraction, we first rotate the light source so that it points directly into the camera. Then, we place a glass sample in the middle of the optical path between the source and the camera. The positions of the source and sample relative to the camera are calibrated as described above. By measuring the displacement of the peak intensity before and after the placement of the sample, and knowing the index of refraction of the glass cell, we can estimate the index of refraction of the material using
simple geometric optics.

The above procedure is exactly accurate only under the assumption that a material is transparent. For scattering materials, the location of the peak intensity also depends on the scattering parameters of the material. However, if the sample used in the above process is sufficiently thin relative to the optical thickness of the material, the above procedure will still produce accurate estimates. In our experiments, we used cells of width 1 mm for the measurement of the index of refraction. From simulations, we verified that for the range of materials we measure ($\sigma_t < 200 \, \text{mm}^{-1}$), the index of refraction estimates we obtain are always within 0.1 of the true value.

**High-dynamic range imaging.** For all of our measurements, we capture a stack of images using 19 exposure times, equally spaced in a logarithmic scale from 0.1 ms to 26 s. For exposure times smaller than 0.5 s, we average across 10 images in order to reduce the effect of noise caused by Brownian motion in the scattering medium. We then combine the raw, 12-bit images from different exposure times into a single, high-dynamic range (HDR) image. For this, we use a variance-based weighting scheme [56], which we found to perform significantly better than the usual linear weighting. In the conversion, we assume a linear response curve for the camera sensor. We validated the accuracy of this assumption, by capturing images of a planar white diffuse target (Spectralon) under white illumination at different exposure times and comparing the recorded intensities.

**Color calibration.** The parameters we measure correspond to three nearly-monochromatic wavelength bands, and not to any true linear color space. To approximate standard linear RGB images, we apply a fixed $3 \times 3$ linear color transform to the rendered output produced from our parameters. To estimate the color transform to be applied, we use our imaging system to capture images of a Macbeth colorchecker target under two illuminations: a 32-channel hyperspectral cube under broadband white light; and a 3-channel image under the trichromatic light produced by our light generation system. We white balance the
hyperspectral cube captured under broadband white light to match D65 illumination, and use it to render an RGB image of the colorchecker. Then, we obtain the $3 \times 3$ scaling matrix $A$ for matching the 3-channel image of the colorchecker under trichromatic light to the RGB image under D65 light, by solving a linear least-squares problem.
Appendix E

Appendix to Chapter 5

In this appendix, we prove the two claims presented in Chapter 5 of the thesis.

E.1 Local Similarity Relations in Pathlength-Resolved Case

Claim 11. Let \( \{ a_{n,l}(x, \tau), n > 0, -n \leq l \leq n \} \) be the coefficients of the spherical-harmonics expansion of the solution \( L(x, \omega, \tau) \) of Equation (3.1) at some point \( x \in \mathcal{M} \), and \( \{ f_{p,n}(x), n > 0 \} \) the coefficients of the Legendre expansion of \( f_p \) at that point. If there exists \( N > 0 \) such that \( a_{n,l}(x, \tau) = 0 \) for all \( n > N \), then two materials \( m, m^* \) will produce equal values \( L(x, \omega, \tau) \) if, for \( 1 \leq n \leq N \),

\[
\begin{align*}
\sigma_a(x) &= \sigma_a^*(x), \\
\sigma_s(x)(1 - f_{p,n}(x)) &= \sigma_s^*(x)(1 - f_{p,n}^*(x)).
\end{align*}
\]

Proof. The proof follows easily from Equation (C.14). As the form of this equation is exactly analogous to the stationary RTE of Equation (3.7), except for replacing \( \sigma_a(x) \) with \( \tilde{\sigma}_a(x, s) \) from Equation (C.12). We can then directly apply to it the similarity relations derived for
the stationary case in [150, 156], to obtain,

$$\sigma_a (x (X_o)) + \frac{\sigma}{c} = \sigma_a^* (x (X_o)) + \frac{\sigma}{c^*}, \quad (E.3)$$

$$\sigma_s (x (X_o)) (1 - f_{p,n} (x (X_o))) = \sigma_s^* (x (X_o)) \left(1 - f_{p,n}^* (x (X_o))\right). \quad (E.4)$$

If we assume that the two materials have the same index of refraction, and therefore the same speed of light $c = c^*$ inside them, then Equation (E.3) reduces to Equation (E.1). \[
\]

Note that, according to Equation (E.3), in materials with different index of refraction, the local similarity relations for the absorption coefficient $\sigma_a$ are different in the pathlength-resolved and stationary cases. This relates to the fact that, at different optical pathlengths, light will have travelled different (geometric) distances inside the two materials, and therefore the absorption coefficient must change in order for the volumetric attenuation in the two materials to be the same. In our paper, we do not consider this case, as we assume that the index of refraction is known.
E.2 Non-Local Ambiguities in Pathlength-Resolved Case

Claim 12. Using pathlength decomposition, the configuration of Figure 5.2(a) provides measurements of the form

\[
I_\tau = \begin{cases} 
Q_\tau + \sum_{p \in k(n)} \sigma_s [nh, p] \int_0^\pi f_p ([nh, p], \theta) R_{\tau, p} (\theta) \, d\theta, & \tau = 2nh, \\
S_\tau + \sum_{p \in k(n)} \sigma_t [nh, p] T_{\tau, p}, & \tau = (2n+1)h,
\end{cases} 
\]  
(E.5)

where \(Q_\tau\) and \(R_{\tau, p} (\theta)\) are functions of material parameters \(m [d, p], d < nh\); and \(S_\tau\) and \(T_{\tau, p}\) of \(m [d, p], d < nh; \sigma_s [nh, p]; f_p ([nh, p], \theta)\).

Proof. We prove this claim in the two-dimensional case for simplicity, with the proof being exactly analogous in the three-dimensional case. We refer to Figure E.1 throughout the proof.

From Equations (A.5)-(A.6), the radiance contribution of a path will depend on the material parameters of a specific voxel \(p, d\) only if the path includes a propagation event at a point inside that voxel.

Consider first a path \(\bar{x}\) with endpoints \(x_i, x_o\) on the material boundary and that includes a propagation event at a point \(x_j\) inside the material layer at depth \(d = nh\), as shown in Figure E.1(a). The distance \(v\) of this point from the volume boundary is \(v > (n-1)h\). The lengths of the sub-paths \(x_i \rightarrow h \cdots \rightarrow h x_j, x_j \rightarrow h \cdots \rightarrow h x_o\), are at least equal to \(v\); given that they must be integer multiples of \(h\), we conclude that each of the subpaths must consist of at least \(n\) segments. Then, the entire \(\bar{x}\) must have at least \(2n\) segments. Conversely, given any point \(x_j\) inside the material layer at depth \(d = nh\), we can always connect it to the material boundary with subpaths of \(n\) segments each. Therefore, we have shown that paths \(\bar{x}\) with endpoints on the material boundary and that have a propagation event inside voxels at depth \(d = nh\) have length \(\tau (p) \geq 2nh\).
Consider again a path $\bar{x}$ with endpoints $x_i, x_o$ on the material boundary. We now assume that it includes two consecutive propagation events at points $x_{j-1}, x_j$ inside the material layer at depth $d = nh$, without a change in direction at $x_j$, and with $x_{j+1}$ being at a layer of depth $d < nh$. This is shown in Figure E.1(b). As in the previous case, the distance of $x_{j-1}$ from the boundary is $v_1 > (n-1)h$, and therefore the sub-path $x_i \xrightarrow{h} \ldots \xrightarrow{h} x_{j-1}$ has at least $n$ segments. The distance of $x_{j+1}$ from the boundary is $v_2 > (n-2)h$, and therefore the sub-path $x_j \xrightarrow{h} \ldots \xrightarrow{h} x_o$ has at least $n-1$ segments. Then, together with segments $x_{j-1} \xrightarrow{h} x_j$ and $x_j \xrightarrow{h} x_{j+1}$, the entire $\bar{x}$ must have at least $2n+1$ segments. Conversely, given any points $x_{j-1}, x_j, x_{j+1}$ as above, we can always create a path of $2n+1$ segments with endpoints on the boundary and with $x_{j-1} \xrightarrow{h} x_j \xrightarrow{h} x_{j+1}$ as a subpath. Therefore, we have shown that paths $\bar{x}$ of this kind have length $\tau(p) \geq 2nh$.

Consider now the measurement $I_{2nh}$ captured for pathlength $\tau = 2nh$ from the source-sensor configuration of Figure E.1(c). By definition, $I_{2nh}$ will be equal to $\mathcal{T}_m(X_s, X_s, 2nh)$ and therefore, using Equation (A.3), equal to the sum of radiance contributions from paths $\bar{x} \in \mathbb{T}_{2n}$ with endpoints $x_s$ and appropriate starting and ending directions. All such paths are completely contained inside a circle of center $x_s$ and radius $r = d = nh$. An arc of this circle is shown in Figure E.1(c).

From the previous discussion, all paths contributing to $I_{2nh}$ will have at most one propagation event in layer $d = nh$, and no propagation events inside layer $d = (n+1)h$. For paths $\bar{x}$ that have no propagation events inside layer $d = nh$, Equations (A.5)-(A.6) imply that their radiance contributions $s_m(\bar{x})$ depend only on material parameters of voxels at smaller depths $d < nh$. If we denote by $B \subset \mathbb{T}_{2n}$ the set of all such paths, then we can define

$$Q_{2nh} = \int_B \Delta(\bar{x}, X_o, X_i) s_m(\bar{x}) \, d\bar{x}. \quad (E.6)$$

Based on the above discussion, $Q_{2nh}$ is a function only of material parameters $\{m[d,p], d < nh\}$. For paths $\bar{x}$ that have one propagation event $x_j$ in layer $d = nh$, from Equations (A.5)-(A.6)
we can write their radiance contributions in the form

\[
s_m(\vec{x}) = \left( \prod_{j=1}^{l-1} s_m(\vec{x}_{j-1} \xrightarrow{h} \vec{x}_j \xrightarrow{h} \vec{x}_{j+1}) \right)
\cdot h\sigma_s[d, p] f_p([d, p], \theta)
\cdot \left( \prod_{j=f+1}^{2n-1} s_m(\vec{x}_{j-1} \xrightarrow{h} \vec{x}_j \xrightarrow{h} \vec{x}_{j+1}) \right)
\]

\[
\triangleq \sigma_s[d, p] f_p([d, p], \theta) R_{2nh,\vec{x}}. \tag{E.7}
\]

where \(\{d\}\) is the voxel at layer \(d = nh\) containing point \(\vec{x}_f\), and \(\theta\) is the change in direction at that point (see Figure E.1(c)). The term \(R_{2nh,\vec{x}}\) is a function only of material parameters \(\{m[d, p], d < nh\}\), given that it corresponds to propagation events at voxels of depth \(d < nh\).

If \(C_{(d,p),\theta} \subset \mathbb{T}_{2n}\) is the set of all such paths for a specific voxel \(\{d, p\}\) and angle \(\theta\), then we can write their total radiance contributions as

\[
L_{(d,p),\theta} = \sigma_s[d, p] f_p([d, p], \theta) R_{2nh,p} (\theta) \tag{E.8}
\]

\[
R_{2nh,p} (\theta) \triangleq \int_{C_{(d,p),\theta}} R_{2nh,\vec{x}} \, d\vec{x}. \tag{E.9}
\]

Given that for every voxel \(\{d, p\}\) there will be non-empty sets \(C_{(d,p),\theta} \subset \mathbb{T}_{2n}\) for multiple values \(\theta\), we can write the total radiance contributions for all paths with \(\vec{x}_f\) in \(\{d, p\}\) as,

\[
L_{\{d,p\}} = \int_0^\pi L_{\{d,p\},\theta} \, d\theta = \sigma_s[d, p] \int_0^\pi f_p([d, p], \theta) R_{2nh,p} (\theta) \, d\theta. \tag{E.10}
\]

The term \(L_{\{d,p\},\theta}\) will be non-zero for all values of \(p\) for which the voxel intersects with the circle in Figure E.1(c). If we denote by \(k(n)\) the set of such values, we can express the contributions of all paths that have one propagation event at layer \(d = nh\) as,

\[
L_d = \sum_{p \in k(n)} L_{\{d,p\}} = \sum_{p \in k(n)} \sigma_s[d, p] \int_0^\pi f_p([d, p], \theta) R_{2nh,p}. \tag{E.11}
\]

Finally, we can write the measurement \(I_{2nh}\) as the sum of \(Q_{2nh}\) and \(L_d\) which, using Equations (E.6) and (E.11), becomes,

\[
I_{2nh} = Q_{2nh} + \sum_{p \in k(n)} \sigma_s[d, p] \int_0^\pi f_p([d, p], \theta) R_{2nh,p}. \tag{E.12}
\]
This is the first part of the desired Equation (E.5).

The derivation of the second part is similar. At pathlength $\tau = (2n + 1) h$, the measurement $I_{(2n+1)h}$ includes contributions from four types of paths: 1) paths that have no propagation events in voxels at depth $d = nh$; 2) paths that only have one propagation event in voxels at depth $d = nh$; 3) paths that have two propagation events in voxels at depth $d = nh$, with a change in direction in both events; and 4) paths of the form of Figure E.1(b).

The contributions for paths of types 1-3 are included in the term $S_{(2n+1)h}$, which therefore is a function only of material parameters \( \{ m [d, p], d < nh \} \) (paths of type 1) and \( \{ \sigma_s [nh, p]; f_p ([nh, p], \theta) \} \) (paths of type 2 and 3).

Paths of \( \bar{x} \) type 4 will have a radiance contribution of the form

\[
\bar{s}_m (\bar{x}) = \left( \prod_{j=1}^{l-1} s_m \left( x_{j-1} \xrightarrow{h} x_j \xrightarrow{h} x_{j+1} \right) \right)
\cdot (1 - h \sigma_t [d, p] + h \sigma_s [d, p] f_p ([d, p], 0))
\cdot \left( \prod_{j=f+1}^{2n-1} s_m \left( x_{j-1} \xrightarrow{h} x_j \xrightarrow{h} x_{j+1} \right) \right)
\triangleq \sigma_t [d, p] T_{(2n+1)h,\bar{x}} + S_{(2n+1)h,\bar{x}}. \quad (E.13)
\]

The terms $S_{(2n+1)h,\bar{x}}$ and $T_{2nh,\bar{x}} (2n + 1) h$ depend only on material parameters \( \{ m [d, p], d < nh; \sigma_s [nh, p]; f_p ([nh, p]) \} \).

The terms $S_{2nh,\bar{x}}$ are included in $S_{(2n+1)h}$, whereas terms $T_{(2n+1)h,\bar{x}}$ for different paths are accumulated into $T_{(2n+1)h,p}$. By summing over multiple voxels $d, p$ for $d = nh$ and $p \in k (n)$, we arrive at the second part of Equation (E.5).
Appendix F

Appendix to Chapter 6

In this appendix, we show proofs of claims used in Chapter 6, we discuss the relationship between the complex and radiometrix Green’s function, and we provide implementation details for the interferometric setup of Figure 6.8.

F.1 Temporal Coherence

We prove Claim 5, reproduced below for reference.

Claim 13. The field correlation magnitude is a Gaussian with standard deviation inversely proportional to the spectral bandwidth $\Delta_k$,

$$\text{corr}(x, \tau) = e^{ik\tau}G^{-1}_{\Delta_k}(\tau), \quad (F.1)$$

where

$$G^{-1}_{\Delta_k}(\tau) = e^{-\frac{1}{2}(\Delta_k\tau)^2}. \quad (F.2)$$

Proof. To evaluate the correlation, we note from Equation (6.3) that the Fourier transform of $u(x,z,t)$ over the time domain is the signal $a_ke^{ikz}$. Similarly the Fourier transform of
\( u(x, z + \tau, t) \) is \( a_k e^{ik\tau} e^{ikz} \). Using Parseval’s theorem, the inner products in the temporal and frequency domains are equivalent,

\[
\text{corr}(x, \tau) = \left\langle \int |a_k|^2 \cdot e^{ik\tau} \, dk \right\rangle. \tag{F.3}
\]

Then, using also Equation (6.4), we have

\[
\text{corr}(x, \tau) = \int e^{-\frac{(k-k_0)^2}{2\sigma_k^2}} \cdot e^{ik\tau} \, dk. \tag{F.4}
\]

Equation (F.4) is (up to a constant) the Fourier transform of a Gaussian with s.t.d \( \Delta_k \). For a Gaussian centered at zero, the Fourier transform is a Gaussian with inverse s.t.d. \( 1/\Delta_k \). As the Gaussian is not centered at zero, this is multiplied by a sinusoid representing the phase shift, resulting in Equation (F.1).

### F.2 Spatial Coherence

We prove Claim 6, reproduced below for reference.

**Claim 14.** For an area light source with an angular range \( \theta \in [-\Theta/2, \Theta/2] \), the correlation decays as

\[
\text{corr}(x, \xi, \tau) \approx W_{\Delta_k}(\xi)e^{i\tau(\xi)} G_{\Delta_k}(\tau), \tag{F.5}
\]

where

\[
W_{\Delta_k}(\xi) = \text{sinc} \left( \frac{\xi}{\Delta_c} \right), \Delta_c = \frac{\bar{\lambda}}{2\Theta}. \tag{F.6}
\]

**Proof.** A plane wave propagating in angle \( \theta \) is

\[
u_{\theta}(x, z, t) = \int a_{k,\theta} \cdot \exp(-ik(\sin(\theta)x + \cos(\theta)z - ct)) \, dk, \tag{F.7}\]

where we assumed \( \eta = 1 \) for simplicity and without loss of generality. For small \( \theta \), \( \sin(\theta) \approx \theta \) and \( \cos(\theta) \approx 1 \), therefore

\[
u_{\theta}(x, z, t) = \int a_{k,\theta} \cdot \exp(-ik(\theta x + z - ct)) \, dk. \tag{F.8}\]
The time delay between the shifted wave $u_{\theta}(x + \xi, z + \tau, t)$ and the original one $u_{\theta}(x, z, t)$ is the projection of the shift $(\xi, \tau)$ on the propagation direction $(\sin(\theta), \cos(\theta)) \approx (\theta, 1)$, 

$$u_{\theta}(x + \xi, z + \tau, t) = u_{\theta}(x, z, t - (\theta \xi + \tau)/c). \quad (F.9)$$

A small adaptation of Claim 5 shows that, for every $\theta$,

$$corr_{\theta}(x, \xi, \tau) = e^{i\theta(x + \xi + \tau)}e^{-\frac{1}{2}(\Delta_k/(\theta \xi + \tau))^2}, \quad (F.10)$$

where the shift between the waves is now $\theta \xi + \tau$ instead of $\tau$ as in Claim 5. If this shift is large, the interference contrast for every $\theta$ decays as a Gaussian with width equal to the temporal coherence length $1/\Delta_k$. However, even if the temporal coherence length is large relative to the spatial shift, so that in Equation (F.10) $e^{-\frac{1}{2}(\Delta_k/(\theta \xi + \tau))^2} \approx 1$, the observable interference still decays if $\xi$ is larger than $\Delta_c$. To see this, we compute the total interference contrast by integrating Equation (F.10) over angles in the source,

$$corr(x, \xi, \tau) = \int_\theta e^{i\theta(x + \xi + \tau)}e^{-\frac{1}{2}(\Delta_k/(\theta \xi + \tau))^2} d\theta. \quad (F.11)$$

Though we can compute Equation (F.11), to obtain a simpler expression we approximate the Gaussian term by its central value,

$$corr(x, \xi, \tau) \approx e^{-\frac{1}{2}(\Delta_k\tau)^2} \int_\theta e^{i\theta(x + \tau)} d\theta$$

$$= W_{\Delta_k}(\xi)e^{i\theta(\tau)}G_{\Delta_k^{-1}}(\tau). \quad (F.12)$$

The last equality follows from the fact that the term $e^{i\theta \xi}$ is a sinusoid varying with $\theta$, whose integral is a sinc. Therefore, for large $\xi$, integrating over $\theta$ can completely eliminate the correlation. \[ \]

**F.3 Light Path Decomposition**

We prove Claim 7, reproduced below for reference.
Claim 15. The correlation of the reference and scattered fields is

\[ C_T(x) \propto \int_{\xi} W_{\Delta,}(\xi) \int_{\xi} \widetilde{T}(f(x) + \xi, x, \xi) e^{i\Delta(\xi - \tau)} G_{\Delta,1}(\xi - \tau) \, d\xi \, d\xi. \tag{F.13} \]

Proof. We denote \( \xi = x + \zeta - f(x) \), implying \( x + \zeta = f(x) + \xi \). Substituting Equations (6.15) and (6.17) in Equation (6.21), we obtain

\[ C_T(x) = \int_{\xi} \int_{\xi} \widetilde{T}(f(x) + \xi, x, \xi) \left( u_{in, \theta}(f(x) + \xi, t - \frac{\xi}{c}) u_{in, \theta}^*(f(x), t - \frac{\zeta}{c}) \right) \, d\theta \, d\xi \, d\xi. \tag{F.14} \]

Using Claim 6, we get Equation (F.13). \( \square \)

F.4 Interference Energy Reduction

When interferometry is applied on a dense scattering media, neighboring particles emit multiple scattered fields. In Section 6.3 we have derived the correlation component of the recorded interference signal and showed that it is a blurred version of the ideal pathlength-resolved light transport function \( \bar{T} \),

\[ C_T(x) \propto e^{-i\Delta(\xi - \tau)} \int_{\xi} \int_{\xi} \bar{T}(x + \xi, x, \xi) e^{i\Delta(\xi - \tau)} e^{-\frac{1}{2}(\Delta(\xi - \tau))^2} \, d\xi \, d\xi. \tag{F.15} \]

When the particle distance is smaller than the wavelength, the interference signal (the blurred \( \bar{T} \) function in Equation (6.22)) averages scalars with different phases. As a result, interference between neighboring points will attenuate the intensity of the measured signal. One may be concerned that this attenuation will destroy all information we hope to capture. To address this concern, in this section we show that the exact same interference phenomenon occurs in the intensity image \( I_s \) that would be captured using a conventional camera, and therefore that the power of the correlation component of the interference signal is proportional to that of the standard intensity image, with no extra attenuation involved.

While interference between particles does occur, the particle locations are sufficiently random so that the interference is not perfectly destructive. The interference between neighboring
paths is the same natural phenomenon happening when measuring a standard intensity image. As most mediums have a sufficiently random structure, the measured intensity image via scattering media contains non-zero energy. We quantify this argument below.

To simplify notation, we assume here coherent illumination and neglect integration over illumination angles. Since in this case the wave is spatially coherent, we shorten notation and write

$$
\mathcal{T}^c(x, \tau) = \int \mathcal{T}(\xi, x, \tau) d\xi.
$$

(F.16)

The intensity of the scattered field. Recall that the measured intensity is $I_s(x) = \langle |u_s(x, t)|^2 \rangle_t$. We can write

$$
\langle u_s(x, t)^*, u_s(x, t) \rangle_t
= \int_t \left( \int_{\tau_1} \mathcal{T}^c(x, \tau_1)^* \int_{k_1} a_{k_1} \cdot e^{-ik_1(ct-\tau_1)} \right) \cdot \left( \int_{\tau_2} \mathcal{T}^c(x, \tau_2) \int_{k_2} a_{k_2} \cdot e^{ik_2(ct-\tau_2)} \right)
= \int_{\tau_1, \tau_2} \mathcal{T}^c(x, \tau_1)^* \mathcal{T}^c(x, \tau_2) \int_{k} |a_k|^2 \cdot e^{i(k_2-k_1)(\tau_2-\tau_1)}
= \int_{\tau_1} \mathcal{T}^c(x, \tau_1)^* \mathcal{T}^c(x, \tau_1 + \epsilon) e^{i(k_1-k_2)\epsilon} G_{\epsilon_2}^A(\epsilon)
$$

where we use $G_{\epsilon_2}(\epsilon)$ as a shorten notation for a gaussian function with variance $\sigma^2$:

$$
G_{\epsilon_2}(\epsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\epsilon^2}{2\sigma^2}}
$$

(F.18)

To express Equation (F.17) let us define $\mathcal{T}^c_{\sigma_2}$ as the signal $\mathcal{T}^c$ times a sinusoid, blurred over a window of width $\sigma$

$$
\mathcal{T}^c_{\sigma_2}(x, \tau) = (\mathcal{T}^c(x, \tau) e^{ik\tau}) * G_{\sigma_2} = \int \mathcal{T}^c(x, \tau + \epsilon) e^{ik(\tau + \epsilon)} G_{\epsilon_2}^A(\epsilon) d\epsilon
$$

(F.19)

We can then express Equation (F.17) as an inner product between $b$ and its blurred version

$$
\langle u_s(x, t)^*, u_s(x, t) \rangle_t = \left[ \mathcal{T}^c(x, \tau)^* e^{-ik\tau} \mathcal{T}^c_{\Delta_2}(x, \tau) \right]_\tau
$$

(F.20)
Equivalently we can think of a blur with a smaller support on both sides

$$\langle u_s(x, t)^*, u_s(x, t) \rangle_t = \left\langle \tilde{T}^{c}_{\Delta^2/2}(x, \tau)^*, \tilde{T}^{c}_{\Delta^2/2}(x, \tau) \right\rangle_{\tau}$$ (F.21)

This holds since

$$G_{\Delta^2/2} * G_{\Delta^2/2} = G_{\Delta^2}$$ (F.22)

(to prove that relationship we can express the convolution in the frequency domain and use the convolution theorem).

In short the intensity of the scattered field equals the magnitude of a blurred version of the pathlength-resolved light transport function $\tilde{T}^c(x, \tau)$. Of course, since epsilon shifts in the path length result in different phases, interference takes place and reduces the intensity of the signal. However, given that when we image scattering materials we measure some intensity, the implication is that we do not have perfectly destructive interference.

**The intensity of the correlation signal.** As derived in Section 6.3 the correlation component of the OCT signal is also a blurred version of the pathlength-resolved light transport function $\tilde{T}^c$, which can be written as:

$$\langle u_{r, \tau_0}(x, t)^*, u_s(x, t) \rangle_t = 2Re\left(e^{-i\kappa\tau_0} \cdot \tilde{T}^{c}_{\Delta^2}(x, \tau_0) \right)$$ (F.23)

That is, the interference signal will correspond to the pathlength-resolved light transport function, blurred with a window of width $\Delta^2$. This blurring reduces contrast since the signal $\tilde{T}^c(x, \tau)$ has changing phase, but it does not wash out the magnitude completely, as it does not wash the intensity of the scattered field itself. As seen in Equation (F.21), the scattered field is blurred by a slightly narrower window of size $\Delta^2/2$, so we can expect that the magnitude of the interference signal will be somewhat weaker. The interference signal will correspond to the pathlength decomposition of a signal with a slightly longer coherence length, but changing the coherence length does not completely wash out the intensity of the scattered wave.
F.5 Acquisition Setup

In this section, we provide details about our implementation of the setup shown in Figure 6.8. For convenience, the schematic is replicated in Figure F.1, along with a photograph of our setup. All references to parts are with respect to the schematic of Figure F.1.

Light source. We use three different light sources (part (a)), depending on the application.
The first source is a high-power white LED, with an output power of 9 W, which we use in combination with color filters. The second is a red (625 nm central wavelength) LED source of output 0.7 W. Even though it has overall less power, all of its output is concentrated within a narrow band of 25 nm, resulting in an overall higher intensity and analogous coherence length compared to the white LED and color filter combination. Additionally, the output of the red LED has a much smaller numerical aperture, making collimation easier and improving interference contrast. On the downside, it is not possible to do RGB scans with this source. We used the red LED to obtain the scattering measurements of Section 6.5, for which contrast is critical. When using either of these two LED sources, immediately in front of the LED we place an adjustable aperture (part (b)), which we use to control the angular extent of the source.

The third source is an SLD of output power 5 mW, central wavelength 650 nm, and bandwidth 20 nm. We found it necessary to use a spatial filter to clean-up the output of the raw diode, which contained undesirable angular components. When using the SLD (with the spatial filter), we remove the adjustable aperture, as the SLD is already practically a pinhole source.

**Illumination lens.** We use a 200 mm prime compound lens (part (c)) to collimate the output of our sources. We also experimented with off-the-shelf, AR-coated achromatic doublets. The compound lens unsurprisingly performed better in terms of reducing chromatic aberration, improving collimation, and additionally slightly improving light efficiency. However, the most critical improvement and the reason why we used the compound lens was that it drastically reduced interreflections compared to the off-the-shelf lens, despite the latter’s coating. We discuss interreflections later. Finally, we only experimented with telephoto-range focal lengths, as beam collimation is critical for interference contrast. Off-the-shelf 100 mm lenses resulted in considerably worse SNR compared to off-the-shelf 200 mm lenses. For this reason, we selected the latter, despite the better light efficiency of lenses with shorter focal lengths.
**Beamsplitter.** Beamsplitters typically come in three types: cube, pellicle, and plate. Pellicle beamsplitters exhibit no chromatic aberration, and are also free of secondary reflections. However, because they are made of a very thin membrane, they are extremely sensitive to vibrations arising from noise or other environment sources. We discuss vibrations later. Cube beamsplitters are rigid, but create many secondary reflections and have large chromatic aberration. We experimented with both types, and in both cases their respective limitations resulted in severely reduced detection SNR. As a compromise between the two, we use a thin plate beamsplitter (part (f)).

**Mirrors.** We use high-quality first surface mirrors of \( \lambda/8 \) flatness (part (h)), though in practice we found no difference in performance when using lower-quality mirrors. For diagonal probing (Figure 6.6(a)), we use a single, circular mirror that can be mounted rigidly, to reduce vibrations. For anti-diagonal probing (Figure 6.6(b)), we use a pair of square mirrors, mounted to form a right angle. For anti-diagonal probing of scenes that are not two-dimensional, we use a monolithic hollow mirror retroreflector (three mirror surfaces forming a cube corner). We use extra square mirrors for the alignment procedure described below.

**Translation stage.** We use a high-end translation stage (l) with minimum incremental motion of 10 nm and low-noise operation. This is necessary in order to minimize surface vibrations induced by the stage’s motion. We discuss this in more detail later.

**Camera lens.** Given the size of our scenes, achieving at least 1 : 1 (real size) magnification is a key specification for the imaging lens. We image using a 180 mm prime compound macrolens (part (k)). The long focal length is necessary in order to have enough working distance, even at large magnifications, for all the optical parts between the camera and the target scene or reference mirror (beamsplitter, polarizers, color filters, and optical mounts). Additionally, the long focal length improves interference contrast.
**Camera.** We use a camera with a high sensitivity CCD sensor (part (l)). The 8 MP sensor has a pixel pitch of 4 μm. As we discuss in Section 6.4, small pixel pitch is an important factor for improving interference contrast. The large sensor size is important for achieving sufficient fields of view, especially at large magnifications. Finally, given the contrast-exposure time trade-off described in Figure 6.4, the high sensitivity of the sensor is critical for keeping capture times reasonable.

**Color filters.** We use hard-coated bandpass color filters (part (e), as their high transmission is helpful for reducing capture times. One downside is that the reflective coating of these filters creates a very strong interreflection that can negatively affect interference contrast. We discuss this later. We use filters of different bandwidths, 3 nm, 10 nm, 25 nm, depending on the desired pathlength resolution (Figure 6.4). To capture measurements at multiple wavelengths, we use filters centered at different wavelengths. The exact center wavelengths change slightly for filters of different bandwidths; for the 25 nm filters we use in most experiments, they are: 625 nm (red), 525 nm (green), and 450 nm (blue). The center wavelengths We use a motorized filter wheel to automatically change filters. This is necessary in order to repeatably perform multiple scans.

**Polarizers.** We use a high-quality wire grid polarizer at the source side (part (d)). Such polarizers have high parallel transmission, which helps reduce capture times. We use simple, laminated film polarizers at the camera and reference arms (parts (g) and (j)), as those polarizers need to be at least as big as the desirable maximum field of view—wire-grid polarizers only come in small sizes. The camera and mirror polarizers are mounted on motorized rotation stages, so that they can be rotated (for instance, to change from parallel to cross-polarization or to do HDR) without user intervention. As with the color filters, it is critical to be able to reconfigure the polarizer positions automatically and repeatably when performing multiple scans.

**Neutral density filters.** We use absorptive neutral density filters. Reflective filters have
strong interreflections, which can severely reduce interference contrast.

**Interreflections.** In order to maximize interference contrast, and therefore SNR, it is important to minimize interreflections between the optics of our setup. These interreflections add a constant intensity component to the captured images, which does not interfere at any position of the reference arm. As a result, a portion of the camera’s dynamic range is used up just by interreflections, suppressing the interference signal and directly reducing the SNR of our measurements. In practice, we found that minimizing interreflections is critical for obtaining useful interferometric measurements in all but the most trivial scenes (mirror or single diffuse planes).

For this reason, we made sure that all of our optical components had multi-layer anti-reflective coating. We used a third-party coating company for components that were not available in coated versions by their providers (such as the neutral density filters).

In practice, we also found it helpful to slightly tilt certain optical components, so that any interreflections caused by them would be directed away from the optical path. This was most important for the color filters, due to the large reflections they created, but it also applied to the camera and mirror polarizers.

**Vibrations.** Another factor that can critically reduce interference contrast and SNR is vibrations of optical components. The motion from such vibrations blurs speckle noise during exposure, resulting in decreased contrast. The decrease is exponential with respect to the vibration amplitude. As a result, even submicron vibrations can effectively remove all interference contrast for scenes that require exposure times of more than a few milliseconds.

For this reason, we built our setup on an optical table with a passively damped table top and a pneumatic isolation system, to reduce surface and environment vibrations respectively. Moreover, we placed an enclosure system around the setup during acquisition, to reduce vibrations from environment sources such as acoustic noise or air flow.
Within the setup, a major source of mechanical vibrations is the translation stage at the reference arm. We experimented with DC stepper motor and piezo-electric stages, as well as a high-end stage designed for nanometer resolutions. The very low-noise operation of the latter significantly improved performance compared to the other two categories, though the piezo-electric stages also gave good measurement contrast.

Finally, rigid and stable mechanical mounts for all optical components can significantly improve interference contrast. We used a laser Doppler vibrometer to identify parts that were vibrating during the scanning process, and improve mounting as necessary.

**Motion control.** The optical setup requires careful alignment (described later) of the various optical components at accuracies of a few microns. To achieve this, we mount most components on manual micro-positioning stages, with various degrees of freedom as necessary. The reference mirror, camera, beamsplitter and source LED have mounts providing five degrees of freedom (XYZ translation, tilt, and rotation), whereas the illumination lens has three degrees of freedom (XY translation and rotation).

**Alignment.** The optical setup requires aligning the optical axis of the illumination, reference mirror, and camera (through the beamsplitter). We perform this alignment as described below.

First, we mount a pinhole in front of the illumination lens, to reduce lighting to a narrow beam. In practice, we have an adjustable aperture permanently mounted at the front of the lens, and close it for alignment.

Second, we mount a mirror between the beamsplitter and the illumination lens, at an angle of 45° relative to the beamsplitter (there exist carefully machined beamsplitter mounts that enable this). We rotate and tilt the beamsplitter mount, until the beam is reflected by the mirror back at the center of the pinhole. This ensures that the beamsplitter is aligned at 45° relative to the illumination direction.
Third, we remove the mirror, leaving the pinhole in place. The beam now travels through the beamsplitter towards the reference mirror, where it is reflected. To align the reference mirror with the optical axis, we rotate and tilt the reference mirror until the reflected beam returns at the center of the pinhole.

Fourth, we remove the pinhole from the illumination lens, and focus the camera lens at infinity. The infinity focus means that the illumination axis will be parallel with the camera’s optical axis when all of the collimated light is imaged at the center of the camera sensor. We can make this happen by rotating and tilting the camera while looking at a live stream of the sensor. Once the two axes are parallel, we refocus the camera lens and translate the camera without changing its direction, in order to spatially center the collimated beam relative to the sensor.

Finally, in many cases it is useful to place another aligned mirror at the target arm, as in the basic Michelson interferometer setup of Figure 6.2. This can be used to measure resolution, through the experiment described in Figure 6.4, or to fine-tune various imaging parameters for maximum interference contrast. The alignment can be done by first focusing the camera lens at infinity, then tilting and rotating the target mirror until its reflection is imaged at the center of the camera sensor.

**Component list.** In Table F.1, we provide a list of the major components used in our setup. We omit standard optical mounts and micropositioning stages, which can be found in most optical equipment providers.
Table F.1: List of major components used in the optical setup of Figure F.1.
References


[34] Duchi, J., Shalev-Shwartz, S., Singer, Y., and Chandra, T. Efficient projections onto the $l_1$-ball for learning in high dimensions. *ICML* (2008). 78


234


[61] Henyey, L., and Greenstein, J. Diffuse radiation in the galaxy. The Astrophysical Journal 93 (1941), 70–83. 6, 14, 111


[107] NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. All-frequency Shadows Using Non-linear Wavelet Lighting Approximation. ACM Trans. Graph. 22, 3 (July 2003), 376–381. 58


238


[133] Sloan, P., Kautz, J., and Snyder, J. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. ACM Trans. Graph. 21, 3 (2002). 194


