NORTHWESTERN UNIVERSITY

Methods for the Imaging, Analysis, and Display of Layered Media

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Applied Mathematics

By

Lionel D. Fiske

EVANSTON, ILLINOIS

March 2023

ABSTRACT

Methods for the Imaging, Analysis, and Display of Layered Media

Lionel D. Fiske

This dissertation is a review of three projects I worked on during my time in the Computational Photography Lab at Northwestern University. First, a source separation problem for the X-Ray Fluorescence images of painted works of art is addressed through the incorporation of Hyperspectral Reflectance data. Following this, a discussion of Optical Coherence Tomography and its applications to cultural heritage science is presented. A rigorous analysis of the depth resolved attenuation coefficient in the presence of speckle is performed and a Bayesian model for the signal is derived. Finally, the problem of speckle in fast temporally multiplexed holographic displays is addressed. In this, the impact of quantization on the reconstructed image quality is analyzed and quantization aware optimization methods to reduce speckle are surveyed.

Acknowledgements

First and foremost, I would like to acknowledge and thank my PhD advisor Oliver Cossairt. Over the past four years Ollie has given me the freedom and encouragement to learn and explore any topic I wanted. This freedom lead to a number of interesting projects, collaborations and existential crises I never could have had in a more restrictive lab. I am grateful for for his guidance, mentorship, and family barbecues. I would also like the thank Maurice Aalders who was my advisor during my time in the Netherlands. Maurice encouraged me, took me in, and gave me a lot of hands on guidance both in my time in the Netherlands as well as the projects we collaborated on in the states. Thank you for all your support.

I want to give a big thanks to all of my collaborators. Marc Walton, who was the head of the Center for Scientific Studies in the Arts gave me many opportunities for which I am grateful, however in particular your help on the XRF project is greatly appreciated. Thank you to Matthias Alfeld who taught me about XRF and was happy to answer any cultural heritage science question I brought to him. I would also like to acknowledge to Dirk Faber who provided a lot of support and interest to my OCT project. Your expertise made this project happen. Thank you to Gang Li, who encouraged me, taught me experimental optics and helped me during my internship at Meta. Thank you to Aggelos Katsaggelos who served as a pseudo-advisor on many of my projects. Additionally, thank you to

Mitra Almasian for including me in her many interesting research problems. I hope we can continue to collaborate in the future.

I want to acknowledge the entire Computational Photography Lab, but in particular Florian Schiffers who provided me with support, friendship, and many interesting conversations over the years. I would like to thank all of the talented staff and students at the center for the scientific studies in the arts, particularly Olivia Dill who has been not only a close friend but also edited the first chapter of my thesis. Additionally, thank you to Leah Wilk for being a friend and collaborator during my time abroad. I would like to thank my all of my friends, but in particular Angelia Wang, Josh Roan, and Maria Canyon for their support and friendship. A big thank you to Kevin Gallagher for teaching me about flies, non-duality, and tolerating my long winded rants about Fourier analysis. I would also like to thank my parents for supporting me through this process.

My PhD was almost entirely funded through the NSF PIRE grant 1743748. This funding allowed me to live and work abroad, take on unlikely and interesting cultural heritage projects, and gave me a community of students with whom to do research. Thank you to all the members of the PIRE community but particularly Ken Shull and Matthew Hsu. Finally, I would like thank my committee for reading this thesis closely, but hopefully not so closely that they catch all the grammar mistakes.

Table of Contents

List of Tables

List of Figures

2.7 This figure shows realization of the posterior distribution for the attenuation coefficient given in equation 2.30 for different values of $\hat{\mu}$. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.

a) This panel shows two unimodal reconstructed posterior distributions. With these distributions, the true parameter is much more likely than the DR estimate. This posterior was constructed with a layer mean of $\langle \mu_{oct} \rangle = .4$ mm^{-1}, $\zeta = 6.87 * 10^{-2}$ mm^{-1} and a DR estimates of $\hat{\mu} = .08$ mm^{-1} and $\hat{\mu} = 1.3$ mm^{-1} .

b) This panel shows a constructed posterior distribution which is Bi-Modal and has two local maxima. For a given layer mean, the constructed distribution develops a second peak if the DR estimate used to construct the posterior is sufficiently small. This second peak can make the Maximum a Posteriori difficult due to non-convexity. In many cases, the maximum value of the Posterior distribution may sit very near the origin on this second peak. As demonstrated in this panel, often the total amount of probability mass under the peak near zero is relatively small, despite being the maximum likelihood. Thus, the Maximum of the posterior distribution is a poor representation for the distribution itself. In these cases an estimate for the mean is a better choice. This posterior was constructed with a layer mean of $\langle \mu_{oct} \rangle = .4$ mm^{-1}, $\zeta = 6.87 * 10^{-2}$ mm^{-1}, and a DR estimate of $\hat{\mu} = .015$ mm^{-1} . 75

b) Plot of the DR estimate for A-scan extracted from a). This estimate shows how highly variable the DR attenuation coefficient appears to be in the presence of speckle.

c) Mean of the Posterior estimate for the extracted A-scan. This was computed using the $\langle \mu_{oct} \rangle = 3.38$ mm^{-1} which is the layer mean for the first 100 rows of pixels of the B-scan., The value of ζ was computed to be $\zeta = 6.0053 * 10^{-2} mm^{-1}$ from a voxel volume of $3.2 * 10^{-6}$ micron and a scattering cross section of $\sigma_{scat} = 1.9 * 10^{-9}$ mm^2.

d) Comparative DR estimate for the average of all 1000 A-lines in the B-scan. The resulting fluctuations are still very large even after averaging 1000 A-scans 78

2.10 This figure demonstrates the accuracy of the likelihood model and estimates the coefficient of variation of the signal to verify the assumptions made in 2.20 for simulated signals. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.

a) This panel is a simulated B-Scan, which was simulated using parameters $\beta_{NA} = .3$, $I_{inc} = 1e7$, $\mu_{oct} = 2.00$ mm^{-1}, lateral resolution of $\Delta x = .022mm$ and $\Delta z = .0068$ mm in a domain which is 3.4 mm deep.

b) This is the DR reconstructed coefficient map. The reconstruction is highly variable around the true attenuation value of 2.00 mm^{-1}. This panel also demonstrates the growth artifact in the bottom 30%

of voxels where estimated values which are much higher than the true value. The estimates below the red line are truncated to avoid the exponential grown artifact.

c) This figure is a histogram of the top 70% of pixels from panel b). As shown in equation 2.23 we expect this to be exponentially distributed with parameter 2.00 . A best fit exponential demonstrates this is accurate to 3 significant figures.

d) Coefficient of Variation for the simulated OCT A-scans in panel a) at different depths. In these simulations, C_v stays near .12 at all admissible depths. 79

2.11 This figure shows estimates of the attenuation coefficient for simulated OCT data using the standard DR and the Bayesian estimator given in 2.32. The OCT data was simulated with parameters $I_{inc} = 1e7$, $\mu_{oct} = 2.00$ $\beta_{NA} = .3$, $\sigma_{scat} = 1.0 - 6$ mm^2, a lateral resolution of $\Delta x = .02$ mm and $\Delta z = 0.0068$ mm in a domain which is 13.6 mm deep. After the attenuation coefficient was inferred using the DR method the bottom 30% of pixels are discarded to avoid reconstruction artifacts. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.

a) This panel shows the ground truth attenuation coefficient for the simulation. This ground truth is a realization of the prior distribution given in equation 2.28.

b) This image shows the reconstructed attenuation coefficient using

Preface

For the majority of my time at Northwestern I have worked in the Computational Photography Lab (CPL) under the guidance of Prof. Oliver Cossairt. This lab focuses on the blending of physical modeling, machine learning, optics, and imaging know-how into new ways to take, display and analyze images. In learning to do research in this lab I became wholly interdisciplinary, learning to speak the languages of art conservators, engineers and physicists in addition to applied mathematics. This dissertation is a review of three of the projects I worked on at during my time in CPL, and represents the culmination of many hours of hard work. While this is an exploration of several topics, the projects are spiritually related through the modeling and understanding of impact of material properties on optical fields.

The first project discussed in this thesis is on the subject of ill-posed source separation of X-ray fluorescence (XRF) elemental images of paintings with multiple layers. I worked on this while jointly appointed in the Northwestern Center for Scientific Studies in the Arts (NU-ACCESS) in collaboration with two talented researchers Prof. dr. Matthias Alfeld and Prof. dr. Maurice Aalders from TU Delft and the Amsterdam UMC respectively. While I learned much in graduate school, the origin of acronym "NU-ACCESS" remains a mystery to me. We developed a technique which identified correlations with another less penetrating spectroscopy method. This differential in penetration depths allows us to gain some 3D information from 2 separate 2D measurements. These results were published these results at the peer reviewed conference ICIP in 2021.

The following chapter, covers work I did while living abroad as a researcher at the Amsterdam UMC in the Netherlands under the guidance of Prof. dr. Maurice Aalders and in close collaboration with Prof dr. Dirk Faber. This project originated as a question

of the applicability of attenuation analysis methods for OCT to the identification of paints in paintings. Instead of that problem, we ultimately addressed an open question modeling the uncertainty of the attenuation analysis methods as a function of the material properties. In this chapter a rigorous statistical model for this system is proposed and and a careful treatment of the signals which considers noise is done. These results were published in Scientific Reports in 2021.

The second chapter additionally covers two papers which I contributed to in collaboration with Dr. Mitra Almasian and several other talented researchers at the Rijks Museum in the Netherlands. This collaboration came about while working in neighboring offices at the Amsterdam UMC and realizing that much of the code I was developing to analyze OCT speckle could be repurposed for the analysis of OCT images of paintings. I helped to develop a robust image processing pipeline with Mitra for the purpose of analyzing varnish and glaze layers layers in painted works of art. While, I continue to provide technical support and help for Mitra in these investigations the spirit of the scientific work comes from her and the other talented scientists at the Rijks Museum. This resulted in a publication in a Cultural Heritage Science in 2021 and a second publication in review as a part of operation Night Watch.

In the final chapter of the thesis, the problem of displaying a 3D object is considered using a technique called Computer Generated Holography. My interest in holography developed as a result of listening to Ollie talk endlessly about Fourier optics and 3D displays. This holographic indoctrination culminated in a research internship at Meta Reality Labs under the supervision of Dr. Gang Li to work on holographic display technology for next generation head mounted displays. Mathematically and conceptually, this problem

is deeply related to the OCT problem. Only, instead of measuring the object as a series of slices instead you must display them. To do this, coherent light is controlled using a spatial light modulator forming a potentially 3D image in the interference pattern. This project studied the speckle introduced into these interference patters from fast spatial light modulators and surveyed different methods which could help to reduce it.

This dissertation is about the numerical modeling of the interaction between optical fields and objects. In this work, I will study two versions of this problem. The source separation problem covered in chapter 1 and the attenuation parameter problem posed in chapter 2 are focused on analyzing and extracting images taken with various modalities. That is, measuring an optical field and extracting information about materials from these measurements. In chapter 3 I focus on the opposite problem; the realities of generating an arbitrary optical field through the use of real world optical components. In the following chapters I will demonstrate an understanding of a wide variety of topics, techniques, and concepts and use those to answer or investigate problems in computational imaging and display. I will call on my background in numerics, modeling, as well as signal processing and machine learning to address these problems in fun and creative ways. I sincerely hope this dissertation is as enjoyable a read as graduate school was to complete.

CHAPTER 1

Computational XRF Delayering: Inferring Depth with Data Fusion

1.1. Introduction

Cultural heritage science is an interdisciplinary field which blends techniques from history, art conservation and hard sciences together in an effort to answer historical questions and to preserve art and artifacts for future generations. While scientific techniques have been applied to art historical questions since at least the 1930s [**92**], there has recently been increased interest in bringing new computational imaging techniques to bear on these problems [**98, 94**]. In the study of painted works of art, a problem of particular interest is material identification and characterization. Solving this problem is necessary for several applications. First and foremost, material identity can be used to inform the preservation of the artwork as different materials will require different light, humidity, and temperature conditions for storage and display and specific conservation treatments. Material identity is also useful for studies of the artist's style and access to different pigments and media. Furthermore, an understanding of both the pigments an artist worked with and the order in which they were applied in layers on a canvas or support can be important for art historical questions. This *stratigraphic* information when coupled with spatial information can be used to identify revisions, preparatory layers, and an artist's working process. Traditional approaches to layered material identification involve the taking of physical paint cross-sections which is damaging to the artwork. Furthermore, these techniques provide only local information and do not give access to information across the entire work of art.

To avoid both of these issues, imaging spectroscopy is now commonly employed. Imaging spectroscopy refers to any technique in which a large 2D 'image' of an object is created where each pixel contains some sort of characteristic spectra (E.g. reflectance,

fluorescence, phosphorescence) of the object. These spectra contain chemical or physical information. In the case of painted works of art or other layered media, imaging spectroscopy often involves the use of deeply penetrating wavelengths such as X-rays or infra-red which can probe information about all of the layers– allowing access much of the same information as a cross section in a non-contact and non-destructive manner. For all of their advantages, these techniques provide large amounts of complex data and as such the interpretation of the acquired images features is still an ongoing area of research.

This chapter addresses the problem of identifying surface and subsurface portions of 2D X-ray fluorescence elemental maps through the incorporation of visible reflectance data. In essence, this is an attempt to infer depth information from two different measurements which are not depth resolved. The XRF measurements are a superposition of the component parts we seek to recover, making this problem essentially like trying to recover two numbers which have been added together. A complementary spectroscopy technique with a different penetration depth profile will contain additional information about the stratiography, which can help to address the ill-posedness of the delayering problem. This chapter shows a proof of principle that penetration depth diversity can be used to gain 3D layering information and the limitations of this simple approach are also discussed.

1.1.1. Chapter Outline and Structure

To better understand the problem, section 1.2.1 provides the background needed to understand this problem. Specifically, section 1.2.1 and 1.2.2 some background on XRF and RIS are given respectively. Section 1.2.4 other data fusion approaches to similar problems

and section 1.2.3 provides context to how these problems arise in the context of cultural heritage science. Following this, a fusion algorithm is then proposed in section 1.3 and results are demonstrated in section 1.5

1.2. Background

1.2.1. X-ray Fluorescence Spectroscopy

X-ray Fluorescence spectroscopy (XRF) is a non-contact and non-destructive material identification technique which gives information on the elemental make up of an object [**102, 42, 6, 33, 4, 7, 32**]. It works by illuminating an object with high energy X-rays with energies ranging from 20–60 kV. This X-ray bombardment causes electron transitions in the inner shells of the constituent atoms in an object, leading to the release of fluorescent X-ray radiation at very characteristic energies. These emission spectra are measured and mapped to different atomic species using a fitting routine [**102**]. By measuring over a large grid of points, a 2D image called an *XRF elemental map* can be created allowing for the visualization of the distribution of various elements across an object, as shown in figure 1.2. XRF has been used in cultural heritage applications since the 1950s [**42**], however in recent years, developments in sensor technology have made XRF much more portable and accessible for cultural heritage scientists and conservators. While this technique is powerful for non-destructive material analysis, X-ray dose can be an issue for cultural heritage applications [**14**]. Additionally, this technique gives information about the distribution of atomic species and not about the chemical state or bonds these atoms are in. This can make distinguishing between different materials with similar atomic make ups difficult. Due to the fact that both the incident illumination and fluorescent light in

this technique have an exceedingly high penetration depth (X-rays) this technique can probe the interior or subsurface of an object. This can be used to study overpaintings where additional paint was applied to change a region in a work of art [**6, 5, 32**]. In layered media, signals from all layers are detected as a superposition. This often leads to ambiguities as to where in the bulk of an object a signal is coming from. Trying to find the signals from different layers given their super-position is an ill-posed source separation problem. Blind source separation techniques such as ICA may be of limited use because of the assumption that the layers are statistically independent which is not true in general for painted images. Confocal XRF can be employed to depth-resolve features in a sublayer [**65**]. However, it is a technically complex method, costly, and requires both the much higher total X-ray dose and longer acquisition times typically associated with synchrotron sources. Confocal XRF is thus not feasible for large paintings or routine applications in cultural heritage science.

1.2.2. Visible Reflectance Imaging Spectroscopy

Visible reflectance imaging spectroscopy (RIS) is a non-destructive and non-invasive technique in which one illuminates an object with broadband white light and measures the reflected light wavelength by wavelength. At each pixel, the *reflectance curve* is continuous in wavelength and has a characteristic shape dependent on the material composition of the underlying object, as demonstrated in figure 1.4. In particular, many commonly used pigments have distinctive absorption features in the visible range which allow them to be unambiguously identified. Since the development of compact RIS cameras in the 1990s

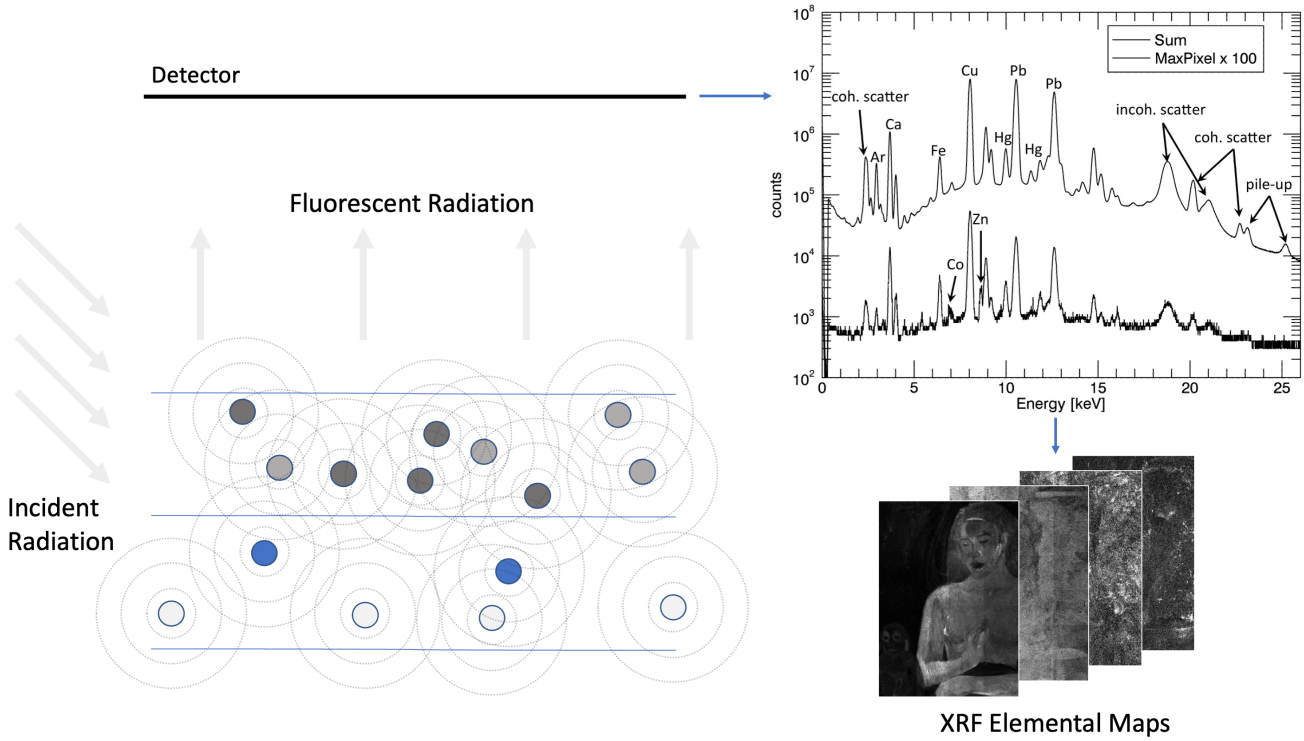


Figure 1.1. Layered structures are irradiated by x-rays and the resulting fluorescence is measured in an energy-resolved way. Elements have characteristic peaks at certain energies which are fitted to qualitatively measure local concentration. These concentrations are then arranged into maps of elemental concentration. Example spectra reproduced from [**121**].

[**88**] these hyperspectral imaging methods have become a workhorse technique for identifying pigments and repairs [**33, 4, 7, 81, 94**]. Typically, measured reflectance curves are compared to a reference database either visually or with quantitative similarity measures such as spectral angle, allowing for the identification of materials such as pigments and additives.

The RIS signals of different pigments mix in a nonlinear fashion based on the absorption and scattering properties of the pigments, fillers, and binding media [**69, 53, 108**].

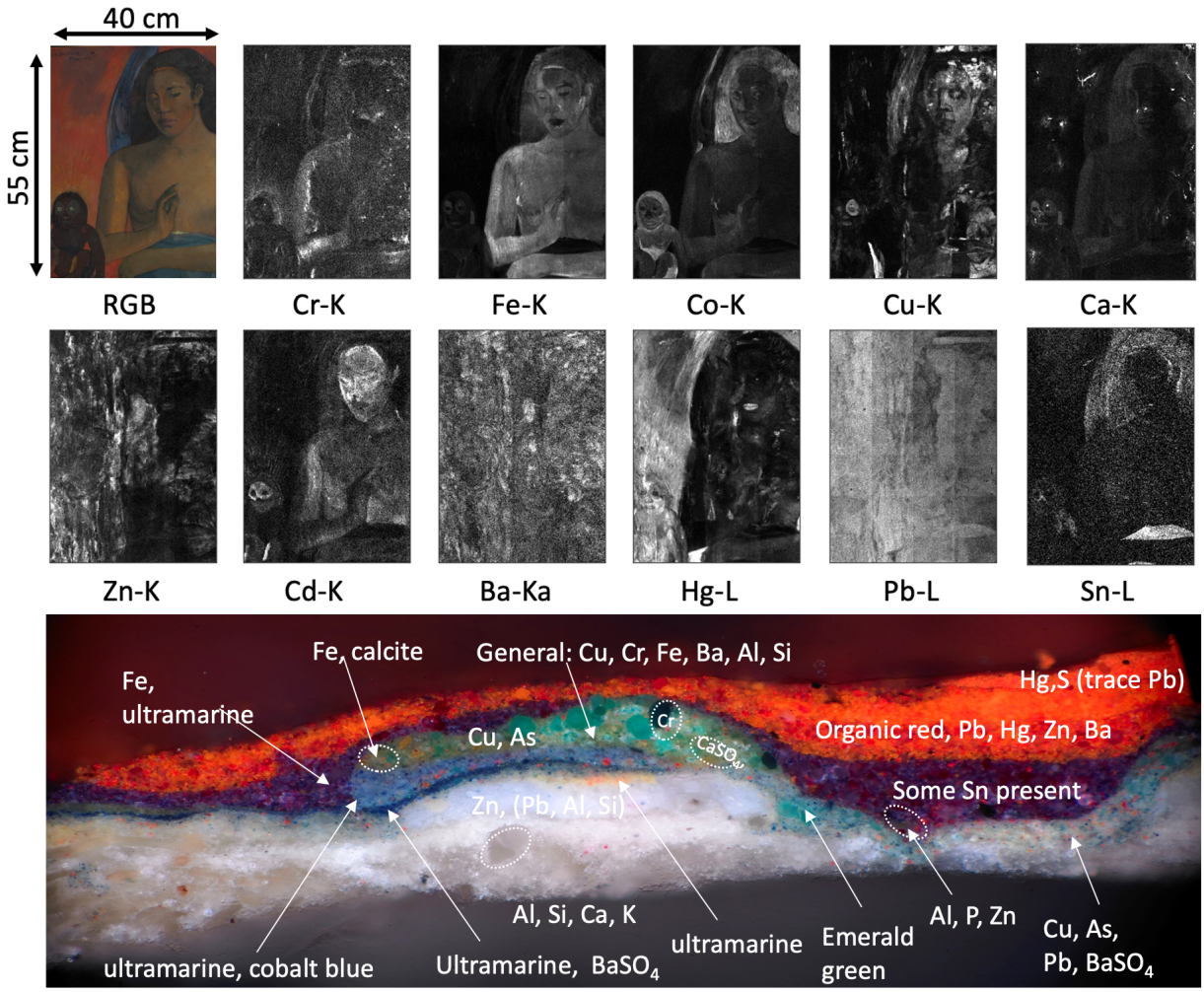


Figure 1.2. Example XRF elemental maps from Poèmes Barbares [**2**] by Paul Gauguin taken as a part of a collaboration between the Straus Center for Conservation and Technical Studies at the Harvard Art Museums and the Northwestern Center for Scientific studies in the arts [**115**]. The top rows show images whose contrast represents the relative abundance of various atomic species. The bottom panel is a real cross section taken from the painting showing the complex composition and layering of the same painting.

However, due to the fact that both the absorption and scattering coefficients are wavelength dependent, the reflectance curve simply does not contain enough information to

recover them unambiguously without additional assumptions or measurements [**99**]. This means that, at least theoretically, two paints can have the same reflectance curves but behave differently when mixed with other paints. Sometimes this ambiguity is addressed by using a dictionary or database of known pigments for regression [**95**].

RIS is often used in conjunction with clustering and nonlinear embedding techniques for unsupervised endmember (pigment) identifications [**7, 115**]. These approaches are essentially trying to find a inverse mapping between pigment identity and relative concentration and the reflectance signal itself. These techniques can find potential pigment candidates which mix either linearly or nonlinearly to produce the measured signal. However, it is not always the case that these pigment candidates are physical and these unmixing problems are still an open area of research. Due to the fact that paint absorbs and reflects light strongly in the visible range the penetration depth of light in this range is generally very small. Thus, in many situations RIS can only provide information about the surface of an object and does not probe the layered structure at all.

1.2.3. XRF Source Separation Problems Arising in Layered Media

In the measurement of XRF elemental maps often the specific layer a pigment signature is coming from can be inferred by reasoning about the techniques in constructing a painting. For example, in the lead elemental map (Pb-L) shown in figure 1.2 the signal is relatively flat and uniform across the surface of the entire painting. Clearly, no single lead containing pigment is used across the entire surface of the painting so this signal must be coming from the subsurface layers. In the case of this painting, the canvas was painted with a lead white ground layer and the image was built on top. However, other signals are more difficult

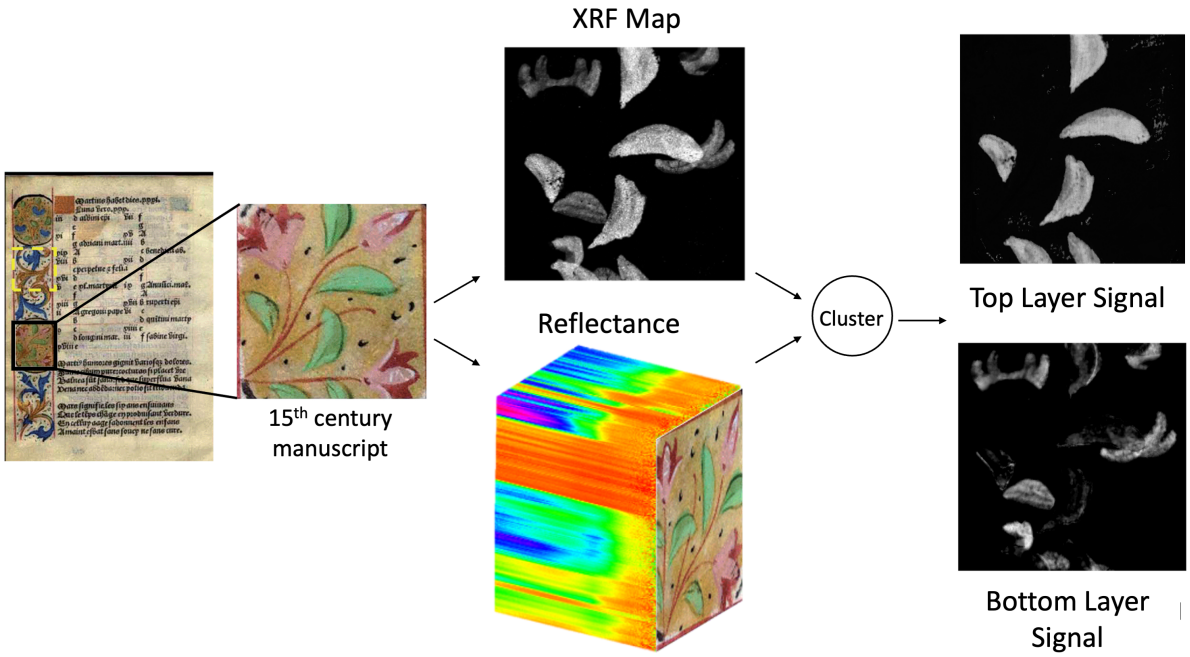


Figure 1.3. The Proposed framework takes co-registered XRF and RIS data cubes. RIS data is clustered into K pigment mixtures and the surface XRF is estimated by calculating the mean XRF response across all clusters. The subsurface XRF signal is estimated by subtracting the surface XRF from the total XRF signal.

to interpret when lower layers and surface layers have elemental composition. Clearly in this case the subsurface layer is visible but the pigments on the surface somewhat obscure and hide a clear image of the elemental map below. It would be desirable from a data visualization and clarity perspective if it were possible to identify and remove the contribution of the surface pigment layer to the elemental map and only view the subsurface.

1.2.4. Related Work

Some approaches to infer depth information has been from XRF and related measurements have been investigated. For example, it has been shown that layering information in XRF can been gained through the incorporation of PXIE, a spectroscopic technique in which a sample is exposed to an ion beam which then induces X-ray Fluoresence which can be measured. One paper found that some layering information can be gained by taking PIXIE measurements with different incident energies [**83**]. These results are impressive but require a particle accelerator making this technique out of reach for most cultural heritage problems. In some cases, depth information can be extrapolated from a series of XRF measurements which are made with increasingly tilted detection angles [**109**]. Signals from deep layers will attenuate more due to interactions with the surface layers allowing depth information to be gained by the ratio of peak heights. However, this can lead to ambiguities when multiple layers have a given element. Additionally, very precise alignment of the illumination between measurements is required. This can be accomplished by using two detectors simultaneously effectively doubling the hardware requirements to use this technique.

Painted works of art are often investigated by using both RIS and XRF. Combined RIS and XRF have been used to study the connection between stylistic changes and materiality in Late Rembrandt's [**34**], study the composition of many illuminated manuscripts [**81, 65**], and to identify hidden or obscured text [**89**]. Currently, both techniques are being employed for the highly publicized *Operation Nightwatch* [**94**] to study and analyze materials used in the Nightwatch by Rembrandt. These modalities are also employed in other disciplines such as Geo-science, where recent work used them to study sediment

cores [**93**]. While both RIS and XRF are often independently analyzed using quantitative techniques, the fusion of the data is typically qualitative and, with a few exceptions [**7**], performed by simply comparing the resulting images by eye. While some preliminary work demonstrated that quantitative fusion is possible [**29**] techniques to solve this problem remain an open area of research.

Some work has been done on the problem of computationally delayering the X-radiography images [**30**]. In this work, a joint dictionary approach based on sparsity is employed to accomplish the separation. While the results are impressive, the problem addressed in this chapter fundamentally differs, as all measurements are made from the surface layer of a painting. A similar one sided problem has been investigated as part of guided XRF super-resolution algorithm [**28**]. In this, an RGB image is used to help interpolate XRF data. This paper split the XRF signal into a portion which is correlated with the RGB and a portion which is unrelated. This splitting suffered from the fact that RGB images typically do not contain enough information for robust pigment identification because many pigment combinations are visually identical. Furthermore, a dictionary approach implicitly assumes a linear mixing model for pigments, which is nonphysical and can lead to poor results. Some newer work has used a Neural Network to find the nonlinear mapping [**91**] however, this approach requires information from both sides.

1.3. Delayering XRF with RIS Data

Light penetrates into a material structure with a depth dependant on wavelength. For the wavelengths used in RIS to analyze most paint systems, we see very little reflectance

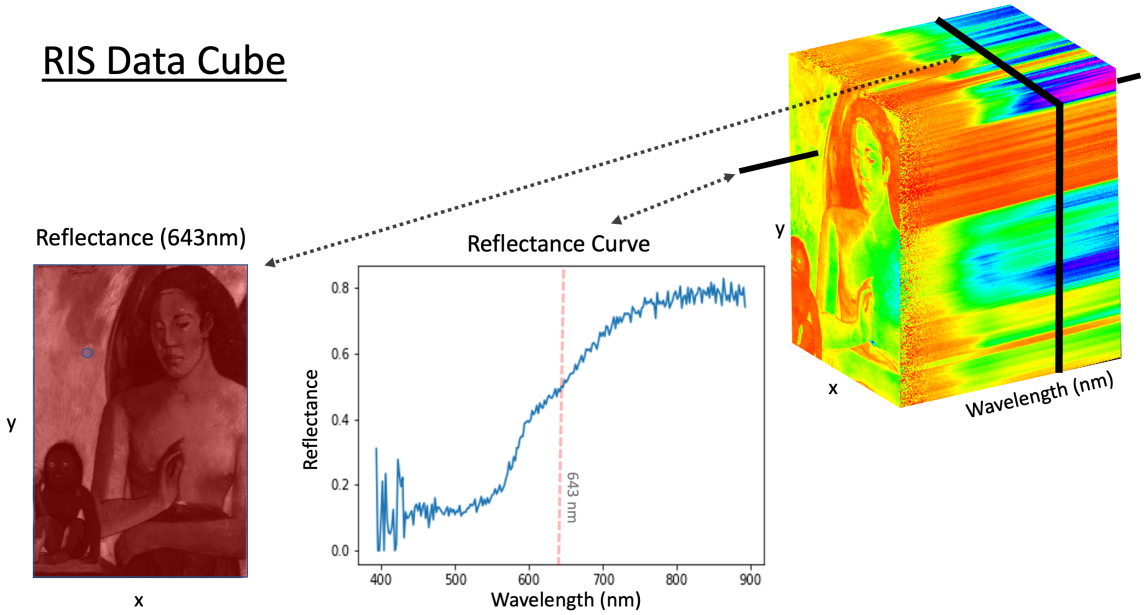


Figure 1.4. An example RIS data cute for the Gauguin dataset. This figure demonstrates the two ways RIS data can be interpreted, as a series of reflectance curves containing data in their shape, or as a series of images each with different relative contrast and brightness between the pixels. The image on the left represents the x,y slice of the data cube at wavelength 643. The blue dot on the red image represents the position of the reflectance curve next to it.

from deep buried layers. Conversely, X-rays, associated with XRF, penetrate through the entire layered structure of a painting with ease. From this perspective, the question of whether it is possible to exploit the diversity in penetration depths alone to isolate the portion of the XRF signal which is closely associated with the surface of the painting.

To gain information about the XRF from the RIS, the data sets must be mapped into a space where they are comparable. Conceptually, a painting is made from physical materials, *i.e.*, different paints and glazes, with a given spatial distribution and chemical structure. Since the chemical structure is made from a defined set of atoms we expect XRF elemental map intensity and RIS data to both be functions of the physical pigment

distribution. Thus, if we can successfully cluster a hyperspectral signal, such that the clusters represent the pigment concentration, we should also expect these clusters to be correlated in some way with XRF signal intensity.

Due to the non-linearity of each underlying mixing problem it is reasonable to expect the relationship between the XRF map and the RIS data cube to be nonlinear and difficult to explicitly model. One obvious mapping would be the composition of the RIS to pigment inverse function which is inferred through clustering and the pigment to XRF function which is either inferred or measured explicitly. Obviously, XRF elemental inverse mapping will not be one-to-one without adding in additional constrains due to the fact that many pigments contain the same elements. Finding RIS to pigment mapping is commonly done via a clustering approach- either a non-negative matrix factorization or a non-linear technique such as TSNE. The goal is to have as many endmembers as pigments used on the surface of the painting, usually around 5-10. However, this 'unmixing' problem is very sensitive to clustering method and parameter choice leading to questions about how physical the resulting distributions and endmembers are in general.

For the delayering problem, the specific distribution of a given pigment is intermediary so solving the difficult and hard to verify nonlinear unmixing problem may be unnecessary. Instead, the assumption can be made that the reflectance signal $R(\boldsymbol{p_i})$ is continuous, such that given two mixtures of pigments $\boldsymbol{p_1}, \boldsymbol{p_2}$, if $||R(\boldsymbol{p_1}) - R(\boldsymbol{p_2})||$ is small then $||\boldsymbol{p_2} - \boldsymbol{p_1}||$ is small as well. This assumption can be used to justify clustering the reflectance data into K groups which are spectrally most similar, and that these spectrally similar groups should also have surface pigment concentrations. Instead of unmixing the pixels, the pixels are separated into groups with the most similar mixtures, without uncovering what

those mixtures exactly are. Pixels with similar surface pigment concentrations should have similar surface contributions to the XRF signal allowing us a route to isolating the proportion of the signal coming from the surface layer of the painting.

Once the initial clustering step has been performed the XRF response for each constituent cluster can be estimated. Given sufficiently large clusters, statistically independent layers, and assuming that the subsurface image is sparse, the mean XRF response should provide a sensible estimate for the surface XRF signal for those pixels. We can then form an image of a predicted surface distribution by replacing each cluster of pixels in an image with the mean XRF response as seen in figure 1.7. The surface concentration can be estimated as

$$\hat{X}_{surface} = \sum_{k} \mathbb{E}_{C_k} [X_{tot}] \mathbb{I}_{C_k} \tag{1.1}$$

Where k is the cluster index and C_k is the k_{th} cluster, and $\mathbb{I}_{C_k}$ is the indicator function on cluster C_k. High frequency error can be introduced both by slight misregistrations as well as from hard cluster boundaries. To account for this we consider the solutions to multiple K-means clusters which have their XRF intensities slightly varied by shifting the XRF signal by single pixels relative to the reflectance cluster and different initial seeds. We estimate the surface pixel value as

$$X^{i}_{surface} = \min_{l} \left\{ \hat{X}^{l,i}_{surface} \right\} \tag{1.2}$$

where $\hat{X}^{l,i}_{surface}$ is the surface estimate for the i_{th} pixel given by the l^{th} clustering seed. We have found that this routine produces results with fewer errors at the boundaries than using a single estimate. The example in figure 1.6 uses 13 different clustering seeds and

shifts. Finally, the subsurface intensity is estimated as

$$X^i_{subsurface} = \min(X^i_{tot} - X^i_{surface}, 0). \tag{1.3}$$

The threshold of 0 is physically motivated from the fact that XRF elemental intensity is strictly positive. This algorithm is written in psuedocode in appendix C.1.

1.4. Data Sets

To test this delayering approach, a page from a 15th/16th century Book of Tides with decoration and writing on both sides of the page is considered. This manuscript has been analyzed and imaged by XRF and RIS as shown in figure 1.3. Since by its nature, a manuscript allows easy access to the rear of the page, features on the backside can easily be compared to the results obtained through visual inspection. While having access to the rear helps with verification, we will only use measurements made from a single side. Confocal XRF data was also obtained for this manuscript [**65**] thus providing a ground truth comparison with a different technique.

To assess the performance of the algorithm, a simulated dataset produced from real reflectance curves for pigments clustered out of the Book of Tides dataset has been produced. The reflectance curves from the RIS data are split into 15 clusters and select an example cluster for red, beige, white, and green regions in figure 1.6 showing a figure which is a simulated painting depicting the ICIP 2021 logo. The respective clusters contain several hundred pixels, each with their own reflectance spectra, which we assign pixel-wise to the image. Each color in the image is then assigned a mean XRF response and a top layer XRF signal is created as a realization of a Poisson distributed around the

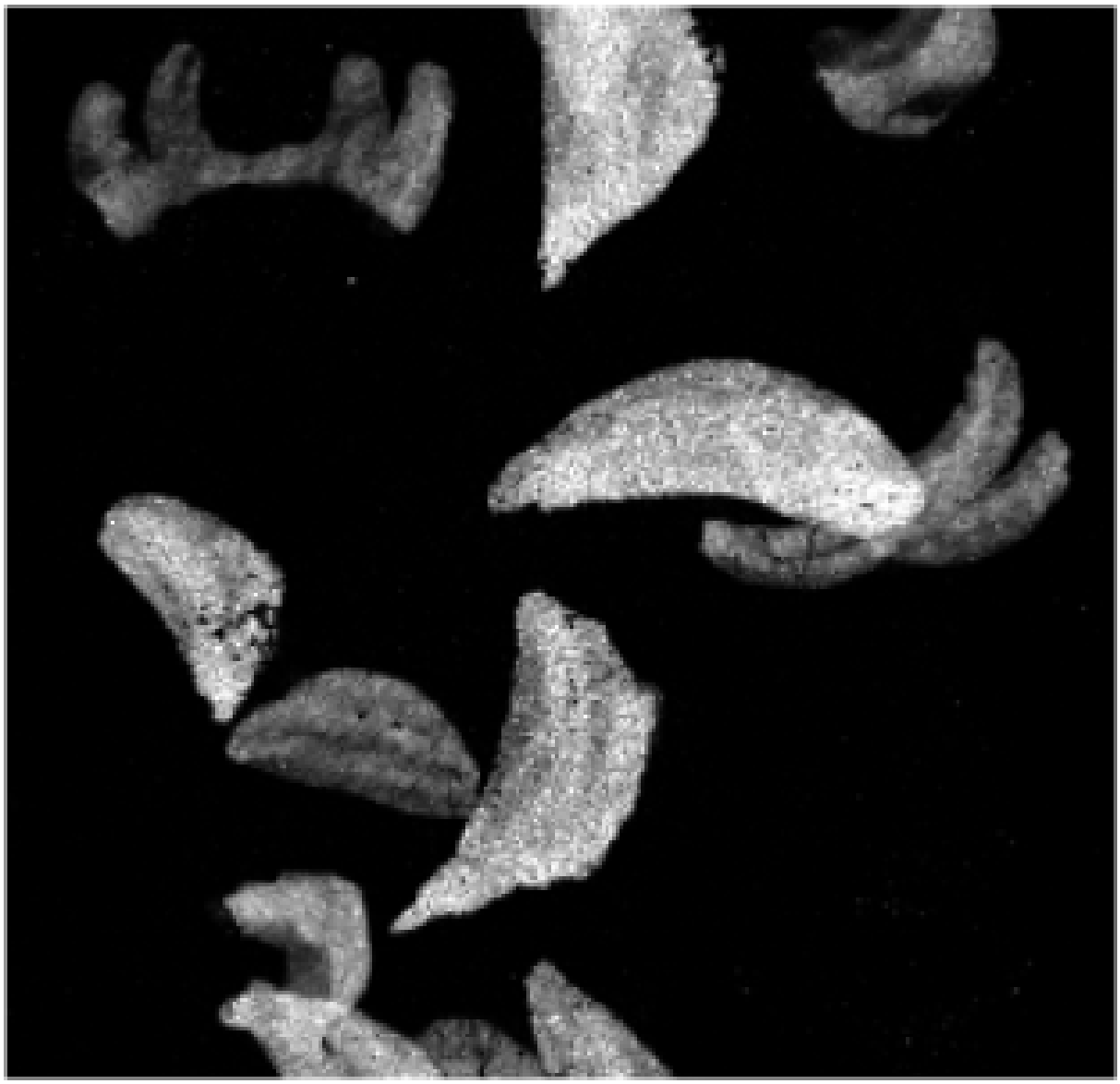


Figure 1.5. Typical X-ray distribution map. Cu-Ka X-ray image showing both surface (bright pixels) and subsurface (darker pixels) features.

mean value of this signal to simulate physical photon counting noise. Finally, a second image is added to the rear which is also realization of a Poisson random variable giving physically realistic random variation.

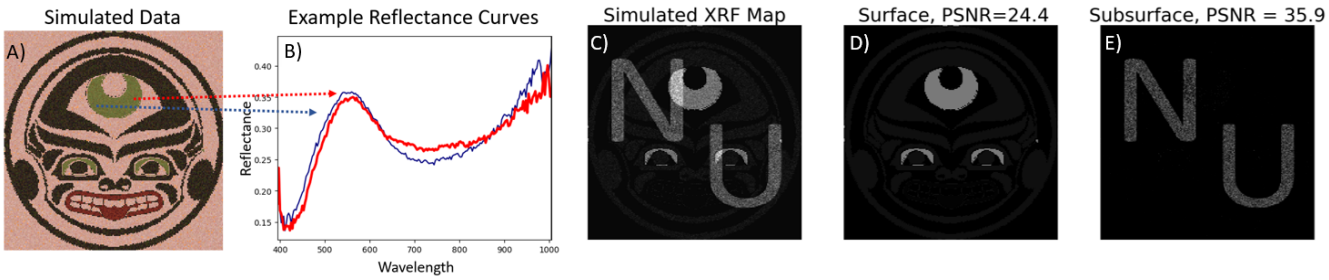


Figure 1.6. A simulated RIS image generated by reassigning pixels from the Book of Tides RIS data cube to an simulated RIS image of a painting of the ICIP 2021 logo and simulate an XRF by assigning each cluster a mean XRF value and adding a subsurface image. A) shows an RGB reconstruction of the simulated RIS data cube, B) examples of two green reflectance curves located at 2 different pixels. The curves have similar shapes but also contain slight differences in intensity and noise. C) shows the simulated XRF signal under Poisson noise, D) a surface XRF estimate with K=5 clusters obtained by computing the average for each cluster, E) a subsurface XRF estimate.

1.5. Results

1.5.1. Simulated Data Results

The results on the simulated data are shown in figure 1.6. In this simulation it is qualitatively observed that the reconstructed surface estimate matches with the true surface XRF signal on average. Due to the fact that this is a simulation there are no uncorrelated surface effects or other sources of error so the shifting routine can be skipped. The resulting delayering is very close to the original image. You can see some residual signal remaining particularly in the regions of strong XRF signal in the green areas. The overall PSNR for the surface and subsurface is good at 24.4 and 35.9 respectively. The subsurface performance is aided in the fact that the physical threshold makes the pixel values match exactly where the image is 0.

1.5.2. Book of Tides Results

The delayering results for the complete dataset is shown in figure 1.7 in which our proposed method is compared to the previously acquired confocal XRF measurements made on the same area and visible images from the front and back of the page (see **[65]** for complete results and conditions of acquisition). In figure 1.7A, the total overlapping XRF signal for the element Cu is observed. Our prediction for the top layer and bottom layer are shown respectively in figure 1.7B and C. In comparison with the previously acquired confocal XRF image of the same area (figure 1.7F) it is found that our result compares surprisingly well with this complex and costly method. Likewise, compared to the visible images (figures 1.7D and E), our algorithm may be seen to faithfully reconstruct the copper signal coming from the front and back of the page.

1.6. Discussion

This chapter considered the problem of virtually delayering XRF elemental maps through the incorporation of RIS data. The problem of interpreting XRF data in layered media is of interest to cultural heritage scientists because it can allow for easier analysis of layering structure and the imaging of underpaintings. An approach which uses an initial clustering step and the identification of a cluster-wise response dictionary which can then be used to estimate the surface image was proposed and tested in three different data sets.

This approach of cluster and compare is computationally simple and easy to implement and provides the end user with an easy to interpret delayered result with few parameters. This serves as an proof of concept that information of the surface XRF intensities can

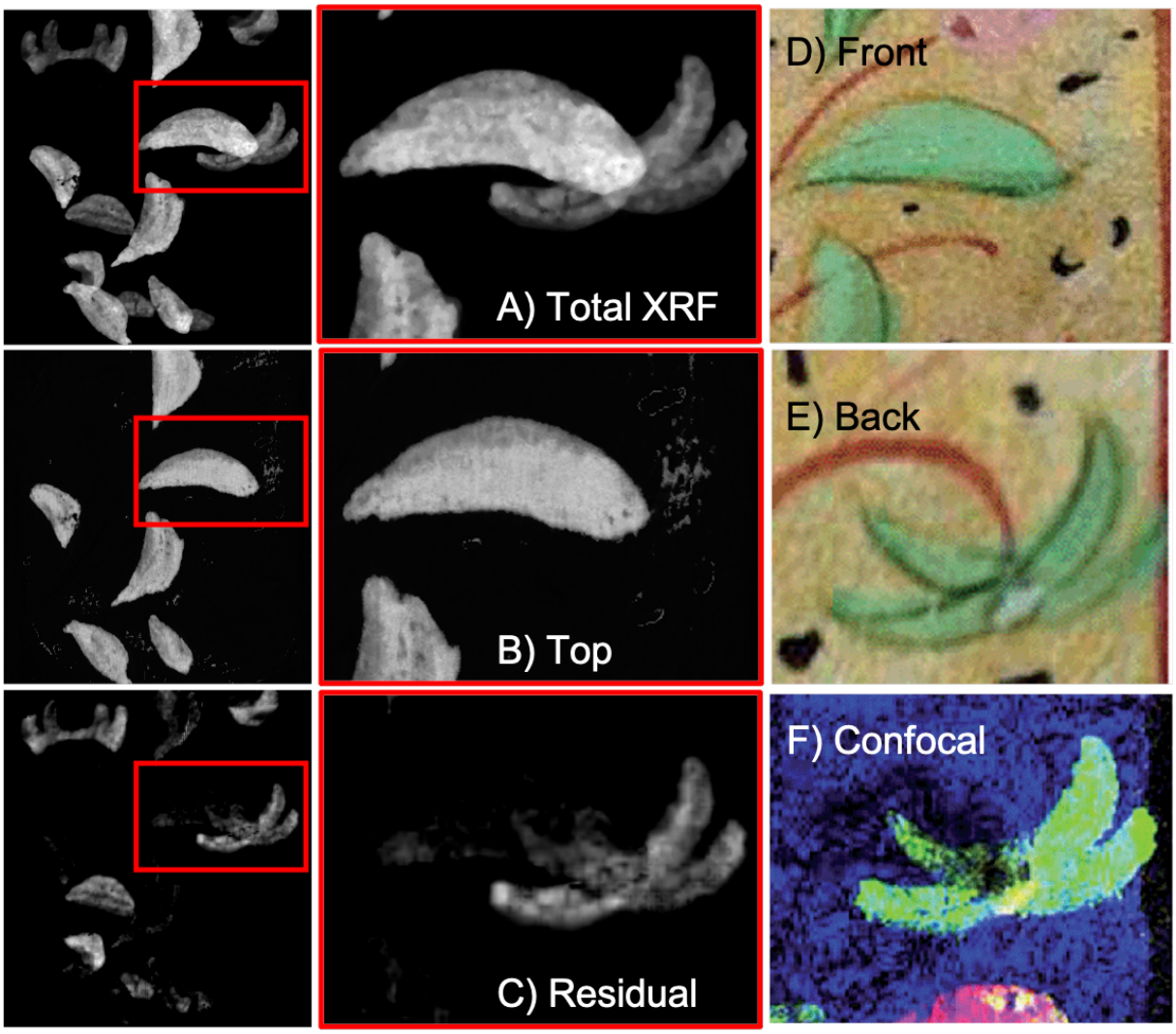


Figure 1.7. Recovery of signal from an illuminated manuscript which is obscured by surface level features. A) Total Cu K XRF signal, B) estimate of Cu top layer, C) estimate of Cu bottom. D and E show visible feature of Front and Rear of manuscript. F) Cu map produced by confocal scan of same area.

be gained by examining clusters in registered RIS data cubes. However, the method only can detect correlations between the XRF signal and the RIS signal so in situations where there is surface signal which is not correlated to RIS spectra the method fails. This occurs

in the case of a thick layer of paint with thickness variations visible in the XRF signal. Furthermore, the method is sensitive to misregistration between the XRF and the RIS data. Which can be a difficult problem in and of itself due to contrast and resolution differences between the two measurements.

This method provides an easy to implement and effective way to estimate delayering for cultural heritage applications involving 2 layered paintings. However, we use a very brute force clustering algorithm and ignore the underlying distributions of the XRF measurements. Several frameworks can more directly handle this sort of statistical prior such as Gaussian Mixture model or a neural network with an appropriate cost function.

While these results can and should be refined, the results demonstrate the difficulty in extrapolating 3D information from a 2D measurement. The conceptual framework of treating a painting as a series of independent depths can sometimes be appropriate but it is nonphysical. An example of a situation where such a conceptual model would not work is when studying objects such as glass particles [**10**] which are embedded within a layer. In situations such as these an understanding of the depth of the particles within a layer will need to be studied with a true 3D imaging method such as Optical Coherence tomography, as discussed in chapter 2.

CHAPTER 2

Optical Coherence Tomography Attenuation Imaging with Speckle Present

2.1. Introduction

Both imaging modalities discussed in chapter 1, are concerned with measuring 2 dimensional projections of an object. When studying questions surrounding stratiography, morphology, shape or structure the relevant information is inherently 3 dimensional. While clever tricks can be developed to infer this depth information from a 2D imaging modality, these measurements often will not contain enough information to address the question being asked. In these cases a true depth resolved imaging modality is needed. Optical Coherence Tomography (OCT) is an imaging technique which allows for the visualization of internal structures of tissues and other translucent materials volumetrically. It manages to do this with a very low dosage of (often) visible light making it appropriate for sensitive applications such as imaging the eye [**10**] or studying the structure of varnish and glazes in old master paintings [**10, 9**].

OCT was developed as an imaging technique in the seminal paper by Huang and Fujimoto [**55**] by extending work in 1-dimensional partial coherence interferometry developed by Fercher [**40, 41**]. Later, the technique was refined again by Fercher into Fourier Domain OCT which uses wavelength diversity to reconstruct an image faster and with less noise [**39**]. OCT creates these depth resolved images through interferometry, effectively measuring where the light is being reflected from the phase of the back-scattered field at the detector. Due to the wavelike nature and constant speed of light the phase of the light carries depth information for where in the sample the light was reflected allowing an image to be constructed where the pixel values are related to the amount of light back-reflected at each point. Modern OCT is a workhorse for visualizing layered media

by probing depth in one dimension called an A scan, generating a virtual cross section called a B scans, or by measuring full volumes called C scans.

In applications such as cultural heritage science or medical imaging, insights on material structure and makeup can be gained by extracting quantitative material parameters from the imaged sample. In particular, the depth resolved attenuation coefficient has been related to tissue structure in cancers [**116**]. However, because OCT is a technique which uses the scattering of coherent light, the measurements suffer from large and effectively random intensity fluctuations due to self interference called speckle.

The goal of this chapter is to construct and validate the *posterior distribution* for the voxelwise attenuation coefficient in the presence of random fluctuations due to speckle. The posterior distribution assigns a meaningful probability to every possible value of the true attenuation coefficient. Here, the true attenuation coefficient is defined as the attenuation coefficient of the mean OCT signal without speckle fluctuations. Using an existing depth resolved method [**114**], the attenuation coefficient at each voxel can be estimated from the measured OCT signal. These estimates depend on the intensity at each voxel which fluctuates due to speckle. Because of these voxelwise fluctuations, the estimated value of the attenuation coefficient at that point will likely differ from the true coefficient. The posterior probability distribution gives the probability that the true value of the attenuation coefficient is equal to μ_{oct} given that our depth resolved estimate from the measured image was equal to $\hat{\mu}$. Utilizing the posterior distribution allows for the identification of estimation biases and the quantification of uncertainty by giving access to statistics about the inferred attenuation coefficient. A better understanding of uncertainty can have direct clinical implications by helping to inform practitioners of how much they

can trust a given inference. Furthermore, this approach opens the door to probabilistic tissue classification tasks such as tumor grading where the likelihood of various outcomes must be compared.

2.1.1. Chapter Outline and Structure

Section 2.2 provides background and examples on OCT and attenuation analysis. A description of back scattering as a contrast mechanism is given in 2.2.1. Two applications of Fourier Domain OCT to the imaging of painted works of art are discussed 2.2.3 with example images of real OCT scans. Before the posterior and prior distributions of the attenuation coefficient can be derived we must first have a mathematical model for the measured OCT signal so we can make depth resolved attenuation coefficient estimates. In section 2.3 a model which describes the mean signal decay is given. This model assumes that the measurements are made on a weakly absorbing medium and that the majority of measured light is single scattered. Next, in section 2.4 the effect of speckle on this OCT signal is considered and the probability distribution for the measurement is given.

The likelihood function is derived in section 2.5 by analyzing the speckle variations and is verified experimentally in section 2.6.1 by measuring the distribution of depth resolved attenuation coefficient estimates for a very homogeneous phantom. In section 2.5.1 the prior probability, $P(\mu_{oct})$, is derived using basic physical principles. This prior gives the background probability for finding a particular value of μ_{oct} at any point in the sample without any additional measurement information. Following this we define a Bayesian estimator for the attenuation coefficient in section 2.5.3. In section 2.6.2 we

simulate OCT signals with realistic variations to test our assumptions and statistical model. Finally, section 2.7 will provide a thorough discussion and propose future work.

2.2. Background

2.2.1. Scattering Coefficients

Fourier domain OCT is a technique which senses the Fourier transform of the depth resolved changes in index of refraction by measuring interference patterns of the back-reflected electric field at many different wavelengths. A rigorous derivation of this method is presented in appendix A.1. For many applications of OCT, the light scattering occurs not from index of refraction variations within one media, but instead from small particles of a different refractive index embedded in the material itself [**11**]. These particles are much much smaller than the imaging resolution of the OCT with any realistic laser bandwidth. As a result, the reconstructed intensity for an OCT depth scan is proportional to the amount of scattering that the aggregate group of particles within a given imaging voxel. This is generally quantified using a single coefficient called the scattering coefficient $\mu_s[L^{-1}]$ which is defined as the amount of light which is scattered out of a path within a given unit length.

Light which is scattered is scattered in all the directions, and the distribution of the amount of light scattered in each direction, called the phase function of the light, depends highly on the particles size and shape. Often, it is assumed that light scatters isotropically although many alternative scattering models exist. From this alternative and more aggregate perspective, not all of the scattered light will reach the detector as some proportion of the energy will be scattered forward or in a direction which does not enter

the camera aperture. This is frequently modeled as a coefficient called the back-scattering coefficient $\beta_{NA}(z)$ as shown in figure 2.1. This coefficient is dependent on the optics of your OCT system, and in particular the numerical aperture (angle of acceptance) of your lens.

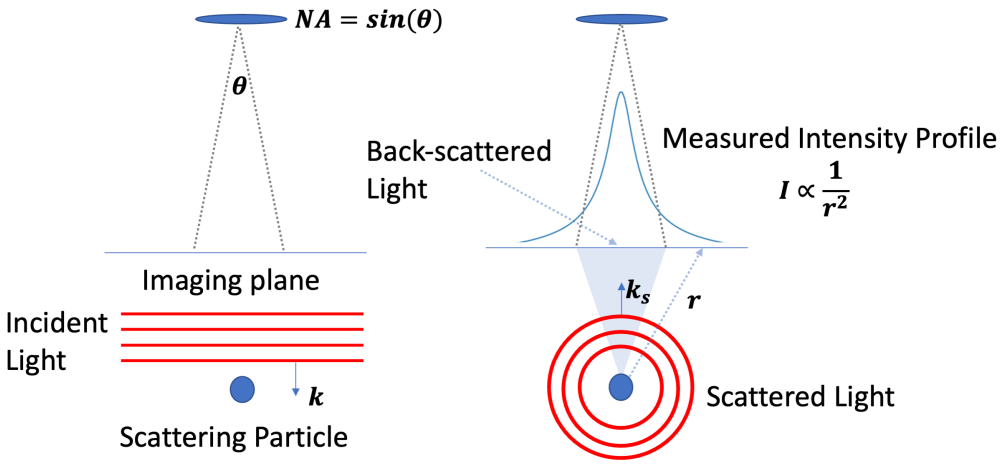


Figure 2.1. A demonstration of the impact of numerical aperture on the ratio of back-scattered light. The NA of the lens controls the acceptance angle of the optics. As light is scattered a certain fraction of light will arrive at the imaging plane. Only a small subset of that light can be detected by your optics. The ratio of the detected intensity and total scattered intensity is β_{NA}

2.2.2. OCT Attenuation Analysis

While qualitative visualization is sufficient to solve many problems there is interest in developing new quantitative metrics which can be used to make inferences in a more robust and reliable way, particularly in the medical world. One such quantitative measure is OCT attenuation analysis. Assuming a constant amount of scattering, as light penetrates

more deeply into a medium the total amount of light there which can be back reflected is lessened due to previous scattering and absorption events reducing the incident intensity leading to a decrease in signal. The precise rate of this signal decay is controlled only by the scattering and absorption of the medium. These are material properties which are imaging system independent and can be correlated with material structure and composition. Currently, methods to extract the attenuation coefficient fall into one of two categories: layerwise extraction through curve fitting [**38**] and depth resolved or voxelwise extraction [**24**].

In the layerwise approach, the layers of media are segmented, and then an exponentially decaying model is fit to each A-scan of the OCT signal in the least squares sense [**38, 24, 82**]. From this perspective, the attenuation coefficient is a bulk measure which assigns a single, deterministic number to each segment of an A-scan. However, a measured A-scan will contain fluctuations due to speckle [**96, 60**]. OCT speckle is the voxel-to-voxel variation of OCT amplitude, due to random variations in the spatial position of scattering particles within the imaging voxel. Randomly placed scatterers within the voxels will thus return scattered fields with random amplitude and phase – leading to intensity fluctuations at the detector. While the origin of speckle is deterministic at the microscopic level, in practice the measured signal is well modeled as a realization of a random process equivalent to randomly varying the exact microscopic position of the scattering particles in the bulk of the media [**11**]. One common technique to overcome the speckle variations is lateral averaging [**111, 82, 36**], where neighboring A-scans are averaged together prior to fitting. Lateral averaging can be an effective technique at reducing speckle variations but at a severe cost to lateral resolution. If the sample is not perfectly static, as is the

case in liquid samples with particles undergoing Brownian motion or sufficiently dynamic living samples, consecutive A-scans taken at the same location can be averaged together to reduce speckle variations at the cost of effective acquisition time [**119**]. In either case, the layerwise fitting assumes complete uniformity in the composition and statistics of the layer segment in depth and lateral averaging makes the same assumption over a volumetric region.

A depth resolved (DR) approach, initially developed for ultrasound image quantification [**56**], was adapted by Vermeer for use in OCT and has become popular in recent years [**114, 120, 100**]. This approach removes the assumption of material uniformity in depth and allows variations in the attenuation coefficient in three dimensions. The DR approach assumes the material is weakly absorbing, making this technique related to voxelwise OCT scattering parameter inference methods[**61, 1, 110**] which have a long history in OCT signal processing. This method has been further refined by Liu [**70**] to better handle boundary effects caused by finite imaging depth. In either formulation, reconstructions of the attenuation coefficient will be highly variable due to the influence of speckle [**100**]. Thus, as before, lateral averaging is often still employed to get a more consistent result [**114**]. Conceptually, the DR approach allows one to recover some amount of the natural variability of optical properties within the tissues. While the advantages of the DR approach are manifest, the result of this approach in the presence of intensity variations due to speckle leads to reconstructions in which the recovered attenuation coefficient itself has large variations.

The propagation of speckle variation into the recovery of an otherwise deterministic coefficient has clear implications for the accuracy of the attenuation parameter inferred at

a single voxel. Since the exact measured intensity is effectively random, one can in general expect the inferred coefficient to be effectively random as well. One way to handle the inference of parameters in these circumstances would be to adopt a Bayesian perspective. In this paradigm, instead of simply seeking an estimate for the value, one seeks the posterior distribution, which quantifies how probable each attenuation value is [**104**]. In these methods, accurate physical models about measurement uncertainty are combined with prior information about the objects which are being measured.

2.2.3. Two Applications of FD-OCT to Cultural Heritage Science

This subsection summarizes work done as a part of a collaboration with the Rijksmuseum. I would like to thank the staff Rijksmuseum, particularly Erma Hermens and Mitra Almasian, for allowing me to reproduce these figures to include in this section.

OCT is primarily thought of as a medical imaging technique and sees consistent use in optometry. However, as a non-contact and low power imaging modality, OCT has become increasingly popular to address questions in cultural heritage science. In particular, glazes and varnishes [**10, 9**] are frequently translucent in the wavelength ranges of most common OCT systems and often contain embedded features and complex layered structures which OCT is well suited to image. In this section two interesting case studies are addressed by using Fourier Domain OCT coupled with an image processing pipeline which we developed for cultural heritage purposes.

2.2.3.1. Imaging Ground Glass Embedded in Glazes. Ground glass was added to paint layers for its sicative (oil-drying) properties throughout 15th-17th centuries [**10**].

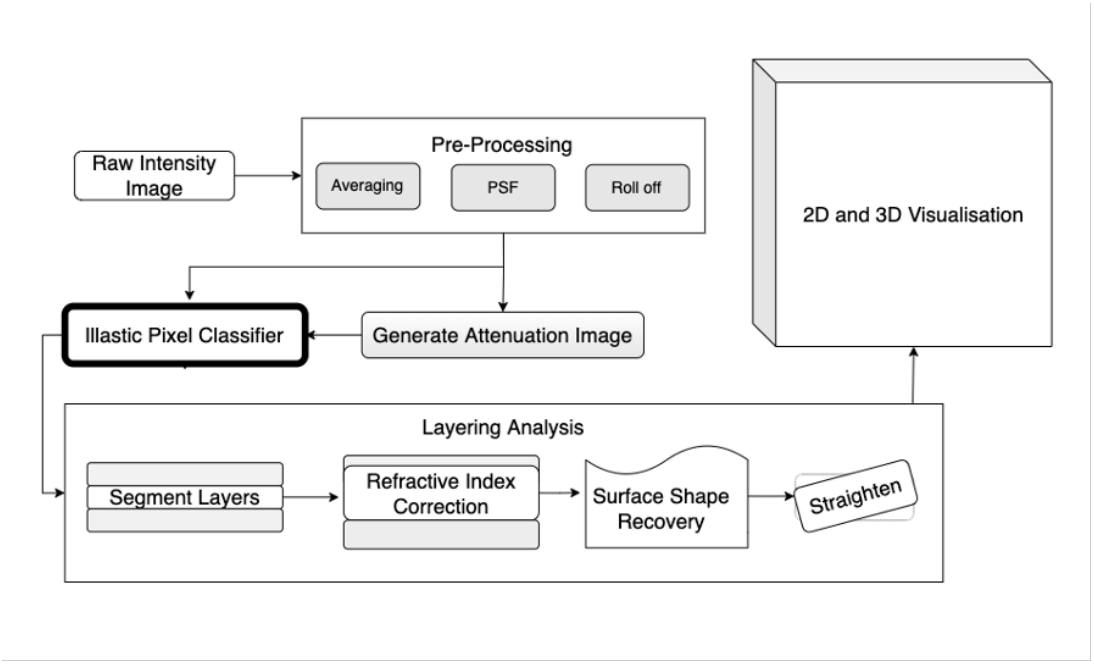


Figure 2.2. An outline from the image processing pipeline developed for OCT images. Initially, some preprocessing and corrections are done to smooth and normalize the signal. After this the OCT layers and or embedded particles are segmented and optionally a depth resolved attenuation image is generated which can increase contrast. Following this, a pixel-wise random forest classifier is trained and outputs a series of probability maps. Layers are then segmented using a maximum likelihood approach and the segmentations are smoothed. Following this an index of refraction correction is applied to each material and the layer thickness and geometry is computed. Finally, any tilt between the OCT and measured object is removed and the visualizations are generated.

When ground very finely, this can have the additional effect of increasing color depth in certain red lake glazes [**73**]. However, altering the composition of the glaze layer could potentially change the hardness or robustness of the layer. The changes in mechanical properties can potentially impact the way the layer changes over time or must be handled. While many example of painted works of art with glaze layers that contain ground glasses exist historically, the available information on the ground glass and glaze recipes comes

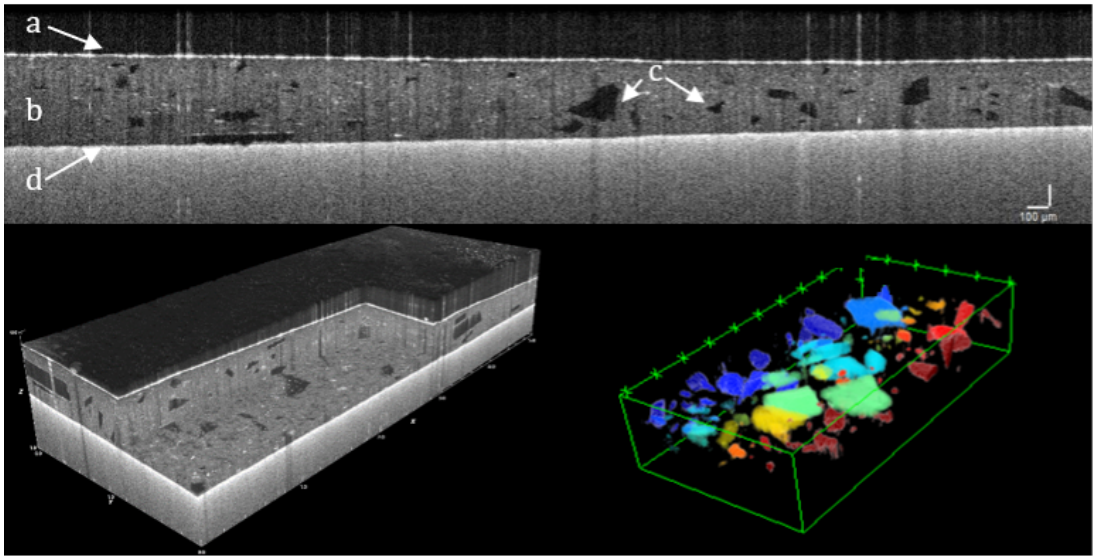


Figure 2.3. Example OCT C-scan **[10]** (volumetric image) of a mock-up glaze layer with embedded glass particles. Clear contrast difference can be seen between the glass (c) and the surrounding medium (b) due to the glass particles not having any volume scattering making them appear darker than the surrounding glaze layer. The layer interfaces (a,d) are bright due to a rapid index of refraction change. The segmentation of the glass particles can be seen on the lower left.

from recipes calling for the glasses to be hand ground confounding any measurements taken. To investigate the mechanical properties of layers we created a series of sample reconstructions following historic recipes. The mechanical properties such as hardness could then be measured using microindentation. However, even controlling the creation of the ground glass as well as possible there was ambiguity in the size and shape distribution of the glass as well as the specific glaze layer thickness.

To gain insight on the size, shape, and distribution of the settles glass in the glaze layer OCT was employed. The resulting images, shown in figure 2.3, show clear contrast between the glass particles and the background glaze layers. However, the images must be analyzed so the glass shape, location and size can be studied. To accomplish this task,

the OCT images were processed using an off the shelf random forest classier called Illastik [**13**]. This classifier creates probability maps, where the probability of each pixel belonging to a specific layer, boundary, or glass particle is measured. These probability maps are then analyzed in Matlab where a rough segmentation is produced using a max likelihood approach and then refined with assumptions about layer continuity and uniformity. The resulting segmentation is then corrected using information about the refractive index of the glaze and glass which can have a large impact on the measured geometry. The segmentation is then processed to extract statistics from the segmented layers such as thickness and shape, as well as the segmented particles particularly the distribution of size and location. The image processing pipeline used here is shown in figure 2.2.

Comparing the size and shape distributions for the ground glass showed a relationship between the total surface area, which is inverse to the particle size which also impacts the mechanical proprieties of the layer. The more homogeneous particle size distribution further increased elastic and viscous moduli of the dried red glazes. Furthermore, it could be seen that the distribution and locations of the particles in the layer were relatively uniform. Showing in the short term little setting or particle migration occurred. Further scientific details can be seen in the full publication [**10**].

2.2.3.2. Studying Varnish Degradation and Thickness in the Night Watch. The Night Watch is a large and famous painting by Rembrandt van Rijn. This painting is arguably the most famous work by Rembrandt and is a part of the rich cultural heritage that the Netherlands relies on for tourism. While this painting is widely beloved, the attention of this painting often makes it a target, which has lead to several attempts at vandalization including two separate knife attacks [**106, 84**] and an acid attack [**107**].

Figure 2.4. A 2D virtual cross section (B-scan) done in a small region of the shirt of Willem van Ruytenburch [**27**]. The varnish layers are clearly visible as well as the surface topography of the paint layers.

The damage to the painting from these attacks has largely been mitigated through hard fought conservation and restoration efforts. One such effort is currently under the way called *Operation Night watch* [**94**]. The goal of [**94**] is three-fold. First, this effort is

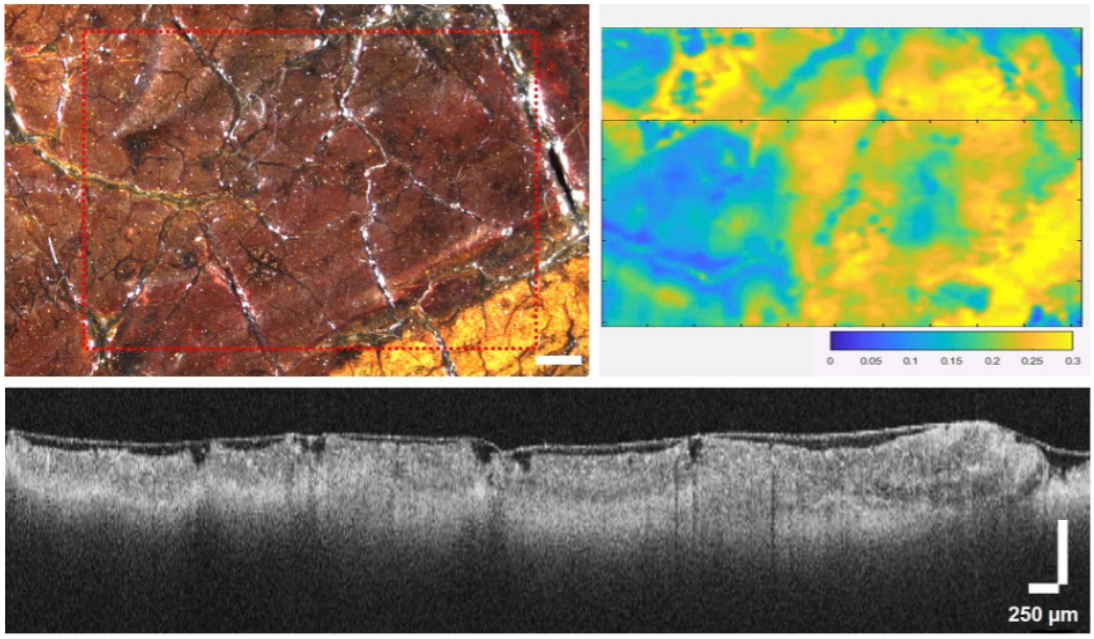


Figure 2.5. A small region of the Night Watch painting which exhibits surface cracking. The cracks can be studied my measuring the surface height using OCT as shown in the top right panel. A virtual cross section (B-scan) clearly visualized the depth and width of the cracks.

restorative with the structure of the painting and the varnish are being repaired. Second, to help connect the broader public to cultural heritage the repairs and measurements are being widely publicised and live streamed. Lastly, and most importantly to this section, the painting is being thoroughly measured and studied by a large team of cultural heritage scientists.

This team of cultural heritage researchers have taken broad variety of measurements including the reflectance spectroscopy and XRF discussed in the previous chapter, a variety of chemical analyses and even ultra high resolution photography at roughly 717 gigapixel. Additionally, to address questions surrounding surface shape, layering, and varnish OCT has been employed. Due to the fact that the painting is roughly 12 ft x

14 ft and OCT has a microscopic field of view the entire painting could not be scanned practically. Instead, 70 separate regions of this room-sized painting were imaged and the subsequently analyzed using the processing pipeline shown in figure 2.2. The regions were chosen due to the presence of damage, repairs, and interesting layering or surface shape.

These OCT scans give insight into how the painting is degrading with age. In some regions lead soaps, where chemical reactions within the paint layers are slowly forming structures which grow and pertrude from the painting are visible and could be measured. Surface cracking also specifically imaged and quantified as shown in figure 2.5. In some regions, multiple varnish layers were discovered as shown in figure 2.4 which shows an imperfect removal of varnish during a previous repair. Overall, the OCT measurements gave insight into current condition of this work of art, but as well as helping to find regions or repairs to the surface of the painting and varnish. A full discussion of this project can be seen in the publication [**9**].

2.3. Mathematical Modeling of the Intensity Decay

Intuitively, due to conservation of energy it is not possible to image infinitely deeply into a media using a technique like OCT. In fact, the amount of light which can be detected from within the media itself will decay rapidly. In most cases, the decay of the reconstructed OCT intensity with depth can be adequately described using a single exponential decay model [**8, 47, 24**]. Understanding the form of the OCT signal is necessary prior to understanding the attenuation itself. The attenuation coefficient is a material property, which depends on the absorption and scattering properties of tissue and is not a function of the measurement system. However, several system dependent

factors can also contribute to measured signal attenuation such as the confocal point spread function and the sensitivity roll off function for OCT systems based on detection in the Fourier domain [**38, 112**]. A model which takes all of these effects into account was described in detail in earlier work [**8, 112**]. Typically for an OCT system, the signal decay due to the confocal PSF and the sensitivity roll off function can be independently measured, and subsequently, the resulting OCT data can be corrected for these effects. For the sake of analysis, we will assume that the measured signal has already been calibrated for these system dependent effects. A more thorough discussion of this can be found in appendix A.3.

We denote the corrected OCT signal at depth z as $I(z; \mu_{b,NA}(z), \mu_{oct}(z))$ where $\mu_{b,NA}(z)$ is the depth dependent back-scattering coefficient (the probability per unit length that light is back-scattered into the detection numerical aperture). The depth dependent attenuation coefficient, $\mu_{oct}(z)$, and the back-scattering coefficients depend on both the scattering coefficient μ_s. These coefficients describe the probabilities of scattering and absorption per unit length, respectively. For weakly scattering samples, with negligible contributions from multiple scattered light, $\mu_{oct} = \mu_s + \mu_a$, where μ_a is the absorption coefficient.

Following Vermeer [**114**], we further assume that the tissue is very weakly absorbing ($\mu_a \approx 0$), and, a constant fraction of the attenuated light is back-scattered at every point in the tissue. We denote this fraction as β_{NA} and define $\mu_{b,NA} = \beta_{NA}\mu_{oct}$. Physically, this implies that the system is highly scattering dominant, i.e., there is very little absorbed light in the system when compared to the total attenuated light. Using these assumptions

and Beer's law we the continuous signal can be modeled as

$$I(z) = I_{inc}\beta_{NA}\mu_{oct}(z)\exp\left(-2\int_0^z \mu_{oct}(\xi)d\xi\right). \tag{2.1}$$

Indefinite integration yields

$$\int dz I(z) = I_{inc}\beta_{NA}\int dz\mu_{oct}(z)\exp\left(-2\int_0^z \mu_{oct}(\xi)d\xi\right) \tag{2.2}$$

$$= -2I_{inc}\beta_{NA}\exp\left(-2\int_0^z \mu_{oct}(z)dz\right) = \frac{-I(z)}{2\mu_{oct}(z)} + C. \tag{2.3}$$

Combining this simple observation with the physical boundary condition that $I(z) \to 0$ as $z \to \infty$ yields the parameter recovery formula

$$\mu_{oct}(z) = \frac{I(z)}{2\int_z^\infty d\xi I(\xi)}. \tag{2.4}$$

In practice, OCT measures a series of discrete samples instead of a continuous signal. Let Δz denote the fixed axial resolution of the system, then the measured intensity is given by quantity

$$I_N = I(N\Delta z; \mu_{oct}(N)) = I_{inc}\beta_{NA}\mu_{oct}(N)\exp\left(-2\sum_{i=1}^{N}\mu_{oct}(i\Delta z)\Delta z\right). \tag{2.5}$$

I_N is defined which describes the mean value of the OCT signal with depth in a certain region at depth $z = N\Delta z$ where N is the pixel index and given an incident intensity I_{inc} . For brevity, the shorthand $\mu_{oct}(N) = \mu_{oct}(N\Delta z)$ is used. Provided that the inverse of the attenuation coefficient is relatively small compared with the pixel size, its discrete value

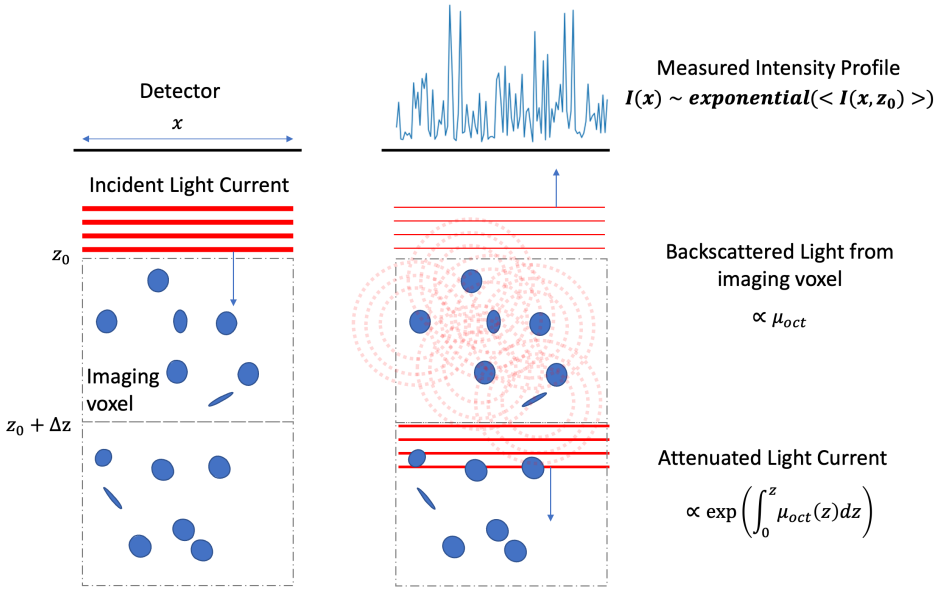


Figure 2.6. This figure demonstrates the imaging principle of OCT. An incident current of coherent light illuminates a volume and some proportion of the light is back scattered. In a weakly absorbing sample the back-scattered light and the attenuation coefficient are proportional to one another. Within a single imaging voxel there are many different scattering particles embedded leading to speckle. This speckle forces the measured intensity to be effectively randomly distributed around the mean refelectance of the voxel.

is given by [**114**]

$$\mu_{oct}(N) = \frac{I_N}{\sum_{i=N+1}^{\infty} 2\Delta z I_i}. \tag{2.6}$$

As noted by Liu[**70**], the tail of the series in the denominator in equation 2.6, meaning all of the terms in the sum after some large term K, can be computed when an estimate for an attenuation coefficient at that point in the sample is available. This is given by

$$\sum_{i=K+1}^{\infty} 2\Delta z I_i = \frac{I_K}{\mu_{oct}(K)}. \tag{2.7}$$

2.4. A Statistical Model of Speckle

For any technique using coherent light, the overall intensity is based on the amount of interference given by the path length of any surface or particle reflecting light to a point. For very smooth surfaces or systems with very few scattering particles the resulting images contain physical and slowly varying fluctuations. However, for volume scatters with many particles tiny variations in path length between the particles compounds and forms high frequency and difficult to model fluctuations called speckle [**48**]. In the case of depth imaging, when there are sufficiently many scattering events within a single voxel, the speckle is called fully developed [**54**] and the measured signal becomes effectively random.

The phase of the measured field at the detector from a randomly positioned particle will be random. In fact both the real and imaginary parts of the total field at the detector will approach i.i.d. Gaussian random variables given a sufficient number of scattering particles due to the central limit theorem with zero mean and variance proportional to the number of particles and their scattering amplitude. A more robust derivation is given in appendix A.2. Let E_z denote a random electric field such that $Re[E_z], Im[E_z] \sim G(0, \langle I \rangle)$. With this statistical model for the electric field, one can derive the distribution of the intensity of the electric field at the detector as a change of variables. Consider the function

$$A = \sqrt{Im[E_z]^2 + Re[E_z]^2} := r \tag{2.8}$$

by the change of variables formula has the PDF

$$A \sim \frac{d}{dA} \int_0^{2\pi} \int_0^A \frac{r}{2\pi\sigma^2} \exp(-\frac{r^2}{2\langle I \rangle}) dr d\theta = \frac{x}{\langle I \rangle} \exp(\frac{-I^2}{2\langle I \rangle}). \tag{2.9}$$

Since OCT is a volumetric imaging modality the measured intensity on the detector must obey these statistics. We can formally write the PDF of the detected amplitude for a given depth as a conditional distribution as

$$P_{amp}(A|\langle I\rangle) = \frac{A}{\langle I\rangle}\exp\left(-\frac{A^2}{2\langle I\rangle}\right) \tag{2.10}$$

where $\langle I\rangle = \sigma^2$ denotes the mean intensity value. This formula gives the probability of measuring amplitude A when the mean signal is given by $\sqrt{\langle I\rangle}$. When OCT measurements are made, typically intensity is measured and not amplitude. Given a Rayleigh distributed amplitude of the form given in equation 2.10 it can be shown by a similar change of variables that the intensity which is the square of the amplitude, follows [**48**]

$$P_{int}(I|\langle I\rangle) = \frac{1}{\langle I\rangle}\exp\left(-\frac{I}{\langle I\rangle}\right) \tag{2.11}$$

which is an exponential distribution with parameter $\langle I\rangle$ as shown in figure 2.6

2.4.1. The Posterior Distribution

Due to the uncertainty from speckle, it is reasonable to expect that the recovered attenuation coefficient for a material to contain some uncertainty as well. A robust understanding of this parameter must contain some way to quantify this uncertainty given a series of measurements. One way to do that is through he use of a Bayesian posterior distribution. The posterior distribution assigns a meaningful probability to every possible value of the true attenuation coefficient. Mathematically, the posterior distribution can be written as the conditional probability distribution $P\left(\mu_{oct}\middle|\hat{\mu}\right)$. This is the probability that the true attenuation coefficient is μ_{oct} given a certain set of estimates denoted $\hat{\mu}$. Conditional

probabilities can be rewritten as product of two easier to derive probability distributions using Bayes' theorem. This theorem states that the posterior distribution is given by,

$$P\big(\mu_{oct}\big|\hat{\mu}\big) = \frac{P\left(\hat{\mu}\big|\mu_{oct}\right) P\left(\mu_{oct}\right)}{P\left(\hat{\mu}\right)}. \tag{2.12}$$

In this expression, $P\left(\hat{\mu}\big|\mu_{oct}\right)$ is called the likelihood function which represents the probability of estimating $\hat{\mu}$ given that the true attenuation coefficient is equal to μ_{oct}. The distribution denoted by $P\left(\mu_{oct}\right)$ is called the prior distribution for the unknown μ_{oct}. The prior probability allows the incorporation of additional information into the statistical model and is often used as a way to establish bounds or to bias solutions towards realistic values. The marginal probability $P\left(\hat{\mu}\right)$ is a normalizing factor and can be computed via integration. Using this relation, we can find the posterior distribution by solving two easier problems: finding the likelihood function and finding the prior distribution.

2.5. Analyzing the DR Reconstruction Distribution

This section considers the estimation of the voxelwise attenutaion coefficients $\mu_{oct}(N)$ from intensity measurements in the presence of speckle modeled by equation 2.11. In this case, instead of measuring the mean intensity $\langle I_N \rangle$ directly we can only measure I_N which is exponentially distributed with parameter $\langle I_N \rangle$. Because the constituent parts of equation 2.6 are now random the estimate will be itself a random variable. The estimated random variable is denoted as

$$\hat{\mu}(N) = \frac{I_N}{\sum_{i=N+1}^{\infty} 2\Delta z I_i} \tag{2.13}$$

Following Vermeer [**114**] we consider the attenuation coefficient at the N^{th} point and truncate the series in the denominator at M which in practice corresponds to the maximum imaging depth Z_{max} with $M > N$ giving

$$\hat{\mu}(N) \approx \frac{I_N}{\sum_{i=N+1}^{M} 2\Delta z I_i}. \tag{2.14}$$

Consider the denominator, and let,

$$D_N = \sum_{i=N+1}^{M} 2\Delta z I_i. \tag{2.15}$$

The variable D_N is the sum of $M-(N+1)$ independent exponentially distributed random variables I_i, taken from distributions parameterized only with average $\langle I_i \rangle$. Thus, D_N will be distributed as a hypoexponential distribution and has mean

$$\langle D_N \rangle = \sum_{i=N+1}^{M} 2\Delta z \langle I_i \rangle, \tag{2.16}$$

because the I_i's are independent. If M is sufficiently larger than N, equation 2.7 implies that

$$\langle D_N \rangle \approx \frac{\langle I_N \rangle}{\mu_{oct}(N)}. \tag{2.17}$$

It is known that reconstruction artifacts [**70, 114**] make the inferred coefficient unreliable near the deepest point of an A-scan. In practice, the reconstructed attenuation coefficient made from this approach must be discarded near the bottom of a scan and estimated using a different method [**70**].

One useful measure of how much a random variable deviates from the mean called the coefficient of variation, and is denoted C_v. This quantity is defined as the standard deviation divided by the mean. The variance of $\langle D_N \rangle$ is given by

$$var\left(D_N\right) = (4\Delta z^2) \sum_{i=1}^{N} \lambda_i^2$$

by the Bienaymé formula where λ_i^2 is the variance of each intensity measurement. Since each constituent λ_i is positive, the binomial theorem implies that

$$var\left(D_N\right) = (4\Delta z^2) \sum_{i=1}^{N} \lambda_i^2 \leq \langle D_N \rangle^2 = \left(\sum_{i=1}^{N} 2\Delta z \lambda_i \right)^2 .$$

Dividing through by the $\langle D_N \rangle^2$ and taking a square root gives

$$C_v := \frac{\sqrt{\left(var\left(D_N\right)\right)}}{\langle D_N \rangle} < 1.$$

In practice, we find that $C_v \ll 1$ as demonstrated in figure 2.10 and described in detail in section 2.6.

Next, letting

$$\eta_N := D_N - \langle D_N \rangle \tag{2.18}$$

allows formula 2.14 to be rewritten as

$$\hat{\mu}(N) = \frac{I_N}{\langle D_N \rangle} \left(\frac{1}{1 + \frac{\eta_N}{\langle D_N \rangle}} \right). \tag{2.19}$$

Because η has zero mean with a very small C_v one can expect $\frac{\eta_N}{\langle D_N \rangle}$ to be small. Using this as justification, consider the Taylor approximation

$$\hat{\mu}(N) \sim \frac{I_N}{\langle D_N \rangle}\left(1 - \frac{\eta_N}{\langle D_N \rangle}\right) + \mathcal{O}\left(\frac{\eta_N^2}{\langle D_N \rangle}\right). \tag{2.20}$$

At leading order, the reconstruction of the attenuation coefficient is given by

$$\hat{\mu}(N) \approx \frac{I_N}{\langle D_N \rangle}. \tag{2.21}$$

Intuitively, this means the denominator of equation 2.14 is approximately constant at the scale set by the mean. Therefore, the probability distribution of $\hat{\mu}$ will be given by rescaling the distribution of I_N. Rescaling equation 2.11 yields

$$P\big(\hat{\mu}(N)\big|\langle I_N \rangle, \langle D_N \rangle\big) = \frac{\langle D_N \rangle}{\langle I_N \rangle} \exp\left(-\langle D_N \rangle \frac{\hat{\mu}(N)}{\langle I_N \rangle}\right). \tag{2.22}$$

Next, using the approximation for the tail of D given in equation 2.17 with $K = N$ and substituting $\frac{\langle I_N \rangle}{\langle D_N \rangle}$ with $\mu_{oct}(N)$ yields the probability distribution

$$P\big(\hat{\mu}(N)\big|\mu_{oct}(N)\big) \approx \frac{1}{\mu_{oct}(N)} \exp\left(-\frac{\hat{\mu}(N)}{\mu_{oct}(N)}\right). \tag{2.23}$$

Therefore, the reconstructed coefficient at leading order will be exponentially distributed around the mean attenuation parameter. The accuracy of this estimate is demonstrated in Fig. 2.10.

This approach can be extended to the time-averaged case, where k independent co-registered measurements have been made. To do this, first the k estimates for the attenuation coefficient, denoted by $\hat{\mu}_i(N), \quad i = 1, 2m..., k$, should be constructed using equation

2.14. Then, assuming the measurements are independent, the likelihood is given by

(2.24)
$$P\left(\hat{\mu}_1(N),...,\hat{\mu}_k(N)\,\middle|\mu_{oct}(N)\right)=\prod_{i=1}^{k}P\left(\hat{\mu}_i(N)\,\middle|\mu_{oct}(N)\right)=\left(\frac{1}{\mu_{oct}(N)}\right)^k\exp\left(-\frac{\sum_{i=1}^{k}\hat{\mu}_i(N)}{\mu_{oct}(N)}\right).$$

2.5.1. Constructing a Prior distribution

In this section, a prior distribution for the variation in attenuation coefficient in a layer is derived based on physical principles. As an initial theoretical step we consider a simplified media of dispersed scattering particles with negligible absorption. Following Chandrasekhar, **[21]** it is assumed that the system is a single layer, with N_p dispersed particles throughout. Let

(2.25)
$$\gamma=\frac{[\text{voxel}]_{vol}}{[\text{scan}]_{vol}}$$

be the ratio of the volume of a single voxel to the volume of the entire scanned layer. Provided that $[\text{voxel}]_{vol}\ll[\text{scan}]_{vol}$ the probability of finding n particles inside the volume defined by a single voxel is given by the Poisson distribution

(2.26)
$$P_{vox}(n):=P(\text{n particles inside voxel})\sim\frac{\langle n\rangle^n e^{-\langle n\rangle}}{n!},$$

where the mean particle number $\langle n\rangle=N_p\gamma$. For very large particle counts, $N_p\to\infty$, the Poisson distribution is well approximated as

(2.27)
$$P_{vox}(n)=\frac{\exp\left(-\frac{(n-\langle n\rangle)^2}{2\langle n\rangle}\right)}{\sqrt{2\pi\langle n\rangle}}.$$

Since absorption can be ignored, the attenuation coefficient can then be computed as $\mu \sim \sigma_{scat}\frac{n}{[voxel]_{vol}}$ and the mean coefficient as $\langle\mu\rangle = \sigma_{scat}\frac{\langle n\rangle}{[voxel]_{vol}}$, where σ_{scat} is the scattering cross section of a particle. Therefore, $P(\mu)$ is a rescaled version of the probability distribution in equation 2.27 given by

$$P(\mu) = \frac{1}{\sqrt{2\pi\zeta\langle\mu\rangle}}\exp\left(-\frac{(\mu-\langle\mu\rangle)^2}{2\zeta\langle\mu\rangle}\right) \tag{2.28}$$

where $\zeta = \frac{\sigma_{scat}}{[voxel]_{vol}}$.

2.5.2. Deriving the Posterior for the Reconstructed Attenuation Coefficients

The posterior distribution for the attenuation coefficient at depth N can be now derived by plugging in the equation 2.24 and equation 2.28 into equation 2.12. Thus, our posterior distribution is proportional to

$$P\big(\mu_{oct}(N)\big|\hat{\mu}_1(N)...\mu_k(N)\big) \propto \left(\frac{1}{\mu_{oct}(N)}\right)^k \exp\left(-\frac{\sum_{i=1}^{k}\hat{\mu}_i(N)}{\mu_{oct}(N)}\right)\frac{1}{\sqrt{2\pi\zeta\langle\mu_{oct}\rangle}}\exp\left(-\frac{(\mu_{oct}(N)-\langle\mu_{oct}\rangle)^2}{2\zeta\langle\mu_{oct}\rangle}\right) \tag{2.29}$$

where k is the number of co-registered scans and $\langle\mu_{oct}\rangle$ is the layer mean of the DR estimates. The specific proportionality constant is given by integrating the numerator of equation 2.12 over all possible values of μ_{oct}. Considering the case where only a single independent scan can be made, the posterior distribution for the attenuation coefficient

at depth N is given by

(2.30)

$$P\big(\mu_{oct}(N)\big|\hat{\mu}(N)\big) \propto \frac{1}{\mu_{oct}(N)} \exp\left(-\frac{\hat{\mu}(N)}{\mu_{oct}(N)}\right) \frac{1}{\sqrt{2\pi\zeta\langle\mu_{oct}\rangle}} \exp\left(-\frac{\big(\mu_{oct}(N)-\langle\mu_{oct}\rangle\big)^2}{2\zeta\langle\mu_{oct}\rangle}\right)$$

This distribution describes the probability of the mean coefficient at voxel N. Assuming that each voxel is independent, a joint posterior distribution for the attenuation coefficient map for the entire A, B or C scan as

(2.31)
$$P\big(\boldsymbol{\mu}_{oct}\big|\hat{\boldsymbol{\mu}}\big) \propto \prod_{i=1}^{R} P\big(\mu_{oct}(i)\big|\hat{\mu}(i)\big)$$

where R is the total number of voxels in the scan, $\boldsymbol{\mu}_{oct}$ is an $R\times 1$ vector of true coefficients and $\hat{\boldsymbol{\mu}}$ is the $R \times 1$ vector of voxelwise estimates for the attenuation coefficient. Figure 2.7 shows two posterior distributions plotted using equation 2.30 which use two different values for the DR estimate. These examples demonstrate the impact that the initial DR estimate has on the shape and position of the posterior distribution for the attenuation coefficient.

2.5.3. Bayesian Parameter Estimator

In Bayesian formulations of parameter estimation problems, when a single number prediction for the coefficient must be made, a Maximum a Posteriori (MaP) approach is often employed **[20, 100]**. This approach gives the attenuation coefficient which maximizes the posterior distribution by solving

$$\boldsymbol{\mu}_{oct}^{MaP} = \underset{\boldsymbol{\mu}_{oct}}{\operatorname{argmin}}\ \log(P(\boldsymbol{\mu}_{oct}|\hat{\boldsymbol{\mu}})).$$

However, as can be seen in figure 2.7 for sufficiently small DR estimates, the posterior distribution becomes bimodal and the MaP estimate will nearly coincide with the low DR estimate for the attenuation coefficient. As demonstrated in figure 2.7 panel b) this peak is relatively narrow and contains little probability mass. Because of this, the maximum a posteriori is a bad representation of the entire probability distribution. The mean of the posterior distribution is agnostic to the bimodality of the distribution and provides a more stable and realistic estimate for the attenuation parameter. Therefore, when a single value estimate is desired, the quantity

$$\hat{\mu}_{mean}(N) := \int_{\mathbb{R}^+} \mu_{oct} P\big(\mu_{oct}(N)\big|\hat{\mu}(N)\big) d\mu_{oct}. \tag{2.32}$$

can be computed.

2.6. Results

2.6.1. Experimental Verification and Results

To verify the likelihood model from equation 2.23, the DR attenuation formula is applied to phantom data and a histogram is computed to compare against theory. The data was collected with a Santec IVS 2000 swept source OCT system with a central wavelength of 1309 nm, axial resolution of 12 micron in air and lateral resolution of 25.5 micron. The phantom was made by suspending silica beads manufactured by BaseClear with mean diameter of 0.47 micron and a refractive index of 1.425 in water at a volume fraction of .08. Water is assumed to have a phase refractive index of 1.32 and a group refractive index of 1.34[**50**]. Using Mie theory, the scattering cross section is given by $1.9 * 10^{-9}$ mm^2 and the total attenuation coefficient is 3.2 mm^{-1} [**8**]. This value is realistic for tissue [**47, 71**].

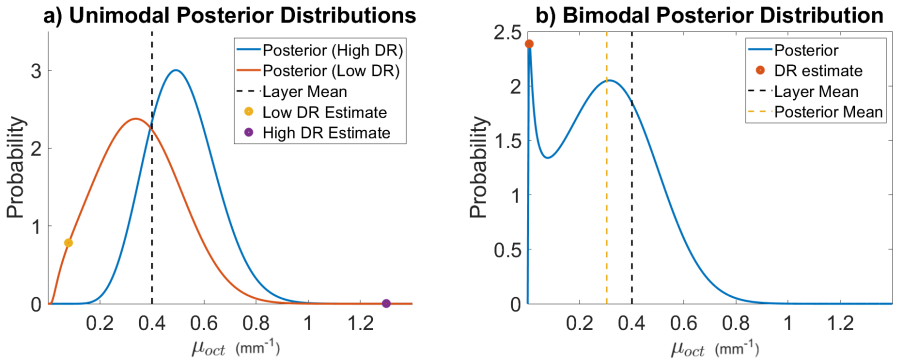


Figure 2.7. This figure shows realization of the posterior distribution for the attenuation coefficient given in equation 2.30 for different values of $\hat{\mu}$. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.
a) This panel shows two unimodal reconstructed posterior distributions. With these distributions, the true parameter is much more likely than the DR estimate. This posterior was constructed with a layer mean of $\langle \mu_{oct} \rangle = .4$ mm^{-1}, $\zeta = 6.87 * 10^{-2}$ mm^{-1} and a DR estimates of $\hat{\mu} = .08$ mm^{-1} and $\hat{\mu} = 1.3$ mm^{-1} .
b) This panel shows a constructed posterior distribution which is Bi-Modal and has two local maxima. For a given layer mean, the constructed distribution develops a second peak if the DR estimate used to construct the posterior is sufficiently small. This second peak can make the Maximum a Posteriori difficult due to non-convexity. In many cases, the maximum value of the Posterior distribution may sit very near the origin on this second peak. As demonstrated in this panel, often the total amount of probability mass under the peak near zero is relatively small, despite being the maximum likelihood. Thus, the Maximum of the posterior distribution is a poor representation for the distribution itself. In these cases an estimate for the mean is a better choice. This posterior was constructed with a layer mean of $\langle \mu_{oct} \rangle = .4$ mm^{-1}, $\zeta = 6.87 * 10^{-2}$ mm^{-1}, and a DR estimate of $\hat{\mu} = .015$ mm^{-1} .

An OCT B-scan of the phantom is shown in figure 2.8 panel a). Using these values and equation 2.28 we can see that the expected variance for the attenuation coefficient is $\langle \mu_{oct} \rangle \cdot \zeta = 0.0020$ mm^{-2} which is very small when compared with the variance of the exponential distribution which is $\langle \mu_{oct} \rangle^2 = 11.5$ mm^{-2}. Since the speckle variance

dominates the distribution of attenuation coefficients the reconstruction should look like equation 2.23. This is demonstrated in figure 2.8 panel c).

Figure 2.9 demonstrates the effect of the posterior mean estimator defined in equation 2.32 when compared with lateral averaging. In figure 2.9 panels a) and b) show the OCT attenuation coefficient B and A-scans respectively generated from the same OCT B-scan used in figure 2.8. This phantom is very homogeneous so we expect that the variation is almost entirely generated from speckle, thus it is reasonable to assume if sufficiently many A-scans are averaged together then the resulting attenuation coefficient should look constant. Figure 2.9 panel d) shows the resulting OCT attenuation coefficient after laterally averaging 1000 A-scans together. Figure 2.9 panel c) shows the result of the mean estimator defined in equation 2.32 applied to the A-scan from panel b). There is little remaining variation in the signal when compared with standard lateral averaging.

2.6.2. Simulated Results

To validate and better understand the statistical model, a series of simulations were preformed. In figure 2.10 panel a) a B-Scan was simulated using equation 2.5 with $\beta_{NA} = .3$, $I_{inc} = 1e7w/m^2$, lateral resolution of $\Delta x = .022mm$ and $\Delta z = .0068$ mm in a 3.4 mm deep domain with a fixed attenuation coefficient of $\mu_{oct} = 2.00$ mm^{-1}. Once the deterministic signal is modeled we generate the OCT signal per voxel as a realization of an exponential random variable with parameter given by the true coefficient as in equation 2.11. This random realization can be seen in figure 2.10 panel a). The attenuation coefficient was estimated using the DR method given in equation 2.14 and is shown in figure 2.10 panel b). The reconstruction equation becomes inaccurate near the bottom of the measurement

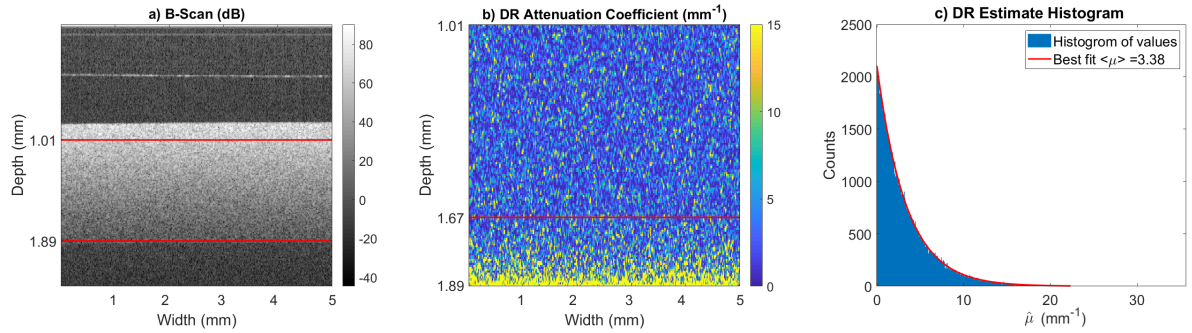


Figure 2.8. This figure demonstrates the validity of the likelihood function given in equation 2.23 by considering the distribution of attenuation coefficients for a very homogeneous phantom. Both the analysis and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.
a) An OCT B-scan of a phantom made by suspending silica beads with mean diameter of 0.47 micron and a refractive index of 1.425 in water which has a phase refractive index of 1.32 [**50**]. The red lines indicate a homogeneous region where the DR estimate is made.
b) The DR estimate for the attenuation coefficient of the B-Scan shown in a). The overestimation artifact is clear towards the bottom part of the scan.
c) Histogram of estimated value for top 100 rows of pixels of the DR estimate on the B-scan. Because the phantom is very homogeneous we expect the histogram to follow equation 2.23 for this region. The exponential fit is in good agreement with the theoretical predicted value of 3.2 mm^{-1} .

volume preventing accurate estimation. To avoid these inaccuracies the deepest 30% of the reconstructed attenuation coefficients were truncated. The 30% value was arrived at by inspection. In figure 2.10 panel c) we fit an exponential model to the histogram of the reconstruction and see that the best fit parameter agrees with our model to the 2nd decimal point.

To avoid artifacts the bottom 30% of the predicted attenuation coefficient is discarded. Figure 2.11 panel c) shows a posterior mean estimate for the attenuation coefficient which was computed with equation 2.32 voxelwise. In general, the mean attenuation coefficient for the layer, $\langle \mu_{oct} \rangle$, would not be known ahead of time to compute the prior distribution.

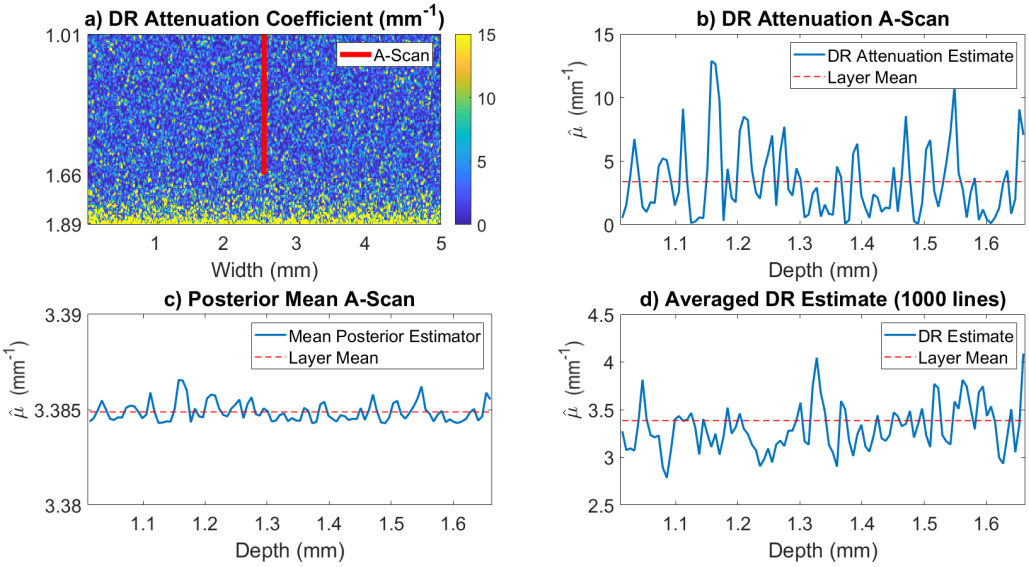


Figure 2.9. This figure demonstrates variability of the DR attenuation estimates in the presence of speckle and the Bayesian estimator for the attenuation coefficient given in equation 2.32 Both the analysis and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.
a) The DR estimate for the attenuation coefficient of the B-Scan shown in figure 2.8. The red line is the portion of the scan considered in b).
b) Plot of the DR estimate for A-scan extracted from a). This estimate shows how highly variable the DR attenuation coefficient appears to be in the presence of speckle.
c) Mean of the Posterior estimate for the extracted A-scan. This was computed using the $\langle \mu_{oct} \rangle = 3.38$ mm^{-1} which is the layer mean for the first 100 rows of pixels of the B-scan., The value of ζ was computed to be $\zeta = 6.0053 * 10^{-2} mm^{-1}$ from a voxel volume of $3.2 * 10^{-6}$ micron and a scattering cross section of $\sigma_{scat} = 1.9 * 10^{-9}$ mm^2.
d) Comparative DR estimate for the average of all 1000 A-lines in the B-scan. The resulting fluctuations are still very large even after averaging 1000 A-scans

To account for this, we used the mean of the truncated DR attenuation estimate for the whole scan in equation 2.32. The estimate given by the mean of the posterior distribution for the attenuation coefficient can give much more accurate estimates for the true coefficient than using the standard DR technique, as demonstrated in figure 2.11.

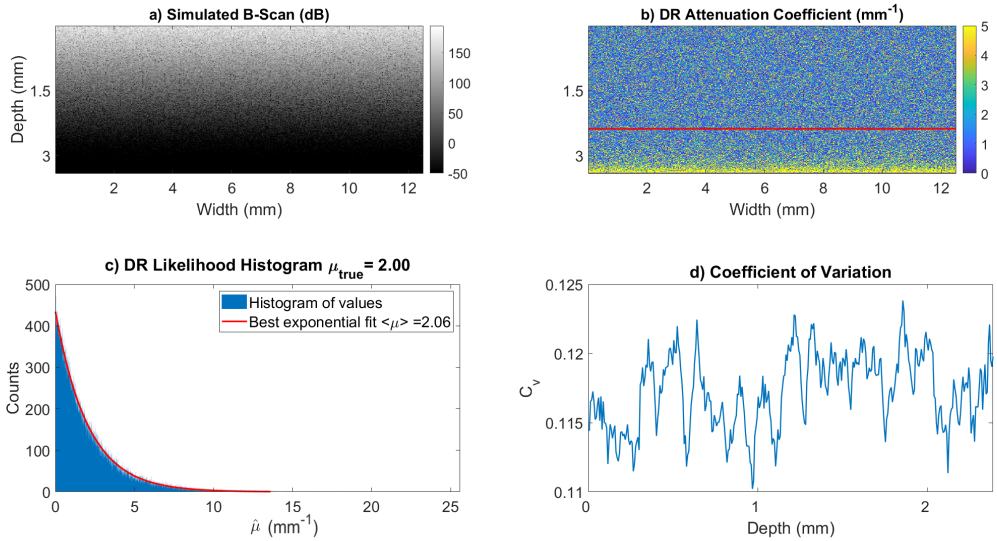


Figure 2.10. This figure demonstrates the accuracy of the likelihood model and estimates the coefficient of variation of the signal to verify the assumptions made in 2.20 for simulated signals. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.
a) This panel is a simulated B-Scan, which was simulated using parameters $\beta_{NA} = .3$, $I_{inc} = 1e7$, $\mu_{oct} = 2.00$ mm^{-1}, lateral resolution of $\Delta x = .022mm$ and $\Delta z = .0068$ mm in a domain which is 3.4 mm deep.
b) This is the DR reconstructed coefficient map. The reconstruction is highly variable around the true attenuation value of 2.00 mm^{-1}. This panel also demonstrates the growth artifact in the bottom 30% of voxels where estimated values which are much higher than the true value. The estimates below the red line are truncated to avoid the exponential grown artifact.
c) This figure is a histogram of the top 70% of pixels from panel b). As shown in equation 2.23 we expect this to be exponentially distributed with parameter 2.00 . A best fit exponential demonstrates this is accurate to 3 significant figures.
d) Coefficient of Variation for the simulated OCT A-scans in panel a) at different depths. In these simulations, C_v stays near .12 at all admissible depths.

2.7. Discussion

This chapter addressed the impact of speckle fluctuations on the depth resolved recovery of the OCT attenuation coefficient. When making an OCT measurement, effectively

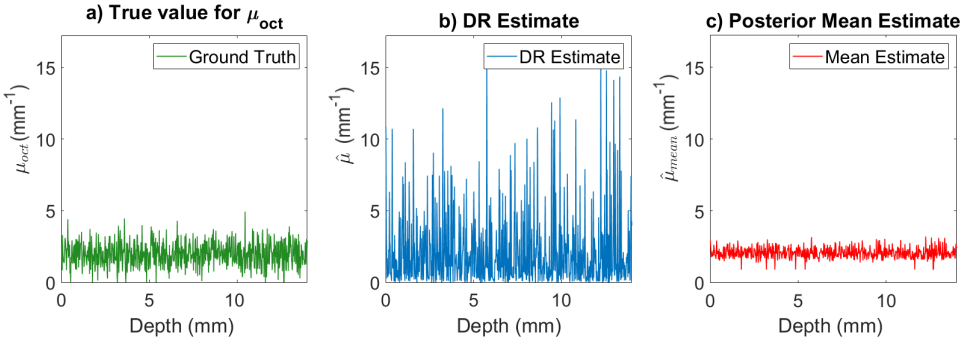


Figure 2.11. This figure shows estimates of the attenuation coefficient for simulated OCT data using the standard DR and the Bayesian estimator given in 2.32. The OCT data was simulated with parameters $I_{inc} = 1e7$, $\mu_{oct} = 2.00$ $\beta_{NA} = .3$, $\sigma_{scat} = 1.0 - 6$ mm^2, a lateral resolution of $\Delta x = .02$ mm and $\Delta z = 0.0068$ mm in a domain which is 13.6 mm deep. After the attenuation coefficient was inferred using the DR method the bottom 30% of pixels are discarded to avoid reconstruction artifacts. Both the simulations and figure creation were done in Matlab 2019a [**75**], *https://www.mathworks.com/*.
a) This panel shows the ground truth attenuation coefficient for the simulation. This ground truth is a realization of the prior distribution given in equation 2.28.
b) This image shows the reconstructed attenuation coefficient using the DR method given in equation 2.14.
c) This panel shows an estimate attenuation coefficient given by the mean of the posterior distribution. This estimate was computed using equation 2.32.

random voxelwise intensity fluctuations are present in the signal due to speckle, and as a result, the voxelwise mean attenuation coefficient can not be exactly determined. Utilizing a statistical understanding of speckle fluctuations and prior physical knowledge, the posterior distribution for the attenuation coefficient was derived from first principles. This probability distribution better characterizes the voxelwise attenuation coefficient because it allows for the weighing of relative likelihoods and the quantification of uncertainty by measuring the variance of the attenuation posterior distribution.

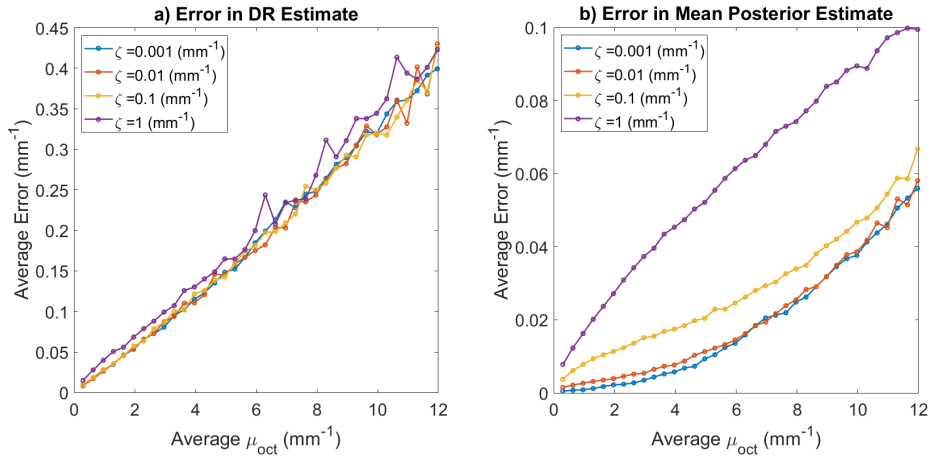


Figure 2.12. This figure demonstrates how problem parameters impact the error in the attenuation coefficient recovery. For each data point a single A-scan was simulated with 2000 depth measurements. After making the initial DR estimate, the deepest 30% of the pixels were discarded to avoid artifacts. The error (L2) estimates were made for each of the 1400 remaining pixels and averaged. Both the simulations and figure creation were done in Matlab 2019a **[75]**, *https://www.mathworks.com/*.
a) This image shows how the DR reconstruction error varies with different mean attenuation coefficients for a variety of ζ values. The ζ value does not appear to greatly impact DR Reconstruction fidelity.
b) This image shows how the mean posterior estimator error varies with different mean attenuation coefficients for a variety of ζ values. This estimate was computed using equation 2.32. Clearly, the incorporation of this value into the prior impacts our uncertainty in our Bayesian estimate.

While the statistical framework derived in this paper is general, the applicability is limited by the assumptions made for the underlying depth resolved reconstruction technique. The DR reconstruction technique, given in equation 2.14, requires that the absorption of light be negligible when compared the total amount of attenuated light. This assumption is restrictive in the materials and wavelengths of light the DR technique can be applied

to. However, for the materials and wavelengths used in most common biomedical applications of OCT this assumption is valid. Furthermore, when the probability distribution for the reconstructed coefficient in equation 2.23 was derived, it was assumed that the coefficient of variation of the denominator in equation 2.14 is sufficiently small such that the denominator can be treated as constant. This does seem to be valid in numerical simulations and experiments, however, it is not clear if this is generally true.

Additional physical assumptions are made during the derivation of the prior distribution for the attenuation coefficient given in section 2.5.1. The prior distribution allows for the use of physical knowledge about the attenuation coefficient to introduce bounds and bias the probabilities towards realistic values. The derivation given in section 2.5.1 was made assuming the measured object contained uniform idealized scattering particles with no absorption. While this assumption may not hold for most tissue systems, a normally distributed prior is still a safe choice due to the fact that superpositions of random fluctuations tend to look normally distributed. In real tissue, the parameter ζ in equation 24 is difficult to define, as the meaning of the effective scattering cross section is ambiguous. However, it is still reasonable to assume that the true attenuation coefficient is normally distributed around the mean. The variance of the prior must be provided or inferred by other methods. There are techniques to estimate this parameter from the data such as empirical Bayesian methods **[80]**, however, the implementation of these techniques can be nontrivial and a robust verification must be performed before the method could be used clinically. While this is outside of the scope of this chapter, the Bayesian model presented here serves an an initial step towards the goal of estimating these parameters more robustly in tissue, and elucidates the impact of speckle on the recovered coefficients.

The use of physically accurate statistical models for the attenuation coefficient has several potential advantages. The variance of the posterior distribution provides a way to quantify uncertainty in reconstructions. Furthermore, estimation bias from higher order moments of the posterior can be quantified as well. The likelihood ratio statistic [**20**] can be computed using the physically accurate likelihood function given in equation 2.23. This statistical test gives a practitioner a sense of how likely a parameter is to fall within a specified range. In situations where a practitioner may want to have a single number to understand the attenuation in a system, the mean of the posterior can be computed as demonstrated in figure 2.11. In figure 2.12 we measure the error in the estimates for both the DR and mean of posterior estimators as the scattering cross section and attenuation coefficient is varied and can see a clear and intuitive relationship between error and $\mu_o ct$.

Another potential application domain is in OCT image segmentation where attenuation analysis is used to correct for signal decay and as a contrast enhancement tool [**120, 18**]. As we have discussed in this manuscript, the resulting attenuation image can be very highly variable due to the speckle fluctuations in the original signal. If the attenuation image is to be segmented, these fluctuations may lead to segmentation inaccuracies which can be an issue in scattering media such as the cultural heritage applications discussed in subsection 2.2.3. Denoising algorithms could combine our exponential likelihood with a spatial priors, such as total variation, [**100**] which would increase the likelihood of the piecewise constant attenuation coefficients. This could be used to improve segmentation accuracy by removing speckle fluctuations from the attenuation image. This

approach may be applicable even in the case of absorbing media because image segmentation does not require extraction of accurate attenuation values, only sufficient contrast between layers.

Speckle forms as a result of wavefront perturbations larger than a quarter of the wavelength of the light. For visible and near infrared light this causes speckle to be present in all but the most specialized applications as most objects have a surface roughness and subsurface scattering that far exceed this threshold. In some arenas of coherent optics, despeckling using hardware considerations may be possible. Techniques such as wavefront shaping for deep tissue imaging or in advanced camera-in-the-loop calibration methods for holographic displays demonstrate this principle. However, speckle in OCT can never be completely eliminated because the contrast of the imaging technique comes from the total light scattered volumetrically within a voxel. Since the volume scattering is the source of the contrast and the source of the speckle, the speckle itself is the fundamental signal formation mechanism. Thus, understanding the behavior, statistics, and impact of the speckle is the best route forward.

This work is an initial theoretical step towards fully quantifying and characterizing uncertainties in voxelwise OCT attenuation coefficient recovery in order to better understand the resulting estimates. The likelihood function from equation 2.23 accurately models the voxelwise measurement uncertainty of the attenuation coefficient due to speckle. This likelihood function gives insight into the voxelwise statistics of the DR attenuation images. The posterior distribution for the mean value of the attenuation coefficient, given in equation 2.30, allows parameter estimation to be performed in a consistent and reliable manner by using the posterior mean estimator given in equation 2.32. Furthermore, the

posterior distribution derived in this paper can be used to quantify the variance in estimates, which gives insight into uncertainty. While this is a promising approach, further research is still needed to find the best way to apply these techniques to clinical practice.

CHAPTER 3

Computational Holography: Suppressing Speckle in Time Multiplexed 3D Displays

3.1. Introduction

Chapter 2 centers around the problem of parameter recovery in a coherent imaging system with speckle. Fundamentally, the underlying imaging problem is about sensing a 3D optical field generated from scattering. The imaging of these scattered fields is at the core of much modern science, however, the inverse of this problem, whether it is possible to generate an arbitrary light field, is often less considered. Precise control over an optical field has many applications, for instance, scattering media can often be imaged through by shaping the incident coherent light beam to account for the occluding media [**68**]. A more straightforward application of this principle would be in the creation of a coherent 3D display. Such a display would generate an image that would contain all the needed visual cues for an observer to perceive depth.

Depth is perceived by human beings in response to a wide number of cues. One of the primary depth cues comes from stereoscopic vision where each eye perceives the scene differently due to parallax effects. 3D images have been displayed using this cue since at least the 1800's when stereoscopic displays were first built by Charles Wheatstone [**101**]. In these, two images which are captured from slightly different perspectives are presented to each eye tricking the viewer into perceiving a natural 3D scene. When an observer looks around these images the eyes will rotate to center the same point in each image into the same part of the visual field meaning the eyes are rotated so that they converge to a point in the scene, the depth of which is called the vergence distance. However, the images shown to each eye are themselves 2 dimensional images meaning no matter where the observer is looking at the eyes themselves are focused to a single plane, the plane that the image is placed. This differs from a natural scene where there is an intrinsic

correspondence between the vergence distance and focus distance (called accommodation distance). In natural scenes, an additional depth cue comes from accommodation state which is in conflict in traditional stereoscopic setup. While the human visual system is far more complex than a standard camera, the eyes still focus light through a lens and have a natural depth of field. When an observer is focused on a single depth, objects far from this plane are out of focus and blurry. Furthermore, the farther an object is from your focus the more severe the blur becomes. This subtle cue can conveys depth information to the brain without the need to scan a scene. More high level depth cues come from object geometry and motion, cue which underlay many 3D effects scene in optical illusions and 3D effects in computer generated scenes. Over the past 10 years, head mounted 3D display technology has become increasingly commonly adopted for entertainment and professional use cases [**57, 77**]. While the current generation displays are promising, they still rely on stereoscopic depth cues only to generate the perceptual 3D experience. The lack of the more subtle 3D depth cues leads to nausea and eye strain [**62, 22**]. These effects prevent people from using current generation VR displays for an extended period of time and is considered a core problem in the field of head mounted displays.

The problem of vergence accommodation conflict and missing depth cues could be solved introduction of more complex display technology such as varifocal displays. A varifocal head mounted display refers to a display where the focus position of a screen changes by either the incorporation of a tuneable lens or a physical shift in the display in depth which better corresponds with the natural vergence depth [**52**]. However, head mounted displays must be high resolution, compact, and light enough to be worn comfortably for an extended period of time; parameters which are difficult to achieve with

these methods. Instead of a varifocal display, a volumetric or multifocal display can be used [**117**] which has several displays stacked up at different depths from the eye allowing for natural accommodation. However, these are bulky and have limited depth planes. Another approach is a near eye lightfield display [**66**] where different images are projected in each direction allowing for parallax and view dependent effects to be rendered. However, the low space bandwidth product of these displays yields a severe tradeoff between angular and spatial resolution making them impractical for high quality imagery.

One developing technology which can potentially address vergence accommodation conflict and include natural depth cues is near eye Holographic Displays [**74, 22, 87, 67, 31, 79, 63**]. A holographic display forms a 3D image using interference patterns generated by a programmable spatial light modulator (SLM) which can control the shape of an incoming beam of light. This requires a relatively small amount of optics helping to maintain a lightweight and compact form factor. Additionally, because the entire light field is being controlled, it is possible to display a true 3D image which contains all necessary focus cues.

Despite their advantages, holographic displays suffer from several physical drawbacks which must be addressed prior to them being a practical option for real headsets. One fundamental drawback of a holographic display is their limited *Etendue* [**63**]. The Etenedue of an display is the product between the maximum solid angle a display can generate and the entry pupil size. This quantity is a Lagrange invariant and is conserved for a given optical system. For a near eye Holographic display this means the the *field of view* (FOV) that the display can produce and the *eyebox* size, which is the area in space where the

eye can be located and see the image, are fundamentally at odds with one another. A display with a large immersive field of view will have a tiny eyebox, meaning that slight eye or head movement can cause an end user to completely lose the image being displayed. Techniques have been proposed to effectively boost the overall Etendue of holographic displays through the use of scattering masks [**63, 22**] which have a very large scattering angle due to their microscopic features sizes. However, this technique introduces a random perturbation to the wavefront causing speckle. Additionally, phase perturbations can be introduced due to model mismatch, dust, quantization, poor alignment, or a number of other factors making speckle an ever present issue in this display technology [**51, 85, 37, 25**]. These physical limitations may prevent a single holographic image from achieving the high SNR, large FOV and large eyebox needed for next generation applications.

It may be possible to circumvent these limitations through the use of temporal multiplexing approaches [**37, 51, 25**] which build up the total high quality image using a series of lower quality images displayed very rapidly. For instance, an approach to this problem would be to use a scanning pupil duplicating display architecture where the optical configuration is chosen such that the FOV is as large as possible and the eyebox is small and a grid of eyeboxes are sequentially projected and tiled. Additionally, by perturbing the optical field deliberately, speckle from multiple frames can be averaged out producing a higher SNR image. These approaches are promising but require an SLM with a considerably higher frame rate than traditional SLM technologies can offer.

A new type of SLM which relies on micromirrors instead of a traditional liquid crystal display has recently been developed by Texas Instruments [**15**]. This *DLP SLM* has an

extremely high refresh rate at around 10kHz, making temporal multiplexing for either image enhancement or Etendue expansion a possibility. While the refresh rate of these DLPs is exceedingly fast, they have a displacement bit depth of 4 bits, meaning they can only delay phase to 16 different levels. This offers very poor control over the shape of the wavefront severely degrading image quality, particularly if the original SLM pattern was optimized for assuming continuous phase delays. The discretization of the phase leads to the formation of speckle. Traditional methods to reduce speckle in temporally multiplexed displays involve averaging a very high number of frames together to improve quality. This is effective but uses a significant portion of the increased temporal bandwidth which could be used to increase Etendue making the display more useable.

The degradation due to quantization will depend on how much the phase delay is perturbed, which depends on the specific SLM pattern being displayed. Typically, to find an SLM pattern to display a given scene, an iterative technique will be used, often involving some form of gradient descent. The overall cost landscape for these problems is highly non-convex [**118**] allowing for many nearly equally optimal SLM patterns to be found. While, each subsequent pattern may be nearly equivalent in quality for the continuous case, there is no reason to believe that this is the case for the quantized patterns. Specifically, it may be useful to find optimization techniques which can take the quantization operation into account and find patterns which generate images with higher quality post quantization.

In the OCT sensing problem, Speckle was inherent to the volume scattering of coherent light so there is no practical way to remove the speckle from the reconstructed volumes. In fact, the speckle was the physical process by which the image is formed. However, in the

case of time-multiplexed low bit holography it may be possible to pre-select SLM phase delay patterns so that the speckle is reduced allowing for fewer frames to be averaged together. In this chapter we consider optimization approaches which explicitly take this discretization into account to reduce the required number of frames.

3.1.1. Chapter Structure

The purpose of this chapter is to propose and investigate a few techniques to incorporate SLM quantization into optimization in the context of Holographic image generation. To begin with, related work is reviewed in section 3.1.2. In 3.2.1 we derive a light propagation model which is useful for on axis holography. In section 3.2.2 a brief primer on holography and its application to display are explored.

The forward model for a general display is derived in section 3.3 and, in particular, quantization is considered in section 3.3.2. The statistical properties of the signal degradation are analyzed in section 3.4. Once this theory is understood we present a number of optimization approaches in section 3.5. Simulated results are considered and compared in section 3.6 and finally a discussion is done in 3.7.

3.1.2. Related work

Near eye displays have been a topic of interest since at least the 1960s when head mounted (but wall tethered) stereo 3D displays with motion tracking were developed by Sutherland [**103**]. Head mounted displays have become of particular interest since the explosion of cheap consumer grade VR headsets in the early-2010s. A typical modern head mounted display contains two LCD panels and the 3D effect is entirely generated by stereo 3D cues

as well as 3D from motion. This type of setup can cause vergence-accommodation conflict (VAC) where the eyes are focused (accommodation) on the display but tilted to look at a virtual object far in the distance (vergence) [**62**]. To address this, 3D near eye displays have been investigated [**12, 45**]. Light field displays, where different images are projected into different viewing directions have been investigated in a near eye configuration [**66**]. This solves VAC but the severe trade-off between spatial pixel resolution and angular resolution prevents high quality immersive images from being displayed. An alternative approach is to either use a varifocal lens with an adjustable focal length or mechanical actuation of the LCD to dynamically move the focal plane of the display to match the depth of the observed object [**35, 52**]. This effectively eliminates VAC without sacrificing any field of view or resolution. These varifocal style displays are promising but they require perfect eye tracking as well as additional moving parts that require precise calibration and add weight and power consumption. Instead of moving the display, it may be possible to use a volumetric technique to accommodate for VAC [**117**]. However these methods are bulky, and typically have limited depth planes and aren't robust to eye movement and rotation [**76**]. Near eye holographic displays have been of increasing interest in recent years as SLMs have become cheaper and algorithms have improved [**23, 113, 87, 86, 19, 25**]. These methods have the capability to create a true 3D light field which contains all real focus cues in an compact package [**74**]. However, they suffer from limited Entendue [**63**] and require coherent light which leads to speckle [**86**] and require precise calibration [**19**].

The impact of hologram quantization in computer generated holography (CGH) has been studied since at least the 1960s when the theory of Binary Fraunhofer Holograms was developed [**72**]. Specific optimization algorithms for binary CGH have been developed in

recent years which use non-convex optimization approaches to arrive at high quality images [**118**]. Many approaches to find good binary patterns for amplitude modulators to do holography have been developed including error diffusion [**78**], optimized random phases [**113**], and spiking NN inspired binary straight through estimation [**67**]. An analysis of the impact of bit depth in Fourier holography found that the image degradation scales with the size of the quantization error [**85**] and that the distribution of the degradation follows speckle statistics at leading order.

In digital holography, the problem of speckle is often addressed through the improvement of models [**87, 86, 19**] using learned holography approaches. In these approaches, the traditional forward modeling is supplemented, or in some cases fully replaced, with a learned forward model which can account for aberrations in the optical path and other sources of model mismatch. This is done with a gradient based optimization scheme where part of the gradient is computed with the camera image itself. This allows the model parameters to be updated properly for the real system despite not having a perfect forward model. Speckle can also be reduced by requiring smoothness on the output phase preventing high frequency errors and phase wrapping from entering the problem at the SLM [**97, 87**], however, this can reduce the ability of the hologram to represent 3D focal stacks.

Speckle is only a major issue when using light sources with a very long coherence length such as a laser. Some work has been done to optically reduce speckles by using light sources with a lower coherence length such as an sLED [**31, 79, 86**]. In these techniques, an SLM pattern which produces a reasonable image for a wide number of wavelengths is found using an optimization method. Due to the fact that this single

pattern must work at many wavelengths there is an overall reduction of image sharpness, however, this also yields a massive reduction in noise [**31**]. This approach has been used in a prototype headset to make a full color and speckle free head-mounted holographic display [**79**]. Partial coherence has been fused with newer, neural holography approaches to create very low noise images which are significantly higher quality than in traditional partial coherence holography [**86**]. This work combined wavelength diversity with their camera in the loop calibration to improve resolution and quality. The impressive results come at a high computational cost and still have lower resolution and contrast compared with traditional CGH.

Temporal multiplexing has been applied to holography to reduce speckle for decades [**51, 85, 37, 25**]. In these approaches, many frames are averaged together with some perturbation which yields independent speckle patterns. This has been directly applied to the speckle generated from quantization noise [**85**]. Recently, an approach was proposed to co-optimize frames using an adapted Gerchberg-Saxton phase retrieval method and reduce speckle via averaging these frames together [**37**]. A similar technique has been developed for binary bit depth CGH optimization. In this paper a proxy gradient method has been used to replace the non-differentiable quantization operator derivative. Concurrently, a similar multi-frame optimization method was also developed which incorporated camera in the loop for even sharper results [**25**]. In this paper, a proxy gradient method was proposed which used annealed Gumball-softmax functions to approximate the derivatives. This was used in addition to camera in the loop training and showed excellent results.

3.2. Background

3.2.1. Angular Spectrum Method

Computational holography is concerned with recreating a given interference pattern by modulating the phase of a given input field. Before this can be addressed, the way that light diffracts and propagates through space must be modeled. To do this, a numerical method called the Angular spectrum method is often employed [**59**]. This method considers an interesting correspondence between the frequency content of an on-axis incident light field over some aperture and an expansion in a series of incident plane waves which have different incident angles relative to the aperture known as the the angular spectrum of the light.

Specifically, let $E(x, y, 0)$ be the incident monochromatic electric field with frequency k over some aperture function. Then the angular spectrum is defined as

$$\mathcal{A}_0(k_x, k_y) = F[E(x, y, 0)]. \tag{3.1}$$

A single plane wave in the expansion can be propagated a distance of z from the aperture plane by simply applying a phase delay of the form $\exp[izk_z]$ where k_z is the propagation direction given by $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. Integrating over all possible directions will yield the propagated field at distance z and the propagation formula

$$E(x, y, z) = F^{-1}[\mathcal{A}_0(k_x, k_y)exp(izk_z)]. \tag{3.2}$$

We assume that the incident field a very small angular spectrum such that the first order Taylor expansion $k_z \approx k - \frac{k_x^2+k_y^2}{2k}$ holds. Plugging this in we see that

$$E(x,y,z) = \exp(ikz)F^{-1}[F[E_0]\exp(iz\frac{-k_x^2-k_y^2}{2k})] \tag{3.3}$$

where F denotes the Fourier transform. Turning our perspective over to computation we denote the FFT operator as $\boldsymbol{F}$ and the propagation operator as

$$\boldsymbol{A}(z) = \exp\left(iz\left(K^2-k_x^2-k_y^2\right)\right) circ(\frac{k_x^2+k_y^2}{\left(K^2-k_x^2-k_y^2\right)}). \tag{3.4}$$

From this we can define the ASM operator as

$$\mathrm{ASM}(E_0,z) = \boldsymbol{F}^{-1}\boldsymbol{A}(z)\boldsymbol{F}E_0. \tag{3.5}$$

A more detailed derivation can be found in appendix B.1.

3.2.2. Holography

Holography is a long standing technique in which all of the information of a wavefront is encoded onto a film or physical media as an interference pattern. Due to the properties of this interference pattern it is possible to illuminate the recorded film again and 'play' the original light field back. This technique was originally invented by Denis Gabor in the late 1940s [**44**] as a new modality by which to do high resolution electron microscopy. What makes holography so unique is the ability to encode 3D information completely in a 2D slice of the field, allowing for the reconstruction of a 3D light field from a 2D measurement. In this subsection the basic ideas of holography are outlined.

3.2.2.1. Intensity Holograms. Consider an 'object' electric field denoted $O(x, y, z)$ and a reference field which is a plane wave denoted by $R(x, y, z) = \exp(-i\boldsymbol{k}_r \cdot \boldsymbol{x})$. The goal of holography is to encode the object field in a medium as an interference pattern and then reconstruct or *playback* the field later. The interference pattern on plane $z = 0$ is given by

$$H_I(x, y) = ||O + R||^2 = ||O||^2 + ||R||^2 + O^*R + R^*O \tag{3.6}$$

where $O = O(x, y, 0)$ and $R = R(x, y, 0)$. Assuming that this reference pattern in encoded in the transmission of a media the media can be then be illuminated with a second plane wave R' yielding the electric field at a distance z from the aperture

$$ASM(R'H_I, z) = (RR')O(x, y, z) + (R^*R')O^*(x, y, z) + DC \tag{3.7}$$

where $R'RO = exp(i(k-k_r)\cdot x)O$ and $R'R = exp(i(k+k_r)\cdot x)O^*$ which are two twin scaled copies of the original object field evaluated at the plane of the interference pattern. This means that the 3D information from the field was successfully encoded in the transmission hologram H_I. A more detailed derivation of this can be seen in appendix B.2.

3.2.2.2. Phase Holograms. For many types of recording media the hologram isn't recorded in a transmission but instead in an index of refraction variation. In this case we would expect the phase hologram to have the phase patterns

$$H = \exp(i\gamma H_I) \sim 1 + i\gamma H_I - \frac{\gamma^2 H_I^2}{2} + ... \tag{3.8}$$

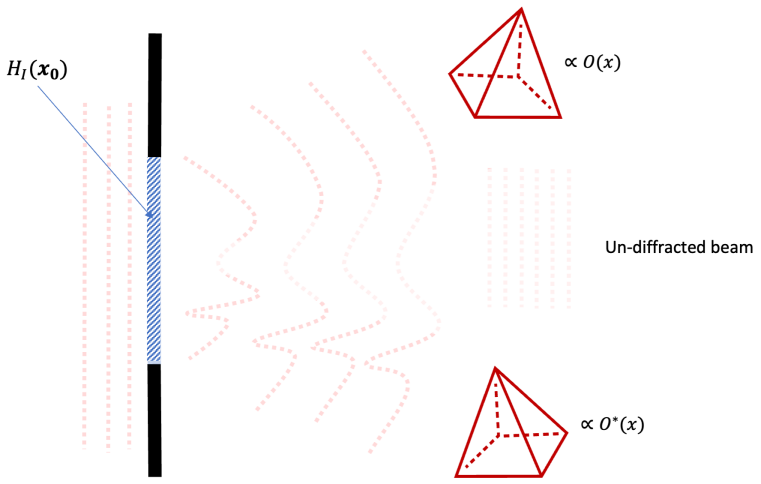


Figure 3.1. Diagram demonstrating the three leading order terms which arise from phase only holography. One term is simply the undiffracted light passing through he aperture. The twin image terms are identical conjugate copies of the object. The coordinates for the conjugate copy of the object field are transposed resulting ins a flipping of the image.

where H_I is defined in equation 3.6. At leading order this is an additional DC term plus the intensity hologram which we showed in the previous section reproduces our object field but you also get a series of higher order diffraction terms. The weaker the index of refraction change in the encoding the less bothersome the higher order diffraction terms will be, but the less contrast you will have to overcome the addition D.C. term.

3.2.2.3. Digital Holography and Holographic Displays. Conventional Holography was a revolution in imaging eventually resulting in Gabor winning the Nobel prize in 1971 [**90**]. However, with modern computational methods having a physical object field for recording may not be strictly necessary. Suppose you want to create an object field

O given some reference beam R'. Using mathematical modeling it should be possible to predict what the interference pattern is that will produce this.

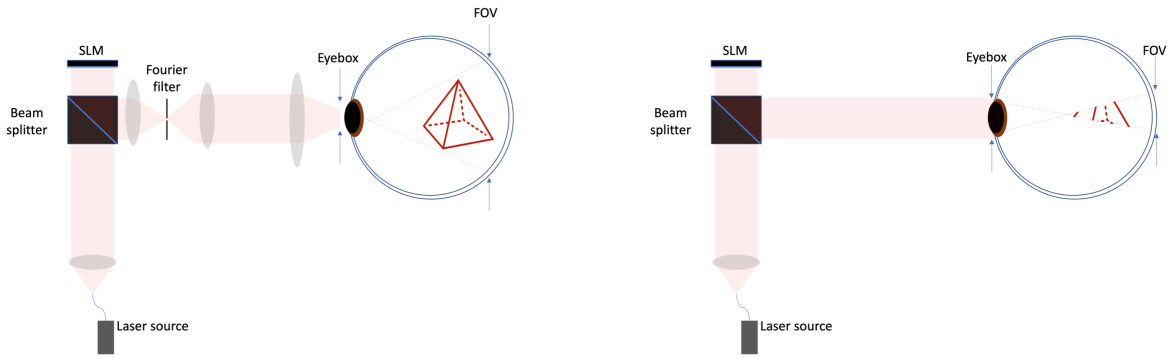


Figure 3.2. Figure demonstrating two common holographic display architectures. On the left a Fourier filter denoises the image and the Etendue trade off is controlled with an additional lens at the end of the optical path. On the right, a simple lensless architecture is shown. In this case, the field of view is governed only by the maximum scattering angle given by the SLM pixel size.

To actually apply the phase delay, the field is controlled with an optical component called a *Phase only Spatial Light Modulator* (SLM). Amplitude SLMs exist as well but we refer in this chapter consider the phase only variety. These optical components traditionally use a liquid crystal on silicon (LCOS) display which has pixels that have a refractive index which can be modulated by applying a voltage. A spatially varying phase delay applied to an incident field by reflecting the field off of the modulator. Newer architectures use micromirrors which physically display by a fraction of a wavelength to accomplish

the phase shift. These SLMs open up the ability to have a fully programmable and user controlled holographic display. This would solve many issues with 3D display technology such as VAC by creating a true 3D light field which is indistinguishable from a real scene.

3.3. Mathematical Modeling

3.3.1. Optical Path Modeling

To optimize for an SLM pattern in CGH we first need a parameterized physical model for a simple optical setup which could be realistically implemented in a display. In this section, it is assumed there is a complex incident field, denoted as E_0, which is a normalized uniform in phase and has intensity with $|E_0|^2 = 1$ everywhere. This incident field is then modulated by the phase-only SLM pointwise as an applied phase delay denoted by ϕ and the peak amplitude of the laser is α. We assume that there is some percentage of undiffracted light due to the incomplete fill factor of the SLM given by p_{fill}. This is modeled mathematically by

$$\text{SLM}[E_0(x), p_{fill}; \phi(x), \alpha] = (1 - p_{fill})\alpha E_0(x) + p_{fill} exp(-j\phi(x))\alpha E_0(x). \tag{3.9}$$

The propagation model is given by a band limited angular spectrum method discussed in section 3.2.1 where

$$\text{ASM}[E, z] = (\boldsymbol{F}^{-1}\boldsymbol{A}(z)\boldsymbol{F})E. \tag{3.10}$$

It is assumed that the detector measures intensity, meaning the whole image formation model for a single SLM pattern is given by

$$I_{single}(E_0, z, p_{fill}; \phi, \alpha) = \left|\text{ASM}[\text{SLM}[E_0(x), p_{fill}; \phi(x), \alpha]\right|^2. \tag{3.11}$$

In the temporally multiplexed case the intensity should be added together yielding a Temporally multiplexed image formation model

$$I(E_0, z, p_{fill}; \phi, \alpha) = \sum_{j=1}^{n_f} \left|\text{ASM}[\text{SLM}[E_0(x), p_{fill}; \phi_j(x), \alpha]\right|^2 \tag{3.12}$$

where n_f is the total number of multiplexed frames.

3.3.2. Quantization

Quantization is the process of taking a continuous signal and 'binning' it into a series of discrete values. Some version of quantization is present in almost every area of computational science such as machine learning [**46**], numerical differential equations [**26**], and digital image processing [**105**]. A major difference between traditional holography and digital holography using SLMs is signal quantization. Spatial light modulators are real devices which are controlled by real circuits and cannot control phase delays arbitrarily. Instead, a SLM can apply a specific set of phase delays with finite accuracy. The potential delays can broadly be separated into uniform and non-uniform quantization categories. Typically, the number of representable phase levels will be of the form 2^{n_b} where n_b is referred to as the *bit depth* of the SLM. For the case of evenly spaced phase levels we can model the quantization operator as

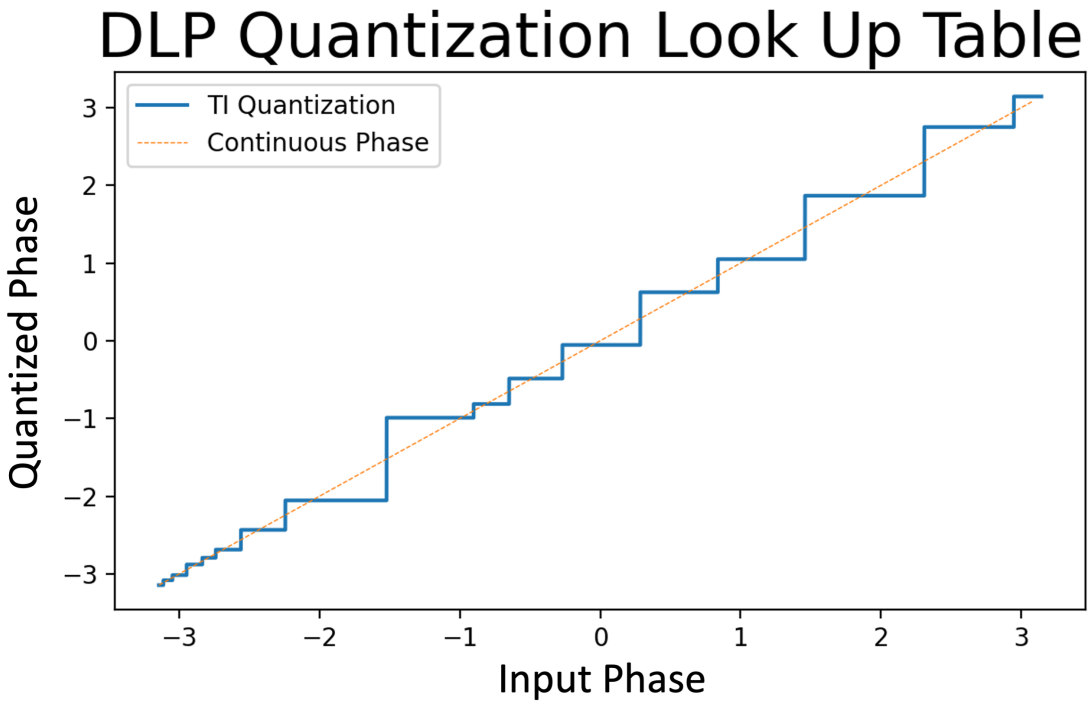


Figure 3.3. This figure shows the quantization look up table for the TI DLP slm. This quantization is uneven with large jumps between certain phase levels and regions of phase which are very well represented buy the SLM. The image degradation is bounded by the maximum phase perturbation, so an unevenly spaced quantiazation will produce a worse image for an arbitrary phase pattern.

$$\text{Quant}(\phi; n_b) = \frac{2\pi}{2^{n_b}-1}\text{round}\big((2^{n_b}-1)\frac{\phi+\pi}{2\pi}\big) - \pi \tag{3.13}$$

where n_b is the number of bits of the SLM and assuming that ϕ is wrapped such that $-\pi \leq \phi \leq \pi$. In this case the phase levels here have an spacing of $\frac{2\pi}{2^{n_b}}$. The quantization for many SLMs is not perfectly uniform and adjusting the precise phase levels [**43**] which are applied for each given signal can represent a core part of a holographic display calibration. In this case, we cannot write down a simple and cheap quantization operator explicitly,

and instead we refer to

$$\text{Quant}(\phi; \text{LUT}) = \underset{\phi^* \in \text{LUT}}{\text{argmin}} \, |\phi - \phi^*| \tag{3.14}$$

In the case of a very low bit depth SLM a search can be performed pixelwise quickly enough to find the optimal quantized encoding.

In either case the quantization operator contains operations which are difficult to differentiate as the derivative is zero almost everywhere except at the phase jumps. Figure 3.3 shows the uneven quantization used in the TI DLP SLM which we will be using to test the optimziation methods later in this chapter. When the image is formed under SLM quantization we can use the model

$$I_Q(E_0, z, p_{fill}; \phi, \alpha) = \sum_{j=1}^{n_f} \left| \text{ASM}[\text{SLM}[E_0(x), p_{fill}; \text{Quant}(\phi_j(x)), \alpha] \right|^2. \tag{3.15}$$

3.4. An Analysis of Low Bit Depth Speckle

As discussed in section 2.4, speckle is a phenomenon in which small, unintended phase perturbations in a coherent wavefront causes high frequency and undesirable intensity fluctuations. In OCT these were introduced by volume scatters having random position within a voxel. However in the case of Holography, similar phase variations can be introduced from poor alignment, model mismatching, finite aperture sizes, or the use of diffuse scatterers. While these sources of speckle are pernicious, they can be addressed with calibration and temporal multiplexing. However, when using a fast SLM such as a DLP, the quantization from the low bit depth causes a large amount speckle to be present

in the system. In this section a brief analysis to understand the statistics of the speckles introduced through signal quantization. This can be seen clearly in figure 3.4.

The magnitude of the phase perturbation will depend on how close the desired phase level is to the nearest quantized value, but can be always bounded by

$$\Delta\phi = |\phi_{ij} - \text{Quant}(\phi_{ij})| < \frac{1}{2}\text{sup}_i(P_i - P_{i+1}) \tag{3.16}$$

where $\{P_i\}$ is the set of all quantized phase levels. Most conventional LCOS SLMs have a bit depth between 8 to 10 meaning they can display 256 - 1024 different phase levels which are nearly uniformly spaced.

In practice, signal degradation could be measured using MSE where the desired or target intensity is given by T_{ij} then the measured MSE is given by

$$MSE = \big|\, |T - I_Q(\phi)| \,\big|^2 \tag{3.17}$$

MSE is a practical and useful way measure error however for the purposes of theory it is more useful to use signal variance given by

$$Var[I] = E[E[I] - \tilde{I}]^2. \tag{3.18}$$

Speckle causes an intensity in the signal which is around the unperturbed mean meaning $E[I_{ij}] = T_{ij}$ as discussed in chapter 2. There is a subtle conceptual difference between the signal variance and MSE in that the variance is discussing the expected variation in an ensemble of measurements while MSE discussing the mismatch in a given realization of the signal undergoing speckle. In practice the signal variance can be thought of as roughly

equivalent to an expected MSE because they are on average they same and display systems have many many pixels making individual realizations have an average like quality.

3.4.1. Quantization Speckle Statistics (How Much Signal Variance will Quantization Introduce)

Let the unquantized phase on the SLM be denoted by H. Heuristically, for a general hologram the histogram of phase delays is well approximated as a uniform distribution. In this case, we can model the quantization error for a bit depth of l in a statistical sense by assuming [**85**] that

$$(\text{Quant}(H) - H)_{ij} \sim \text{Unif}\left(-\frac{\Delta\phi}{2}, \frac{\Delta\phi}{2}\right). \tag{3.19}$$

where $(\text{Quant}(H) - H)_{ij}$ is the quantification error for pixel ij and all pixels are iid. This is sensible because the different between the actual phase level and the quantized phase level shouldn't a priori favor any values. By modeling the quantization error in this way it is possible to get statistical estimates for the variance of the image formed at the observation plane. The standard deviation of the error in a Fourier Holography digital setup is given by $\sigma_q = \frac{\Delta\phi}{\sqrt{12}}$. We can generalize this to a lensless ASM based model. We note that the ASM operator is unitary such that

$$ASM^{-1} = \left(\boldsymbol{F}^{-1}\boldsymbol{A}(z)\boldsymbol{F}\right)^{*} \tag{3.20}$$

Unitary operators transform covariance as

$$\text{Cov}\left[\text{ASM}(\text{quant}(H) - H)\right] = \left(F^{-1}AF\right)^{*} \text{Cov}\left[\text{quant}(H) - H\right] \left(F^{-1}AF\right). \tag{3.21}$$

However, since we assumed $(\text{Quant}(H) - H)_{ij}$ are pairwise independent we know the covariance must be $Cov[\text{quant}(H) - H] = \sigma_q^2 I$. This means that

$$\text{Cov}\left[\text{ASM}(\text{quant}(H) - H)\right] = \left(F^{-1}AF\right)^* \text{Cov}\left[\text{quant}(H) - H\right]\left(F^{-1}AF\right) = \sigma_q^2 I \tag{3.22}$$

allowing us to understand our field variance based on our introduced phase perturbation.

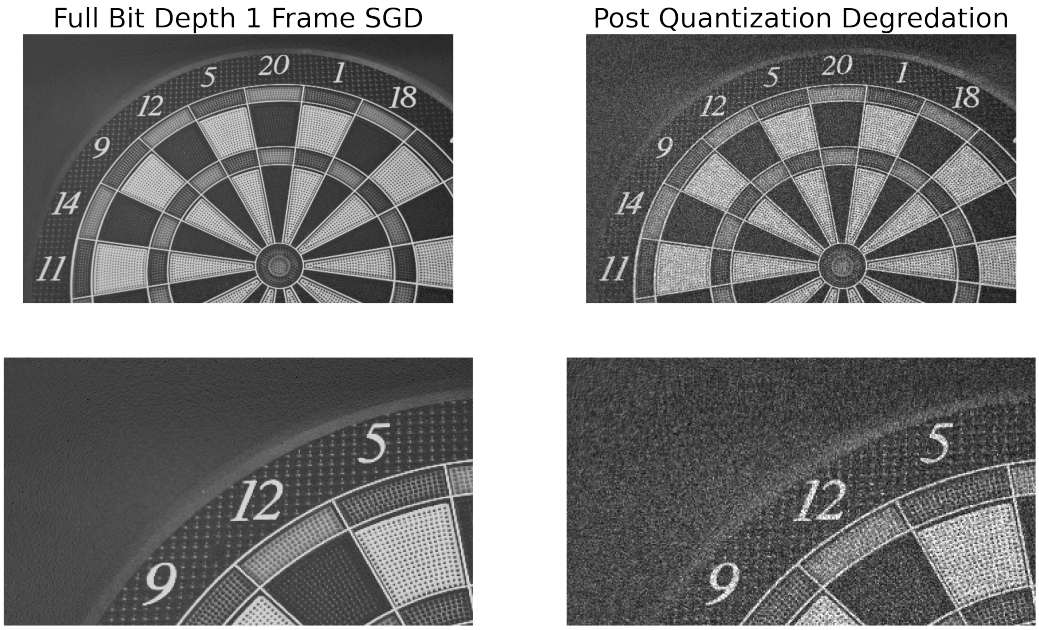


Figure 3.4. Figure demonstrating the impact of 4 bit quantization using the TI LUT. The predicted full bit depth image is in the right most column and shows an image can be created using no temporal multiplexing with high quality by finding an SLM pattern with stochastic gradient descent. The right column shows the resulting image after the weights are quantized simulating displaying the image on a real SLM. The overall image quality plummets and speckle becomes very noticeable.

3.5. Approaches to Optimization

Given an desired complex electric field at the detector plane, the electric field at the SLM plane can be computed simply by propagation. However, because an SLM can only modulate the phase the of the incident electric field, it is in general not possible to produce the exact electric field to form the desired image. Furthermore, since detectors only measure intensity, the SLM phase pattern must be solved for with measurements which disregard phase. In line with most phase retrieval approaches, an iterative technique, typically some sort of regression is used to find an SLM phase delay pattern which creates as close of an image as possible. However, the regression problem is highly nonconvex with many local optima which do not all produce equivalently high quality images after quantization.

To find good SLM patterns using the mathematical model outlined in section 3.3 we must find good optimization techniques which can take this quantization into account. However, the quantization operator halts most gradient based optimization schemes because the derivative is zero almost everywhere. In this section, several gradient based optimization methods which incorporate quantization are compared. Initially, we investigate ignoring the quantization operator and solving for the phase in a continuous quantity. This is standard practice for high bit depth SLMs with very little quantization error. Two variable splitting approaches for temporal multiplexing are then considered, one where each frame is solved for continuously while the other frames are quantized and one where pixels are solved for iteratively allowing for quantization to be used on the remaining pixels. We consider a soft categorical reparametization search method called a Gumbel Softmax optimization. Finally, We consider two proxy gradient methods, straight through

estimation where the quantization is used for a portion of the gradient only and a second method which uses the gradient from the Gumbel method fused with the function evaluations of the straight through estimator.

3.5.1. Continuous Stochastic Gradient Descent

The simplest approach to the problem is to optimize for a non-quantized pattern with stochastic gradient descent and then rounding. The patterns are found assuming the introduced phase delays are continuous variables. The computational graph is shown in figure 3.5 and loss function is shown in equation 3.24. Once the optimization has stalled the derived phase patterns are then wrapped and quantized to the look up table for the TI DLP SLM shown in figure 3.3 and table in appendix B.3 yielding

$$\text{Loss}(\phi^*) = ||I(\phi^*) - \text{Target}||^2 \tag{3.23}$$

$$\phi = \text{Quant}(\text{argmin}(\text{Loss}(\phi^*))). \tag{3.24}$$

3.5.2. Framewise Predictor Corrector

A simple approach to incorporating quantization into a temporally multiplexed SGD approach would be to only optimize for one frame at a time. In this case, the frame which is currently being optimized for would not be quantized allowing for SGD to update the phase pattern. All other frames would be quantized allowing the current frame being updated to correct the speckle from the other frames. This splitting method which we call Framewise predictor corrector (FPC) allows for the incorporation of quantization

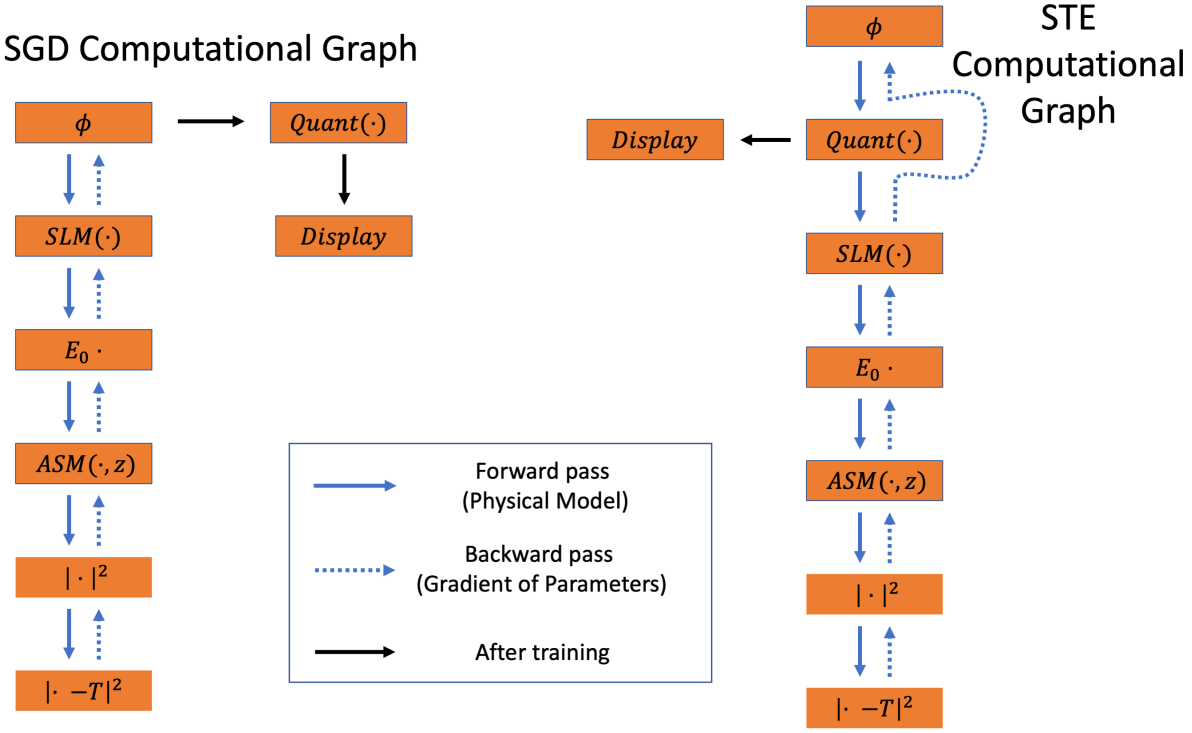


Figure 3.5. These figures show the SGD and STE computational graphs. The graphs are the same except the STE graph replaces the derivative of the quantization node with a 1 effectively skipping that node in the backpropagation. The psuedocode for these algorithms are written in section C.2.2 and C.2.3.

information into the problem while not causing the optimization to halt. To do this the optimization would have n_f loss functions of the form

$$\text{Loss}_i(\phi^*) = || \sum_{j<n_f, j\neq i} \left| \text{ASM}[\text{SLM}[E_0(x), p_{fill}; \text{Quant}(\phi_i(x)), \alpha] \right|^2 - \text{Target}||^2. \tag{3.25}$$

which will be iterated through in sequence. The computational graph is shown in figure 3.6. While conceptually, splitting the SLM patterns into groups framewise may seem obvious it has the drawback of only allowing quantized and non-quantized variables to interact incoherently.

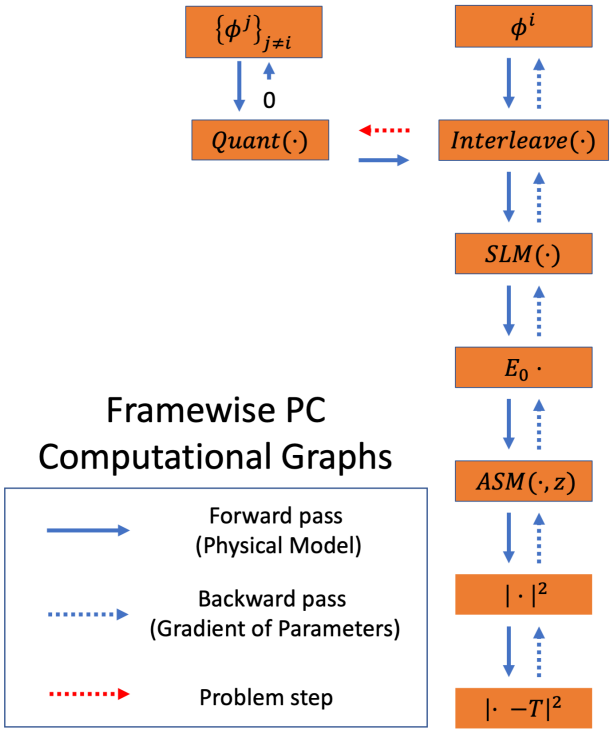


Figure 3.6. The Framewise Predictor Corrector computational graph. Each frame is quantized except for the frame i which is unquantized and optimized for. The frames are then interleaved into a single tensor and each frames phase delay and output is computed and summed. The red arrow shows where the gradient is forced to zero causing the SGD for those variables to halt. The psuedocode for this algorithm is written in section C.2.4.

3.5.3. Single Frame Predictor Corrector

The prospect of splitting the variables into groups which can impact each other coherently may improve results of a splitting method by allowing for pixels with too much intensity to be reduced. For this reason we propose dividing the SLM pixels into odd and even rows, quantizing one at a time and stepping the other in an alternating fashion. Because there are some pixels which are both quantized and unquantized within a single frame coherent effects are allowed.

To accomplish this approach denoted as single frame predictor corrector (SFPC), let $\phi^e = \{\phi_{ij} \quad \text{s.t } i \text{ is even}\}$ and $\phi^o = \{\phi_{ij} \quad \text{s.t } i \text{ is odd}\}$ and note that $\phi = \phi^e \cup \phi^o$. The total predicted field for the even and odd steps can be computed as

$$E_{Q,e,j}(E_0, z, p_{fill}; \phi_j, \alpha) = \text{ASM}[\text{SLM}[E_0; \phi^e \cup \text{Quant}(\phi_j^o), \alpha] \tag{3.26}$$

$$E_{Q,o,j}(E_0, z, p_{fill}; \phi_j, \alpha) = \text{ASM}[\text{SLM}[E_0; \text{Quant}(\phi_j^e) \cup \phi^o, \alpha] \tag{3.27}$$

Two loss function of the form

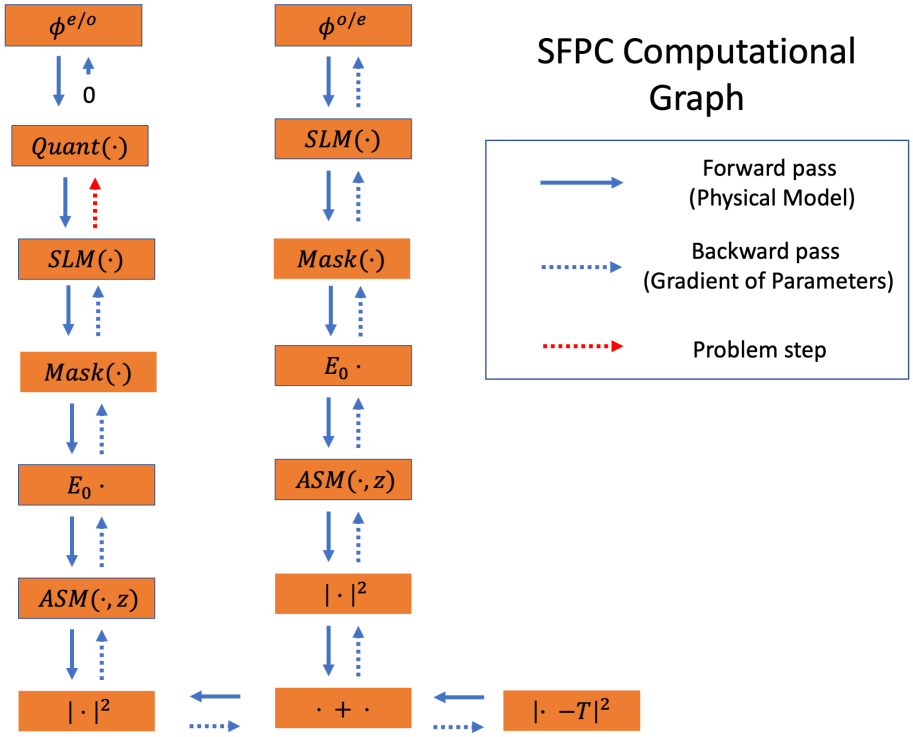


Figure 3.7. The Single Frame Predictor Corrector computational graph. This splits the variables into two groups separated at odd and even rows. Quantization is applied to only one group and the alternating group is updated. The red arrow shows where the gradient is forced to zero causing the SGD for those variables to halt. The psuedocode for this algorithm is written in section C.2.5.

$$\text{Loss}_o(\phi^*) = \left|\left|\,||E_{Q,e,j}(\text{Quant}(\phi_j)) + E_{Q,o,j}((\phi_j))||^2 - \text{Target}\right|\right|^2 \tag{3.28}$$

$$\text{Loss}_e(\phi^*) = \left|\left|\,||E_{Q,o,j}(\text{Quant}(\phi_j)) + E_{Q,e,j}((\phi_j))||^2 - \text{Target}\right|\right|^2 \tag{3.29}$$

can then be partially optimized using SGD iterations and iterated between until convergence. Like in the method discussed in section 3.5.2, this can be implemented as the two versions of the computational graph shown in Fig. 3.7 and alternated between using SGD. Formally, because quantization is differentiable but 0 almost everywhere the SGD can be done for all of ϕ and only the even and odd rows will update respectively. This has some drawbacks with respect to memory and computational however.

3.5.4. Straight Through Refinement

Another field where quantization is frequently encountered is in neural networks, particularly those which must run quickly and efficiently on hardware with memory and processing limitations. Adapting pre-trained networks to this environment is an active and interesting field of research [**46**]. For a neural network, typically training will be done on faster hardware and then the weights of the network will be quantized and copied to the lower power device to run–severely reducing network performance. To ameliorate this, a standard technique called 'quantization aware training' using a straight through estimator. This technique begins by using stochastic gradient descent to train the weight which are then perturbed via quantization. After the quantization step the weights are readjusted using a new optimization process which uses a modified computational graph. This graph uses the non-differentiable quatization operator for the forward pass allowing quantization information to enter the problem. Then, when the gradients are being

backpropagated the backwards graph skips the quantization operator allowing it to be unmodified. This seemingly arbitrary refinement approach has proven to be incredibly effective in reducing the impact of quantization on lightweight networks [**46**].

Modifying this approach for our heavily quantized SLM we can split the optimization into two stages. In the first stage we compute a solution to stochastic gradient descent with full bit depth, and then quantize after convergence as in section 3.5.1. Following this, a refinement process begins using the computational graph shown in figure 3.5. This allows quantization information to be taken into account in our optimization process. Essentially, a quantization operation is written that is custom and the derivative is assumed to be 1 at all times. This allows the quantization node to be effectively 'skipped' during the backpropagation. More mathematically, consider the loss function

$$\text{Loss}_{STE}(\phi^*) = \text{Loss}(\text{Quant}(\phi^*)) = ||I(\text{Quant}(\phi^*)) - \text{Target}||^2$$

we have from the chain rule

$$\frac{\partial \text{Loss}_{STE}(\phi)}{\partial \phi} = \frac{\partial \text{Loss}(\text{quant})}{\partial \text{quant}} \frac{\partial \text{quant}(\phi)}{\partial \phi} := D. \tag{3.30}$$

However since $\frac{\partial \text{quant}(\phi)}{\partial \phi} = 0$ gradient based methods halt. The straight through estimator does SGD using an approximate derivative where $\frac{\text{quant}(\phi)}{\partial \phi}$ is replaced with unity

$$\hat{D} = 1 \frac{\partial L(\text{quant}(\phi))}{\partial \phi} = \frac{\partial L(x)}{\partial x} \tag{3.31}$$

and SGD is performed with the proxy gradient $\hat{D}$ in place of D.

3.5.5. Gumbel-Softmax

Optimizing for Heavily quantized SLM patterns can be though of as a categorization problem instead of a simple regression problem. Each SLM pixel must be assigned to a discrete phase level (category) such that the ensemble of pixels generates the highest quality image possible. These problems are difficult because they must sample a discrete categorical distribution searching for an optimum, an operation which is highly non-differentiable. However, it is often possible to rewrite a stochastic variable with a given mean and variance as a variable which has a parameterized deterministic component (which can be optimized for using a gradient based method) and added noise of fixed statistics. A important result from the theory of rare events [**49, 58**], allows the sampling of a categorical distribution whose probabilities are proportional to some energy functional $\propto \exp(-log(p_i))$ as

$$z = \text{one\_hot}\left[\underset{i \leq n}{\text{argmax}}\left(g_i + log(p_i)\right)\right] \tag{3.32}$$

where g_i $Gumbel(0, 1)$ is a Gumbel random variable with fixed mean and variance for any probabilities. The output z is an n dimensional binary variable with only one non-zero value. The quantity p_i is the un-normalized probability for class i.

In the case of SLM optimization, p_i should be a function which is largest for quantization states similar to the current SLM pattern being displayed. To do this, define the

score function

$$p_i(\phi; w, q_i) = \text{Sigmoid}(w(\phi - q_i))\,(1 - \text{Sigmoid}(w(\phi - q_i))) \tag{3.33}$$

where q_i is the i_{th} quantization level. The one hot encoding in Eq. 3.32 is non-differentiable, however this function can be replaced with the softmax operator

$$softmax_i(x; t) = \frac{\exp(-x_i/t)}{\sum_j \exp(-x_j/t)}. \tag{3.34}$$

This function normalizes and assigns a probability to each category that is near one for the largest values and near zero for the smaller values. As the temperature parameter $t \to 0^+$ the softmax becomes $softmax(x; t) \to \text{one\_hot}\left[\underset{i}{\text{argmax}}(x_i)\right]$ and as $t \to \infty$ the $softmax(x; t) \to 1/n_l$ where n_l is the number of phase levels. The softmax effectively returns a probability distribution that each pixel should be in each quantization state given it's current estimated phase delay value that can be annealed to allow for more or less variation from the current state. We substitute this into equation 3.32 to arrive at the 'soft' category sampling function

$$z(p; t) = \frac{\exp(-\left(g_i + log(p_i)\right)/t)}{\sum_j \exp(-\left(g_j + log(p_j)\right)/t)}. \tag{3.35}$$

The variable z is a probability distribution and not a phase delay. To generate phase patterns we simply compute a probability weighted average of the different quantization states, denoted q_l, as

$$\text{Gum\_quant}(\phi; t) = \sum_l q_l z(p; t)_l \tag{3.36}$$

This function is effectively a differentiable quantization function which introduces stochasticity into the problem to prevent getting stuck in local optima. The overall computation graph for this method is shown in Fig. 3.8. This stochastic differentiable quantization node can be incorporated into a loss function as

$$\text{Loss}(\phi^*) = ||I(\text{Gum\_quant}(\phi^*))\text{Target}||^2 \tag{3.37}$$

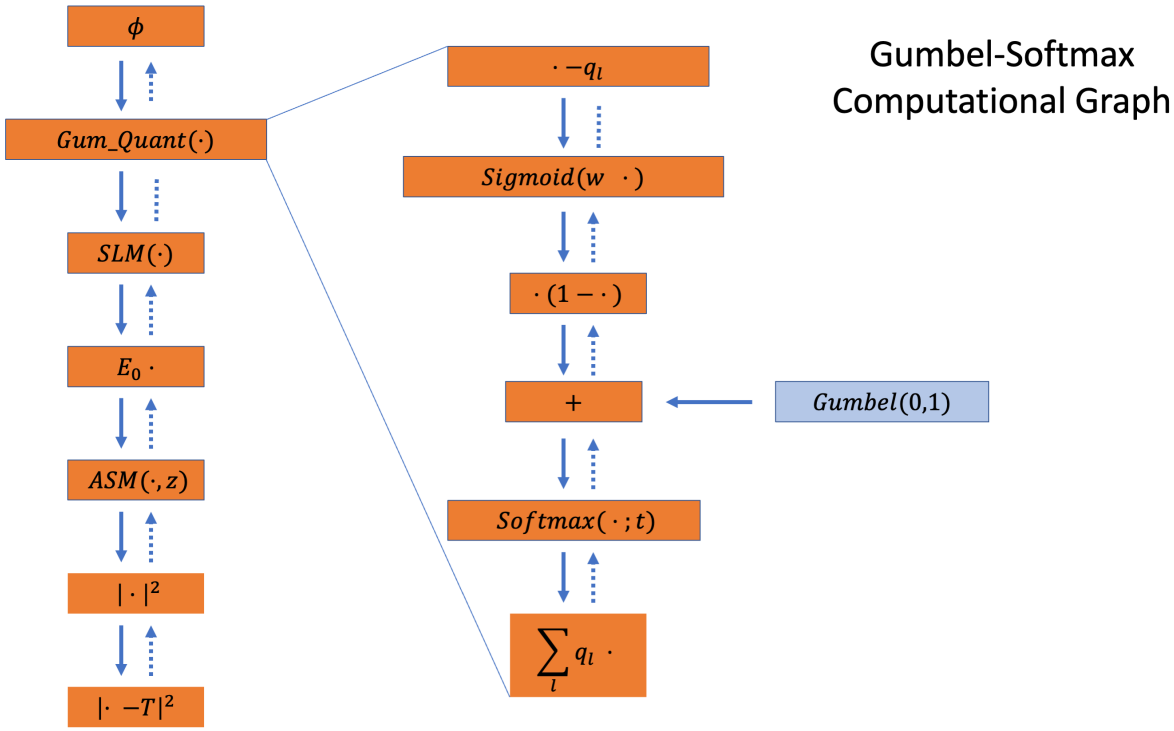


Figure 3.8. The computational graph for the Gumbel Softmax optimization method. The Gumbel soft-quant layer is a stochastic reparaterization of a categorical sampling process where the category probabilities depend on the similarity to the current SLM pattern. The psuedocode for this algorithm is written in section C.2.6.

3.5.6. Proxy Gumbel

Using the same conceptual framework as the straight through estimator it is possible to introduce a more complex proxy gradient where instead of skipping the backwards node an alternative function is differentiated. An interesting implementation of this is in [**25**] where the forward computational graph allows for true computation but the backwards pass was computed using the backwards computational graph from a Gumbel-Softmax method. This is shown in figure 3.9

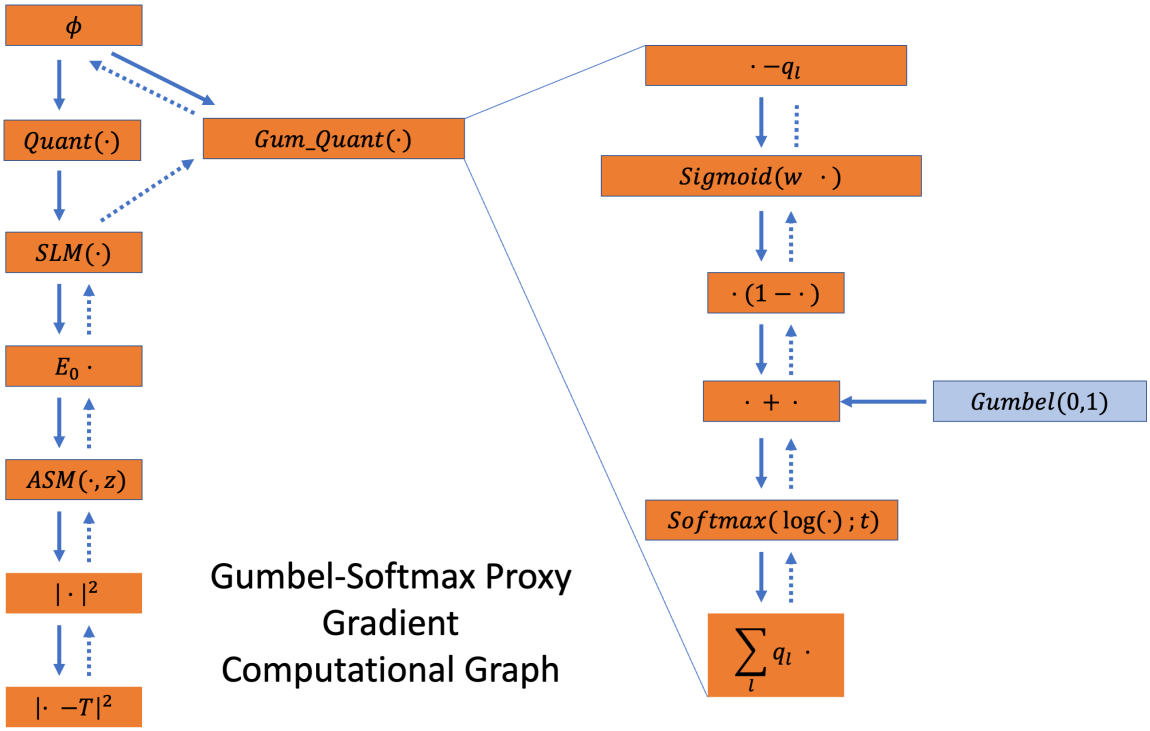


Figure 3.9. The computational graph for the Gumbel Proxy method. This is a fusion of the STE and full Gumbel-Softmax methods which uses the true quantization state on the forward pass and the random gradient updates on the backwards pass. The psuedocode for this algorithm is written in section C.2.7.

3.6. Results and Implementations

To evaluate the effectiveness of the different approaches to optimization the mathematical model from equation 3.15 was implemented using Pytorch 1.9.0. All gradients are computed using Pytorch's automatic differentiation framework. The optimization schemes were employed by altering the training loops from a standard gradient descent approach using the in build ADAM stochastic gradient descent optimization method. Fifty images were chosen from the Div2k data set **[3]** and converted into grayscale using the formula

$$\text{Im}_{gray}(x,y) = \frac{.3\text{Im}(x,y,0) + .6\text{Im}(x,y,1) + .11\text{Im}(x,y,2)}{\max_{x,y}(.3\text{Im}(x,y,0) + .6\text{Im}(x,y,1) + .11\text{Im}(x,y,2))}. \tag{3.38}$$

The optimization is run for each image with a variable number of frames to be multiplex from 1 to 10. The results are then evaluated using PSNR as a metric

$$PSNR(I,G) = 10\log\left(\frac{max(G)}{MSE(I,G)}\right). \tag{3.39}$$

A head to head comparison can be seen in figure 3.11 and 3.10 as well as in appendix B.4.

3.6.0.1. Stochastic Gradient Descent. Due to the fact that this implementation of SGD does not factor gradient information into the problem we should expect the speckle variance to reduce at a rate predicted by the law of large numbers. That is the variance should decrease like $\sim 1/N$ and the PSNR should increase $\sim 2\log(N)$. We can see that this behavior matches SGD quite well in figure 3.10. The average 10 frame PSNR was 29.68 dB for this method.

3.6.0.2. Framewise Predictor Corrector. To implement this approach in pytorch N computational graphs were constructed, each one for a specific frame for an N-frame multiplexing approach. Each graph will have $N-1$ branches which have the non-differentiable quantization operators. After doing n_{it} iterations of SGD for frame i, the current SLM patterns are loaded into the next computational graph which has frame $i+1$ tagged as a parameter and no quantization operator on this branch. Cycle through each frame until desired accuracy is reached. This resulted in a significant improvement over standard SGD with a roughly 5 db improvement in PSNR. However, the scaling as moving to higher and higher frames is logarithmic like SGD resulting in diminishing returns. Another issue is that due to the fact that quantization is being introduced into the problem the overall PSNR does not increase monotonically and can have large drops when switching between frames. To account for this check-pointing is used. The average 10 frame PSNR was 31.48 dB for this method as shown in table 3.1.

3.6.0.3. Single Frame Predictor Corrector. To implement this approach in pytorch two separate graphs were defined and the optimization alternated steps between them. The SLM pattern of each forward model was a tensor of size $(N_x/2, N_y)$ and the forward model will take an additional flag variable to tell it if the model represents the odd or even rows denoted ϕ^o and ϕ^e respectively . When the parameter is passed in it is embedded in the even or odd rows of a zero matrix the size of the actual SLM. Then when the phase delay is applied the the input field pointwise, a mask is applied to the field forcing the magnitude of the field to zero for all un-modulated pixels. The odd and even fields are then propagated and added together before the intensity operation is applied. This implementation strategy allows for reuse of previous modules and makes it

easier to change between optimization methods at a slight cost to memory. As shown in figure 3.10 results show that for low frame counts this method is roughly equivelent to FPC. However, the scaling as more frames are used shows that this method handles large frames better. Furthermore, this method can improve a single frame and is considerably computationally cheaper than FPC as more frames are added because there are only two groups of variables instead of the frame number. The average 10 frame PSNR was 35.64 dB for the 50 images as seen in table 3.1.

3.6.0.4. Straight Through Estimation. The straight through estimator is implemented as a variation of the stochastic gradient descent optical path. A custom quantization module is written which produces a quantized SLM pattern where every pixel is assigned to the nearest quantized state. However, when backpropagating this module just multiplies the input gradient by 1 effectively skipping this node in the computational graph. There is now no guarantee that the gradient point towards an optimum so a robust checkpointing scheme is crucial which tracks the quantized PSNR and saves the best pattern. To initialize the STE alghorithm an initial SGD algorithm is run past convergence and then the weights are modified using STE. Fig. 3.10 demonstrates that this approach greatly improves overall PSNR of the reconstruction. The average 10 frame PSNR was 38.91 dB for the 50 images as seen in table 3.1.

3.6.0.5. Gumbel Softmax. The Gumbel-Softmax method was implemented in Pytorch by augmenting the existing SLM model to include the phase difference, score functions and softmax equation from equation 3.36. The parameter values were hand tuned for two

images to $w = 25$ and the temperature was annealed from $t_0 = 5$ to $t_f \approx .111$

$$t_k = t_0 \exp\left(k \log(45/N_{its})\right). \tag{3.40}$$

The optimization was run for a total of $N_{its} = 2500$ iterations. Overall, for higher frame counts this was the highest performing method on average. In numerical experimentation and tuning the width and temperature scheduling for specific images can improve results further, although is not possible for large batches of images such as the 50 images tested. The method performed poorly for low numbers of multiplexed frames underperforming even SGD with rounding until 3 frames. The 10 frame average PSNR was 39.30 dB for the 50 images as seen in table 3.1.

3.6.0.6. Gumbel Proxy. The Gumbel-Proxy method was implemented in Pytorch by augmenting the existing STE-SLM model to include the backwards pass from equation 3.36 as shown in figure 3.9. The parameter values were hand tuned for two images to $w = 26$ and the temperature was annealed from $t_0 = 5$ to $t_f = .111$

$$t_k = t_0 \exp\left(k \log(45/N_{its})\right). \tag{3.41}$$

The optimization was run for a total of $N_{its} = 2500$ iterations. Like the Gumbel-Softmax this method performed poorly for low numbers of frames in our batch images and improved with higher frame counts to an average of 33.90 dB for 10 frames as seen in table 3.1. Overall, this method was both harder to implement and worse performing than either STE or the Gumbel-Softmax.

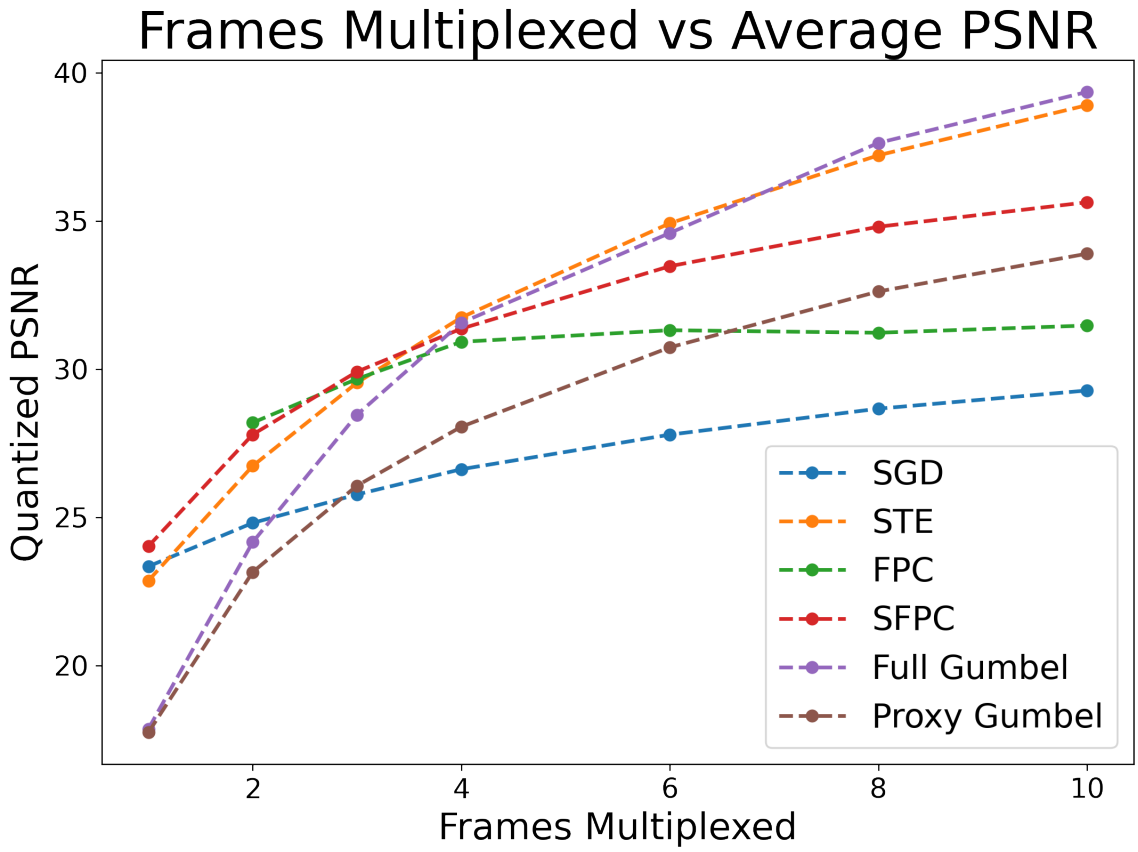


Figure 3.10. A head to head comparison of a number of different quantization aware optimization strategies at different multiplexing levels. The baseline method would be to use a standard gradient descent approach and then quantize after determining the SLM pattern. This is shown in blue and matches the logarithmic profile which would be expected from theory. The variable splitting methods FPC and SFPC both show improvements over the whole range of multiplexed frames. Proxy methods such as STE and Proxy Gumbel seem to consistently produce better results at higher frame counts. It may be possible to improve low frame performance with different parameter regimes which fit better. The full Gumbel Softmax approach suffers for low frame counts but quickly becomes the highest performing method at moderate to high multiplexing numbers.

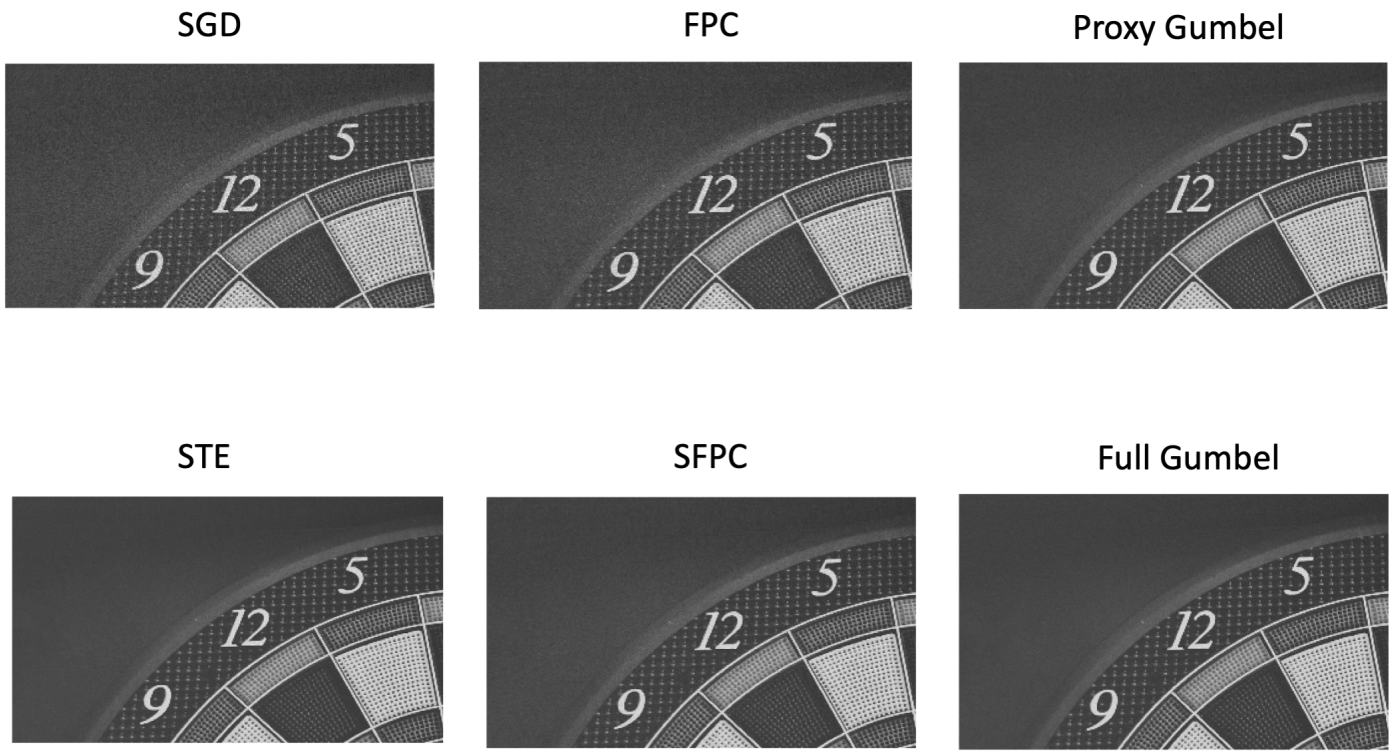


Figure 3.11. Blowup of reconstructed images at 10 frame multiplexing for a number of different optimization methods. All methods show a visual improvement for this image against naive SGD as expected. The Full Gumbel Softmax as well as STE are visually the highest quality.

3.7. Discussion

. This chapter discussed the impact of quantization on image quality for near eye holographic displays and computationally investigated some techniques which can be used to improve the degradation. Near eye holographic displays are a topic of increasing interest at the moment and with the advent of fast spatial light modulators the need for efficient, hardware realistic optimization schemes is greatly needed. Many papers have begun working on incorporating different quantization aware optimization schemes into their results but little head to head discussion between different methods has been done. This chapter laid out a concise theoretical description of the statistics of the speckle generated from SLM phase delay quantization and showed the value of incorporating quantization into the image reconstruction. Quantization aware optimization methods improved the PSNR by huge margins for all methods. The straight through estimators and Gumbel-Softmax estimators can both be easily implemented in a automatic differentiation package and allow for far fewer frames to be multiplexed to attain a desired quality.

While this chapter tested multiple methods head to head it did not show one method is superior in all cases. In particular there are a number of hyperparameters which were set through hand tuning. In the variable splitting methods the number of iterations per group can impact reconstruction quality. The grouping can be seen has a massive impact as SFPC grouped the variables to allow for coherent effects to account for speckle and improved PSNR signficantly. In the cases of both FPC and SFPC the overall computational cost was very high as many iterations need to be done between the groups. Due to the fact that the variable grouping in SFPC allow for a constant number of groups to iterate between no matter the frame count causing the total computational cost to remain

lower. The straight through estimator represents an interesting direction of improving proxy gradients. The concept of a proxy gradient lends itself well to modern automatic differentiation technique and is widely used in machine learning. A proper choice for a proxy gradient can allow for the incorporation of true quantization information into the iteration without halting a gradient based method. The straight through estimator is simple to implement and highly effective. However, as with all proxy gradient methods there is no reason to expect the augmented gradient to point to a local minima meaning early stopping or a consistent check pointing scheme must be used as sometimes the iteration can begin to decrease quality over time. The Gumbel Proxy gradient was an implementation inspired by the impressive results in a recent paper which showed that the Proxy Gumbel would outperform the straight through estimator an several other optimization methods. This paper incorporated a neural network into the model so the results aren't directly comparable but the numerical experimentation in this chapter did not see the same increase in performance. It is also possible that this is due to parameter tuning problems. Further work is needed, however the full Gumbel-Softmax is simpler to implement and not much more computationally expensive. The Gumbel Softmax method performs very well in the high frame case which is very relevant to holographic displays. The method does incorporate random noise meaning that there are issues with inter run reproducability. However, as with all stochastic search methods, the randomness can help prevent getting caught in local minima. The method has a relatively low computational cost and can be hand tuned to produce individual images.

This chapter does not address the incorporation of Neural elements such as CNN layers which can be calibrated in the loop. It is unclear whether the CNN layers can account for

system quantization effects. These camera in the loop calibration methods are becoming standard for all cutting edge high quality displays so the interaction between optimization methods for highly quantized displays and calibration is a topic which will hopefully be addressed in future work. These optimization methods can be used to reduce speckle in traditional SLMs which are calibrated so the primary source of phase error is from quantization. Furthermore, other non differentiable components such as amplitude DMDs can potentially be optimized over in a similar way. Another area where quantization aware methods might be useful is in beam shaping, where an incident coherent light is modulated to account for the deformation from a media such as the layers of atmosphere or tissue. The use of ultra high speed modulators could better account for rapid changes from blood flow or atmospheric fluctuations.

This chapter serves as an exploratory step towards improving holographic displays through computational methods. Circumventing the Etendue trade off with temporal multiplexing allows for improved images with both a large eyebox and field of view making this a realistic possibility for next generation 3D displays. These displays would prevent nausea and eye strain in a form factor which is a fraction of the current size of modern head mounted displays. These methods still must be explored in the context of Neural holography and evaluated in a well calibrated experimental holographic display.

Method	1 Frame	2 Frames	3 Frames	4 Frames	6 Frames	8 Frames	10 Frames
SGD	23.355	24.822	25.782	26.634	27.800	28.676	29.289
FPC	NA	28.204	29.6800	30.932	31.323	31.237	31.481
SFPC	24.037	27.804	29.924	31.374	33.482	34.814	35.639
STE	22.877	26.745	29.552	31.752	34.931	37.224	38.914
Proxy Gumbel	17.757	23.164	26.078	28.066	30.744	32.6300	33.903
Full Gumbel	17.821	24.116	28.162	31.471	34.378	37.672	39.294

Table 3.1. Average results of the different optimization strategies for different levels of temporal multiplexing. The highest 10 frame PSNR is the Full Gumbel Softmax method.

CHAPTER 4

Conclusion

This dissertation was an exploration of several topics in the field of computational imaging and display. The core this thesis is about the numerical modeling of the interaction between optical fields and materials. In this thesis we study two important versions of this problem, one where we don't know the material properties but measure the fields, as in the layering problem in chapter 1 or the attenuation parameter problem posed in chapter 2 and one where the phase of the field is modulated to produce the desired image in the interference pattern as studied in chapter 3.

The source separation problem in chapter 1 represented one corner of a much larger area of research of imaging through a scattering and absorbing media. In this problem we automatically identified portions of the XRF signal which correlated with RIS measurements of the surface of the painting. This simple solution showed how spatial correlations between different signals can be extracted and interpreted. The extracted information relied on the different penetration depths of the spectroscopy methods as well as the registration of the signals. This method can be easily refined with a more complex forward model, physical or learned, which better represents interaction between material properties and light. Specifically, auto-florescence, scattering, and surface layer attenuation could be accounted for to improve recovered signal contrast. While these fine tuned improvements should be done, the essential principle was well demonstrated in this chapter.

Chapter 2 covered my work in OCT signal processing. Two applications of OCT to cultural heritage science are discussed in section 2.2.3. In these papers cross sectional images of painted works of art were taken in a non-contact and non-destructive way. These applications demonstrate the flexibility and usefulness of the technique when trying

to investigate delicate and one of a kind objects. Section 2.2.2 through 2.7 consider the uncertainty in the parameter recovery of the attenuation coefficient. The voxelwise attenuation coefficient is an easy to access metric which can be strongly correlated to useful categories such as tissue type and structure. This chapter modeled the uncertainty in such a technique putting this approach on firmer footing. Basic questions such as whether the parameter is right on average and how variance scales are addressed and a Bayesian framework for denoising applications was proposed. It is unclear how generally applicable the prior derived in that section is and further work is needed to develop this as a despeckling method.

Chapter 3 was an exploratory study which considered how to best incorporate quantization, a real physical system limitation, into the mathematical modeling and optimization. This chapter built out the theory and concepts surrounding holographic displays and then tested a large number of optimization methods which might be able to improve signal quality for temporally multiplexed holographic display. All methods which incorporated quantization improved overall reconstruction PSNR, however the Gumbel Softmax approach had the highest PSNR for moderate to high frame counts, however, the simpler straight through estimator performed almost as well. There was not a lot of literature comparing these methods for holographic displays however, the one paper with such a study for a phase only SLM showed that Proxy Gumbel should outperform STE. This was not reproduced in my numerical experiments. The reason why may come down to several modeling differences and choices and a large parameter space meaning a more precise and mathematical study must be done to understand these different results.

OCT is done in practice by imaging a single line in depth and the raster scanning across the face of an object. However, it is possible to define a 3D OCT technique where uniform plane waves at different frequencies illuminate an object and the scattered field is recorded and then played back. From this perspective, OCT is a holographic imaging technique. Similarly, one might be able to think of the Holographic display problem as optimizing for a hologram which generates the scattered field which would be present in the 3D OCT problem. The simplified 1D approach is done instead because it has a higher lateral resolution and fewer artifacts. These deep connections between OCT and holography helped to enable the transition between the two topics and to help understand the speckle formation in the quantization problem.

The numerical modeling of light material interaction is a topic I have broadly explored and am deeply interested in pursuing further in my career. This thesis represents several completed projects, but is a hopefully just a first step into my future career.

References

[1] 1, Y. Y., Wang, T., Biswal, N. C., Wang, X., Sanders, M., Brewer, M., and Zhu, Q. Optical scattering coefficient estimated by optical coherence tomography correlates with collagen content in ovarian tissue. J. Biomed Opt 16 (2011).

[2] (1848–1903), P. G. Poèmes barbares.

[3] Agustsson, E., and Timofte, R. Ntire 2017 challenge on single image super-resolution: Dataset and study. In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017), pp. 1122–1131.

[4] Alfeld, M., and de Viguerie, L. Recent developments in spectroscopic imaging techniques for historical paintings - a review. Spectrochimica Acta Part B: Atomic Spectroscopy 136 (2017), 81 – 105.

[5] Alfeld, M., and de Viguerie, L. Recent developments in spectroscopic imaging techniques for historical paintings - a review. Spectrochimica Acta Part B: Atomic Spectroscopy 136 (2017), 81–105.

[6] Alfeld, M., Nolf, W. D., Cagno, S., Appel, K., Siddons, D. P., Kuczewski, A., Janssens, K., Dik, J., Trentelman, K., Walton, M., and Sartorius, A. Revealing hidden paint layers in oil paintings by means of scanning macro-xrf: a mock-up study based on rembrandt's "an old man in military costume. J. Anal. At. Spectrom. 40, 51 (2013), 432–444.

[7] Alfeld, M., Pedetti, S., Martinez, P., and Walter, P. Joint data treatment for vis–nir reflectance imaging spectroscopy and xrf imaging acquired in the theban necropolis in egypt by data fusion and t-sne. Comptes Rendus Physique 19, 7 (2018), 625 – 635. Physics and arts / Physique et arts.

[8] Almasian, M., Bosschaart, N., van Leeuwen, T. G., and Faber, D. J. Validation of quantitative attenuation and backscattering coefficient measurements by optical coherence tomography in the concentration-dependent and multiple scattering regime. Journal of Biomedical Optics 20, 12 (2015), 1 – 11.

[9] Almasian, M., Fiske, L. D., Raven, L., Smelt, S., Keune, K., and Aalders, M. C. G. Revealing accurate paint surface topologies of the night watch by optical coherence tomography. Cultural Heritage Science (2022 (in preparation)).

[10] Almasian, M., Tiennot, M., Fiske, L. D., and Hermens, E. The use of ground glass in red glazes: structural 3d imaging and mechanical behaviour using optical coherence tomography and nanoindentation. Heritage Science 9 (2021).

[11] Almasian, M., van Leeuwen, T., and Faber, D. J. Oct amplitude and speckle statistics of discrete random media. Nature Scientific Reports 7 (Nov 2017).

[12] Banks, M. S., Hoffman, D. M., Kim, J., and Wetzstein, G. 3d displays. Annual Review of Vision Science 2, 1 (2016), 397–435. PMID: 28532351.

[13] Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A., and Kreshuk, A. ilastik: interactive machine learning for (bio)image analysis. Nature Methods (Sept. 2019).

[14] Bertrand, L., Schöeder, S., Anglos, D., Breese, M. B., Janssens, K., Moini, M., and Simon, A. Mitigation strategies for radiation damage in the analysis of ancient materials. TrAC Trends in Analytical Chemistry 66 (2015), 128–145.

[15] Blanche, P.-A., and Ketchum, R. S. Texas instruments phase light modulator for holography. In OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP) (2021), Optica Publishing Group, p. DW4B.3.

[16] Blugel, S. Scattering theory: Born series, February 2012.

[17] Born, M., and Wolf, E. Principles of optics : electromagnetic theory of propagation, interference and diffraction of light, 7th (expanded) ed. ed. Cambridge University Press, Cambridge ;, 1999.

[18] Callewaert, T., Dik, J., and Kalkman, J. Segmentation of thin corrugated layers in high-resolution oct images. OSA 25, 26 (Dec 2017), 32816–32828.

[19] Chakravarthula, P., Tseng, E., Srivastava, T., Fuchs, H., and Heide, F. Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. ACM Transactions on Graphics (TOG) 39, 6 (2020), 186.

[20] Chan, A. C., Kurokawa, K., Makita, S., Miura, M., and Yasuno, Y. Maximum a posteriori estimator for high-contrast image composition of optical coherence tomography. Opt. Lett. 41, 2 (Jan 2016), 321–324.

[21] Chandrasekhar, S. Stochastic problems in physics and astronomy. Reviews of Modern Physics 15, 1 (Jan 1943), 81–82.

[22] Chang, C., Bang, K., Wetzstein, G., Lee, B., and Gao, L. Toward the next-generation vr/ar optics: a review of holographic near-eye displays from a human-centric perspective. Optica 7, 11 (Nov 2020), 1563–1578.

[23] Chang, C., Bang, K., Wetzstein, G., Lee, B., and Gao, L. Toward the next-generation vr/ar optics: a review of holographic near-eye displays from a human-centric perspective. Optica 7, 11 (Nov 2020), 1563–1578.

[24] Chang, S., and Bowden, A. K. Review of methods and applications of attenuation coefficient measurements with optical coherence tomography. Journal of Biomedical Optics 24, 9 (2019), 1 – 17.

[25] Choi, S., Gopakumar, M., Peng, Y., Kim, J., O'Toole, M., and Wetzstein, G. Time-multiplexed neural holography: A flexible framework for holographic near-eye displays with fast heavily-quantized spatial light modulators. In Proceedings of the ACM SIGGRAPH (2022), p. 1–8.

[26] Chopp, D. Lecture notes for numerical methods for partial differential equations, November 2011.

[27] Commons, W. File:the nightwatch by rembrandt - rijksmuseum.jpg — wikimedia commons, the free media repository, 2021. [Online; accessed 10-June-2022].

[28] Dai, Q., Pouyet, E., Cossairt, O., Walton, M., and Katsaggelos, A. K. Spatial-spectral representation for x-ray fluorescence image super-resolution. IEEE Transactions on Computational Imaging 3, 3 (2017), 432–444.

[29] de Almeida Nieto, L. XRF and RIS for semi-quantitative sub-surface layer detection and composition analysis of easel paintings. Master's thesis, TU Delft, the Netherlands, 2020.

[30] Deligiannis, N., Mota, J. F. C., Cornelis, B., Rodrigues, M. R. D., and Daubechies, I. X-ray image separation via coupled dictionary learning. In 2016 IEEE International Conference on Image Processing (ICIP) (2016), pp. 3533–3537.

[31] Deng, Y., and Chu, D. Coherence properties of different light sources and their effect on the image sharpness and speckle of holographic displays. Scientific Reports 7, 7 (2017).

[32] Dik, J., Janssens, K., Van Der Snickt, G., van der Loeff, L., Rickers, K., and Cotte, M. Visualization of a lost painting by vincent van gogh using synchrotron radiation based x-ray fluorescence elemental mapping. Analytical Chemistry 80, 16 (2008), 6436–6442. PMID: 18662021.

[33] Dooley, K. A., Conover, D. M., Glinsman, L. D., and Delaney, J. K. Complementary standoff chemical imaging to map and identify artist materials in an early italian renaissance panel painting. Angewandte Chemie International Edition 53, 50 (2014), 13775–13779.

[34] Dooley, K. A., Gifford, E. M., van Loon, A., Noble, P., Zeibel, J. G., Conover, D. M., Alfeld, M., der Snickt, G. V., Legrand, S., Janssens, K., Dik, J., and Delaney, J. K. Separating two painting campaigns in saul and david, attributed to rembrandt, using macroscale reflectance and xrf imaging spectroscopies and microscale paint analysis. Heritage Science 6, 46 (2018).

[35] Dunn, D. Required accuracy of gaze tracking for varifocal displays. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (2019), pp. 1838–1842.

[36] Evelyne, C. e. a. Quantitative measurement of attenuation coefficients of bladder biopsies using optical coherence tomography for grading urothelial carcinoma of the bladder. J. Biomed. Opt 15, 6 (2010), 1 – 6.

[37] Eybposh, M. H., Curtis, V. R., Moossavi, A., and Pégard, N. C. Dynamic computer generated holography for virtual reality displays. In Frontiers in Optics + Laser Science 2021 (2021), Optica Publishing Group, p. FM3B.3.

[38] Faber, D. J., van der Meer, F. J., Aalders, M. C., and van Leeuwen, T. G. Quantitative measurement of attenuation coefficients of weakly scattering media using optical coherence tomography. Opt. Express 12, 19 (Sep 2004), 4353–4365.

[39] Fercher, A., Hitzenberger, C., Kamp, G., and El-Zaiat, S. Measurement of intraocular distances by backscattering spectral interferometry. Optics Communications 117, 1 (1995), 43–48.

[40] Fercher, A., Hu, H., Steeger, P., and Briers, J. Eye deformation measurement by laser interferometry. Optica Acta: International Journal of Optics 29, 10 (1982), 1401–1406.

[41] Fercher, A. F. Fercher, verfahren und anordnung zur messung der teilstrecken des lebenden auges, Deutsches Patentamt, Germany (1983).

[42] Ferretti, M. X-ray Fluorescence Applications for the Study and Conservation of Cultural Heritage. TU Delft, 12 2000, pp. 285–296.

[43] Fuentes, J. L. M., Fernández, E. J., Prieto, P. M., and Artal, P. Interferometric method for phase calibration in liquid crystal spatial light modulators using a self-generated diffraction-grating. Optics Express 24, 13 (2016), 14159–14171.

[44] Gabor, D. A new microscopic principle. Nature 161 (1948), 777–778.

[45] Geng, J. Three-dimensional display technologies. Adv. Opt. Photon. 5, 4 (Dec 2013), 456–535.

[46] Gholami, A., Kim, S., Dong, Z., Yao, Z., Mahoney, M. W., and Keutzer, K. A survey of quantization methods for efficient neural network inference, 2021.

[47] Gong, P., Almasian, M., van Soest, G., de Bruin, D. M., van Leeuwen, T. G., Sampson, D. D., and Faber, D. J. Parametric imaging of attenuation by optical coherence tomography: review of models, methods, and clinical translation. Journal of Biomedical Optics 25, 4 (2020), 1 – 34.

[48] Goodman, J. W. Speckle Phenomena in Optics. Roberts and Company, 2007.

[49] Gumbel, E. J. Statistical theory of extreme values and some practical applications: a series of lectures, vol. 33. US Government Printing Office, 1954.

[50] Hale, G. M., and Querry, M. R. Optical constants of water in the 200-nm to 200-µm wavelength region. Appl. Opt. 12 (1973).

[51] Harthong, J., Sadi, J., Torzynski, M., and Vukicevic, D. Speckle phase averaging in high-resolution color holography. J. Opt. Soc. Am. A 14, 2 (Feb 1997), 405–410.

[52] Hasnain, A., Laffont, P.-Y., Jalil, S. B. A., Buyukburc, K., Guillemet, P.-Y., Wirajaya, S., Khoo, L., Deng, T., and Bazin, J.-C. Piezo-actuated varifocal head-mounted displays for virtual and augmented reality. In Advances in Display Technologies IX (2019), J.-H. Lee, Q.-H. Wang, and T.-H. Yoon, Eds., vol. 10942, International Society for Optics and Photonics, SPIE, pp. 40 – 50.

[53] Hébert, M., and Emmel, P. Two-flux and multiflux matrix models for colored surfaces. M. Kriss. Handbook of Digital Imaging, vol. 2. Wiley, 2015.

[54] Hillman, T. R., Adie, S. G., Seemann, V., Armstrong, J. J., Jacques, S. L., and Sampson, D. D. Correlation of static speckle with sample properties in optical coherence tomography. Opt. Lett. 31, 2 (Jan 2006), 190–192.

[55] Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A., and Fujimoto, J. G. Optical coherence tomography. Science 254, 5035 (1991), 1178–1181.

[56] Hughes, D. I., and Duck, F. A. Automatic attenuation compensation for ultrasonic imaging. Ultrasound in Medicine and Biology 23, 5 (1997), 651 – 664.

[57] Inc., M. Quest 2, 2022. [Online; accessed 22-June-2022].

[58] Jang, E., Gu, S., and Poole, B. Categorical reparameterization with gumbel-softmax, 2016.

[59] Kim, M. K. Holographic Microscopy: Principles, Techniques, and Applications. 1st ed. Springer, New York, 2011.

[60] Kirillin, M. Y., Farhat, G., Sergeeva, E. A., Kolios, M. C., and Vitkin, A. Speckle statistics in oct images: Monte carlo simulations and experimental studies. Opt. Lett. 39, 12 (Jun 2014), 3472–3475.

[61] Knuettel, A. R., and Boehlau-Godau, M. Spatially confined and temporally resolved refractive index and scattering evaluation in human skin performed with optical coherence tomography. J. Biomed Opt 5, 1 (2000), 83–92.

[62] Kramida, G. Resolving the vergence-accommodation conflict in head-mounted displays. IEEE Transactions on Visualization and Computer Graphics 22, 7 (2016), 1912–1931.

[63] Kuo, G., Waller, L., Ng, R., and Maimone, A. High resolution Étendue expansion for holographic displays. ACM Trans. Graph. 39, 4 (jul 2020).

[64] Kärtner, F. Fundamentals of photonics: Quantum electronics. Massachusetts Institute of Technology: MIT OpenCouseWare https://ocw.mit.edu/ (Spring 2006), License: Creative Commons BY–NC–SA.

[65] Lachmann, T., van der Snickt, G., Haschkec, M., and Mantouvalou, I. Combined 1d, 2d and 3d micro-xrf techniques for the analysis of illuminated manuscripts. J. Anal. At. Spectrom., 31 (2016).

[66] Lanman, D., and Luebke, D. Near-eye light field displays. ACM Trans. Graph. 32, 6 (nov 2013).

[67] Lee, B., Kim, D., Lee, S., Chen, C., and Lee, B. High-contrast, speckle-free, true 3d holography via binary cgh optimization. Scientific Reports 12, 1 (Feb 2022), 2811.

[68] Li, R., Peng, T., Zhou, M., Yu, X., Min, J., Yang, Y., and Yao, B. Full-polarization wavefront shaping for imaging through scattering media. Appl. Opt. 59, 17 (Jun 2020), 5131–5135.

[69] Liang, H. Advances in multispectral and hyperspectral imaging for archaeology and art conservation. Appl. Phys. A 106 (2012), 309–3230.

[70] Liu, J., Ding, N., Yu, Y., Yuan, X., Luo, S., Luan, J., Zhao, Y., Wang, Y., and Ma, Z. Optimized depth-resolved estimation to measure optical attenuation coefficients from optical coherence tomography and its application in cerebral damage determination. Journal of Biomedical Optics 24, 3 (2019), 1 – 11.

[71] Liu, S., Sotomi, Y., Eggermont, J., Nakazawa, G., Torii, S., M.D., T. I., M.D., Y. O., M.D., P. W. S., Lelieveldt, B. P. F., and Dijkstra, J. Tissue characterization with depth-resolved attenuation coefficient and backscatter term in intravascular optical coherence tomography images. Journal of Biomedical Optics 22, 9 (2017), 1 – 16.

[72] Lohmann, A. W., and Paris, D. P. Binary fraunhofer holograms, generated by computer. Appl. Opt. 6, 10 (Oct 1967), 1739–1748.

[73] Lutzenberger, K., Stege, H., and Tilenschi, C. A note on glass and silica in oil paintings from the 15th to the 17th century. Journal of Cultural Heritage 11, 4 (2010), 365–372.

[74] Maimone, A., and Wang, J. Holographic optics for thin and lightweight virtual reality. ACM Trans. Graph. 39, 4 (jul 2020).

[75] MATLAB. version R2019a. The MathWorks Inc., Natick, Massachusetts, 2019.

[76] Mercier, O., Sulai, Y., Mackenzie, K., Zannoli, M., Hillis, J., Nowrouzezahrai, D., and Lanman, D. Fast gaze-contingent optimal decompositions for multifocal displays. ACM Trans. Graph. 36, 6 (nov 2017).

[77] Microsoft. Hololens, 2022. [Online; accessed 22-June-2022].

[78] Min, K., and Park, J.-H. Quality enhancement of binary-encoded amplitude holograms by using error diffusion. Opt. Express 28, 25 (Dec 2020), 38140–38154.

[79] Moon, E., Kim, M., Roh, J., Kim, H., and Hahn, J. Holographic head-mounted display with rgb light emitting diode light source. Opt. Express 22, 6 (Mar 2014), 6526–6534.

[80] Morris, C. N. SParametric Empirical Bayes Inference. Journal of the American Statistical Association, 1983.

[81] Mounier, A., and Daniel, F. Hyperspectral imaging for the study of two thirteenth-century italian miniatures from the marcadé collection, treasury of the saint-andre cathedral in bordeaux, france. Studies in Conservation 60, sup1 (2015), S200–S209.

[82] Muller, B. G., van Kollenburg, R. A. A., Swaan, A., Zwartkruis, E. C. H., Brandt, M. J., Wilk, L. S., Almasian, M., Schreurs, A. W., Faber, D. J., Rozendaal, L. R., Vis, A. N., Nieuwenhuijzen, J. A., van Moorselaar, J. R. J. A., de la Rosette M.D., J. J. H., de Bruin, D. M., and van Leeuwen, T. G. Needle-based optical coherence tomography for the detection of prostate cancer: a visual and quantitative analysis in 20 patients. Journal of Biomedical Optics 23, 8 (2018), 1 – 11.

[83] Neelmeijer, C., Brissaud, I., Calligaro, T., Demortier, G., Hautojärvi, A., Mäder, M., Martinot, L., Schreiner, M., Tuurnala, T., and Weber, G. Paintings— a challenge for xrf and pixe analysis. X-Ray Spectrometry 29, 1 (2000), 101–110.

[84] Newspaper, T. A. Trimmed, splashed and slashed: the anatomy of rembrandt's the night watch.

[85] Pandey, N., and Hennelly, B. Quantization noise and its reduction in lensless fourier digital holography. Appl. Opt. 50, 7 (Mar 2011), B58–B70.

[86] Peng, Y., Choi, S., Kim, J., and Wetzstein, G. Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration. Science Advances 7, 46 (2021), eabg5040.

[87] Peng, Y., Choi, S., Padmanaban, N., and Wetzstein, G. Neural holography with camera-in-the-loop training. ACM Trans. Graph. 39, 6 (nov 2020).

[88] Picollo, M., Cucci, C., Casini, A., and Stefani, L. Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios. Sensors 20, 10 (2020).

[89] Pouyet, E., Devine, S., Grafakos, T., Kieckhefer, R., Salvant, J., Smieska, L., Woll, A., Katsaggelos, A., Cossairt, O., and Walton, M. Revealing the biography of a hidden medieval manuscript using synchrotron and conventional imaging techniques. Analytica Chimica Acta 982 (2017), 20 – 30.

[90] Prize, N. Dennis gabor.

[91] Pu, W., Sober, B., Daly, N., Sabetsarvestani, Z., Higgitt, C., Daubechies, I., and Rodrigues, M. R. Image separation with side information: A connected auto-encoders based approach. Arxiv 3, 3 (2017), 432–444.

[92] Rampley, M., Lenain, T., Locher, H., Pinotti, A., Schoell-Glass, C., and Zijlmans, C. K. Art History and Visual Studies in Europe: Transnational Discourses and National Frameworks. Brill, Leiden, The Netherlands, 2012.

[93] Rapuc, W., Jacq, K., Develle, A.-L., Sabatier, P., Fanget, B., Perrette, Y., Coquin, D., Debret, M., Wilhelm, B., and Arnaud, F. Xrf and hyperspectral analyses as an automatic way to detect flood events in sediment cores. Sedimentary Geology 409 (2020), 105776.

[94] Rijksmuseum. Rijksmuseum: Operation nightwatch. Accessed: 2021-01-11.

[95] Rohani, N., Pouyet, E., Walton, M., Cossairt, O., and Katsaggelos, A. K. Nonlinear unmixing of hyperspectral datasets for the study of painted works of art. Angewandte Chemie 130, 34 (2018), 11076–11080.

[96] Schmitt, J. M., Xiang, S. H., and Yung, K. M. Speckle in optical coherence tomography. Journal of Biomedical Optics 4, 1 (1999), 95 – 105.

[97] Shi, L., Li, B., Kim, C., Kellnhofer, P., and Matusik, W. Towards real-time photorealistic 3d holography with deep neural networks. Nature 591, 7849 (Mar 2021), 234–239.

[98] Shull, K., Katsaggelos, K., Casadio, F., Walton, M., and Cossairt, O. Computationally based imaging of structure in materials. Accessed: 2022-05-19.

[99] Simonot, L. Etude expérimentale et modélisation de la diffusion de la lumière dans une couche de peinture colorée et translucide. Dissertation (11 2002).

[100] Smith, G. T., Dwork, N., O'Connor, D., Sikora, U., Lurie, K. L., Pauly, J. M., and Ellerbee, A. K. Automated, depth-resolved estimation of the attenuation coefficient from optical coherence tomography data. IEEE Transactions on Medical Imaging 34, 12 (Dec 2015), 2592–2602.

[101] Smithsonian. [Online; accessed 30-December-2022].

[102] Solé, V., Papillon, E., Cotte, M., Walter, P., and Susini, J. multiplatform code for the analysis of energy-dispersive x-ray fluorescence spectra. Spectrochim. Acta Part B 62 (2007), 63–68.

[103] Sutherland, I. E. A head-mounted three dimensional display. In Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (New York, NY, USA, 1968), AFIPS '68 (Fall, part I), Association for Computing Machinery, p. 757–764.

[104] Tarantola, A. Inverse Problem Theory. SIAM, 2005.

[105] Thamali, W. Sampling; quantization in digital image processing, Nov 2020.

[106] Times, N. Y. Rembrandt's 'the night watch' slashed.

[107] Times, T. L. Rembrandt's 'night watch' painting vandalized : Artwork: Jobless dutchman sprays chemical on painting; damage reported minimal.

[108] Tonon, C., Rozé, C., Girasole, T., and Dinguirard, M. Four-flux model for a multilayer, plane absorbing and scattering medium: application to the optical degradation of white paint in a space environment. Appl. Opt. 40, 22 (Aug 2001), 3718–3725.

[109] Trojek, T., and Trojková, D. Several approaches to the investigation of paintings with the use of portable x-ray fluorescence analysis. Radiation Physics and Chemistry 116 (2015), 321–325. Proceedings of the 9th International Topical Meeting on Industrial Radiation and Radioisotope Measurement Applications.

[110] Turchin, I. V., Sergeeva, E. A., Dolin, L. S., and Richards-Kortum, N. M. S. M. R. R. Novel algorithm of processing optical coherence tomography images for differentiation of biological tissue pathologies. J. Biomed Opt 10, 6 (2005).

[111] van der Meer, F. J., Faber, D. J., Perrée, J., Pasterkamp, G., Baraznji Sassoon, D., and van Leeuwen, T. G. Quantitative optical coherence tomography of arterial wall components. Lasers in Medical Science 20, 1 (Jul 2005), 45–51.

[112] van Leeuwen, T. G., Faber, D. J., and Aalders, M. C. Measurement of the axial point spread function in scattering media using single-mode fiber-based optical coherence tomography. IEEE Journal of Selected Topics in Quantum Electronics 9, 2 (March 2003), 227–233.

[113] Velez-Zea, A., Quinchia, S. B., Barrera-Ramírez, J. F., and Torroba, R. Fast computation of binary amplitude holograms with optimized random phases. In OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP) (2021), Optica Publishing Group, p. DTu7B.5.

[114] Vermeer, K. A., Mo, J., Weda, J. J. A., Lemij, H. G., and de Boer, J. F. Depth-resolved model-based reconstruction of attenuation coefficients in optical coherence tomography. Biomed. Opt. Express 5, 1 (Jan 2014), 322–337.

[115] Vermeulen, M., Smith, K., Eremin, K., Rayner, G., and Walton, M. Application of uniform manifold approximation and projection (umap) in spectral imaging of artworks. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 252 (2021), 119547.

[116] Yuan, W., Kut, C., Liang, W., and Li, X. Robust and fast characterization of oct-based optical attenuation using a novel frequency-domain algorithm for brain cancer detection. Sci Rep 7 (2017).

[117] Zhan, T., Xiong, J., Zou, J., and Wu, S.-T. Multifocal displays: review and prospect. PhotoniX 1, 1 (Mar 2020), 10.

[118] Zhang, J., Pégard, N., Zhong, J., Adesnik, H., and Waller, L. 3d computer-generated holography by non-convex optimization. Optica 4, 10 (Oct 2017), 1306–1313.

[119] Zhang, P., Miller, E. B., Manna, S. K., Meleppat, R. K., Pugh, E. N., and Zawadzki, R. Temporal speckle-averaging of optical coherence tomography volumes for in-vivo cellular resolution neuronal and vascular retinal imaging. Neurophotonics 6, 4 (2019), 1 – 13.

[120] Zhou, H., Chu, Z., Zhang, Q., Dai, Y., Gregori, G., Rosenfeld, P. J., and Wang, R. K. Attenuation correction assisted automatic segmentation for assessing choroidal thickness and vasculature with swept-source oct. Biomed. Opt. Express 9, 12 (Dec 2018), 6067–6080.

[121] Łach, B., Fiutowski, T., Koperny, S., Krupska-Wolas, P., Lankosz, M., Mendys-Frodyma, A., Mindur, B., Świentek, K., Wiącek, P., Wróbel, P. M., and Dąbrowski, W. Application of factorisation methods to analysis

of elemental distribution maps acquired with a full-field xrf imaging spectrometer. Sensors (Basel) (2021).

APPENDIX A

Appendix A: OCT Derivations

A.1. Fourier Domain OCT Derivation

To understand the working principle of a modern Fourier domain OCT, we must model the interaction between a small semi-translucent object of interest **[39]**. Consider a single frequency, time harmonic Gaussian beam of the form

$$E^{inc}(\boldsymbol{x}) = A(x,y)exp(i\ \boldsymbol{k}\cdot\boldsymbol{x})$$

which is sent into a media with index of refraction variations from air given by $f(\boldsymbol{x})$. The dynamics of the electric field, denoted by E, is well modeled by the inhomogenious Helmholtz equation

$$\nabla^2 E + \left(f(\boldsymbol{x}) + k^2\right) E = 0$$

with the Sommerfield radiation conditions

$$\lim_{|\boldsymbol{x}|\to\infty} \left(\frac{\partial}{\partial|\boldsymbol{x}|} - ik\right) E \to 0.$$

It is will established that such a beam will approximately obey the homogeneous equation

$$\nabla^2 E^{inc} + k^2 E^{inc} \approx 0.$$

when the paraxial approximation is valid **[64]**. Using linearity, the related problem difference between the scattered and unscattered fields can be considered equivalently

$$\nabla^2 \left(E - E^{inc}\right) + k^2 \left(E - E^{inc}\right) = fE. \tag{A.1}$$

This system is well studied and has the fundamental solution

$$G(\boldsymbol{x};\boldsymbol{x}_0) = \frac{e^{ik^s\cdot|\boldsymbol{x}-\boldsymbol{x}_0|}}{|\boldsymbol{x}-\boldsymbol{x}_0|}.$$

Multiplying both sides of equation A.1 by G and integrating results in the integral equation

$$E(\boldsymbol{x};\boldsymbol{k}) - E^{inc}(\boldsymbol{x};\boldsymbol{k}) = \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0) E(\boldsymbol{x}_0) d\boldsymbol{x}_0$$

or

$$\left(\mathbb{I} - \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)\right) E = E^{inc}.$$

Which has the formal solution

$$E = \left(\mathbb{I} - \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)\right)^{-1} E^{inc}.$$

Using the assumption $\sup_{g\in L_2} |\int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(x;x_0) f(\boldsymbol{x}_0)| < 1$ then we can safely compute the solution with a Nuemann expansion of the inverse operator:

$$\left(\mathbb{I} - \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)\right)^{-1} = \sum_{n=0}^{\infty} \left(\mathbb{I} - \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)\right)^{n}.$$

This type of solution is for a light scattering system is called a Born series and is a topic which has been well studied **[16, 17]**. Each term in this series represents a higher order scattering event. OCT is typically used for systems which are only weakly backscattering, meaning that the vast majority of the light is given by the first scattering event **[39]**. Thus, we are justified in keeping only the first two terms in the series

$$\left(\mathbb{I} - \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)\right)^{-1} \sim \mathbb{I} + \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0)$$

and computed the scattered field as:

$$E \sim E^{inc} + \int_{\mathbb{R}^3} d\boldsymbol{x}_0 G(\boldsymbol{x};\boldsymbol{x}_0) f(\boldsymbol{x}_0) E^{inc}(\boldsymbol{x}_0).$$

Plugging in the incident field and G we see that

$$E \sim E^{inc} + \int_{\mathbb{R}^3} d\boldsymbol{x}_0 A(x,y) \frac{f(\boldsymbol{x}_0)}{|\boldsymbol{x} - \boldsymbol{x}_0|} \exp\left(i\left(\boldsymbol{k} - \boldsymbol{k^s}\right) \cdot \boldsymbol{x}_0 - i\boldsymbol{k}^s \cdot \boldsymbol{x}\right)$$

If we assume our object of interest is relatively small and some distance away from the origin of our coordinate system we can approximate $\frac{1}{|\boldsymbol{x}-\boldsymbol{x}_0|} \approx D$ which is constant. Furthermore, since in practice an incident beam with a very narrow waist is used, the support of $A(x,y)$ can be very small, such that the scattering potential f doesn't vary much in x and y in depth. Thus, we can approximate the integration in these dimensions as a fixed multiplicative constant w. Combining these assumptions we have

$$E \sim E^{inc} + \frac{A(x,y)Dw}{4\pi} \exp\left(-i\boldsymbol{k}^s \cdot \boldsymbol{x}\right) \int_{\mathbb{R}} dz_0 f(z_0) \exp\left(iKz_0\right). \tag{A.2}$$

where $K = \left(\boldsymbol{k} - \boldsymbol{k^s}\right)_z$. More simply for the scattered field $E^s = E - E^{inc}$ we have

$$E^s(\boldsymbol{k};\boldsymbol{x}) \approx \frac{A(x,y)Dw}{4\pi} \exp\left(-i\boldsymbol{k}^s \cdot \boldsymbol{x}\right) \mathcal{F}[f(z)] \tag{A.3}$$

From this it is clear that if the scattered field can be measured for a range of wavelengths, the scattering potential (index of refraction) for a given object.

The scattered electric field is complex, and most detectors can only sense the intensity of the electric field losing the phase information. The intensity of the scattered field relates to the scattering potential through

$$F^{-1}\left[||E^s(\boldsymbol{k};\boldsymbol{x})||^2\right] \propto F^{-1}\left[|\mathcal{F}[f(z)]|^2\right] = corr(f(Z), f(Z))(z) \tag{A.4}$$

which is the autocorrelation of the scattering potential. Clearly, measuring the signal intensity will lead to reconstruction ambiguities, as auto-correlation is not a one to one function. This is sensible since the intensity operation is throwing away half or more of the information in the Electric field. To avoid this issue, we can use interferometry to gain additional information about the signal. Using a mirror and a beam splitter we can add a known incident beam to the the detector causing an interference pattern. Assume the reference beam is given by $E^r = R\exp(ikz_r) = RF[\delta(z - z_r)]$ where R is a constant depending on the reflectivity and beam intensity. The interference pattern measured on the detector is

$$I(k; x, y, z) = ||E^r + E^s||^2 = ||E^r||^2 + ||E^s||^2 + E^s E^{r*} + E^{s*} E^r$$

The Fourier transform of the correlation terms is

$$F^{-1}[I] \propto corr(f(z), f(z)) + R^2\delta(z - z_r) + F^{-1}[E^s E^{r*}] + F^{-1}[E^{s*} E^r]. \tag{A.5}$$

The first term in this expression is the auto correlation of the scattering potential and is not impacted by the reference mirror position given by z_r. The second term is a delta peak from the beam intensity which can be in practice subtracted out from the signal. The remaining two terms are conjugate mirrored copies of the scattering potential. This is clear as

$$F^{-1}[E^s E^{r*}] = corr(f(Z), R\delta(Z + z_0))(z)$$

$$F^{-1}[[E^{s*}, E^r] = corr(f(Z)^*, R\delta(Z - z_0))(z).$$

The specific mirror position z_0 impacts the location in the Fourier domain that the copies of f(z) are centered. If we make the compactness assumption that $f(z) = 0$ for $z < 0 \cup z > L$ and assume the mirror position is $z_0/2 > L$ then none of these terms will overlap in the Fourier domain and $f(z)$ can be viewed directly.

In practice, the wavelengths which was used to illuminate a sample are given by a certain bandwidth. Because of this, the measured signal will be a sampled copy of this FFT with an effective high pass filter and resolution depending on the center wavelength and bandwidth.

A.2. Derivation of Scattered Electric Field Statistics

Given a volumetric scattering material backscattering and incident electric field, the real and imaginary parts of the electric field at the detector are well modeled by pixelwise gaussian random variable. To see this, assume that there are N_p particles with positions which are randomly (uniformly) distributed throughout an imaging voxel. The electric

field on the detector from the i^{th} particle will have the phase

$$\phi_i = mod(\frac{z + \xi_i}{\lambda}, 2\pi) \tag{A.6}$$

where ξ_i is a uniformly distributed random variable. Assuming the wavelength of light is sufficiently smaller than the axial depth resolution, then the phase is effectively random and uniformly distributed between $(-\pi, \pi)$ after phase wrapping. This means that the real and imaginary part of the field at the detector from each particle will be a random variables with the constraint that the intensity is fixed. The complex representation of the total electric field incident on the detector will be proportional to

$$E_z = \sum_{i=1}^{N} E_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |a_i| \cos(\phi_i) + \frac{i}{\sqrt{N}} \sum_{i=1}^{N} |a_i| \sin(\phi_i). \tag{A.7}$$

By the central limit theorem we should expect the real and imaginary parts of this field to approach Gaussian random variables as $N \to \infty$, which is fully defined by the mean and variance. We can compute these quantities for the real and imaginary parts as

$$\begin{aligned} \langle Re[E_z] \rangle &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \langle |a_i| \cos(\phi_i) \rangle = 0 \\ \langle Im[E_z] \rangle &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \langle |a_i| \sin(\phi_i) \rangle = 0 \end{aligned} \tag{A.8}$$

and by orthogonality

$$\langle Re[E_z]^2\rangle = \sum_{i=1}^{N} \frac{|a_i|^2}{2N} \tag{A.9}$$

$$\langle Im[E_z]^2\rangle = \sum_{i=1}^{N} \frac{|a_i|^2}{2N}$$

$$\langle Im[E_z]Re[E_z]\rangle = 0.$$

Letting $\sum_{i=1}^{N} \frac{|a_i|^2}{2N} = \sigma^2$ Thus we have that the real and imaginary parts of the electric field at the detector are, $Im[E_z] \sim G\left(0, \sigma^2\right)$ and $Re[E_z] \sim G\left(0, \sigma^2\right)$ respectively. The detector then can detect the real valued intensity of this field as $I_z = |Im[E_z]|^2 + |Re[E_z]|^2$ which is the length of a 2D Gaussian random vector.

A.3. Correcting Beam Profile and Sensitivity Roll Off

An OCT signal model to account for OCT signal decay, confocal point spread function (PSF) and sensitivity roll off effects was derived by Faber **[8, 112]**. This model defines the OCT signal in depth as

$$I_T(z;\mu_B,\mu_{oct}) = \alpha \cdot T(z - z_f) \cdot H(z) \cdot \mu_B \exp\left(-2\int_0^z \mu_{oct}(\tau)d\tau\right) \tag{A.10}$$

where $I_T(z;\mu_B,\mu_{oct})$ is the measured signal at depth z, μ_B is the potentially spatially varying back-scattering coefficient, μ_{oct} is the potentially spatially varying attenuation coefficient. The quantity T is the confocal PSF defined by

$$T(z - z_f) = \frac{1}{1 + \left(\frac{z - z_f}{2nz_{R0}}\right)^2}$$

where z_f is the focus position in depth and z_{R0} is the apparent Rayleigh length given by

$$z_{R0} = \pi w_0^2/\lambda_0.$$

In this expression, w_0 is the radius of the beam waist, λ_0 is the center wavelength of the OCT system, and n is the refractive index of the media. The sensitivity roll off function is defined by

$$H(z) = \text{sinc}^2(.5\Delta k_{samp}) \exp\left(-\frac{\Delta k_{opt}^2 z^2}{8 \log 2}\right)$$

is a depth dependent function of wave number where Δk_{samp} is the spectral resolution and Δk_{opt} is determined by the dispersion line width of the spectrometer.

The model used in this paper assumes all of the attenuation comes from scattering events, the frequency of which depend on material properties. However, because a significant amount of signal attenuation comes from the sensitivity roll off and confocal PSF functions applying our model directly to an uncalibrated signal will result in a large overestimation of true material dependent attenuation. Therefore, to use the DR model presented in this paper these functions must be measured and the resulting data must be corrected as

$$I(z; \mu_{B,NA}, \mu_{oct}) = \frac{I_T}{H(z)T(z - z_f)}. \tag{A.11}$$

This formula is valid such that $sin^2(.5\Delta k_{samp})$ is non-zero, although for small spectral resolutions this could lead to numerical instability.

APPENDIX B

Appendix B: Holography Derivations

B.1. Angular Spectrum Method

Steady state light propagation in space or through an aperture is well described by Huygen's principle. This states that the wavefront can be considered as the superposition of outgoing spherical waves originating at every point. Mathematically, this is often written as the Fresnel diffraction integral [**59**]

$$E(x,y,z) = \frac{-ik}{\pi}\int_{R^2} dx_0 dy_0 E(x_0,y_0,0)\frac{\exp(ikr)}{r} \tag{B.1}$$

where $E(x_0, y_0, 0)$ is the known 'input' electric field in a system and the distance $r =$

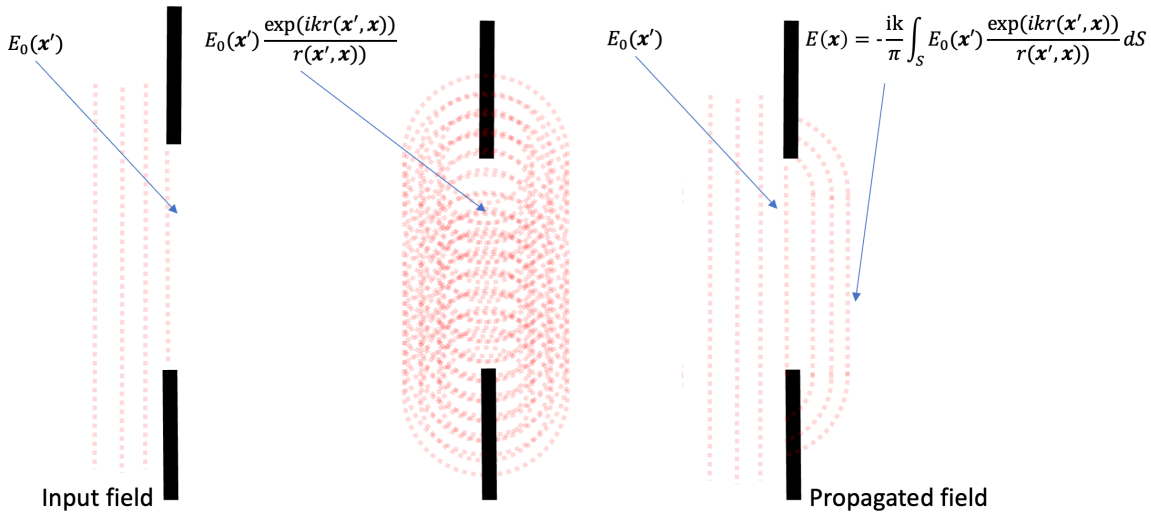


Figure B.1. Figure demonstrating visually how Huygen's principle can be used to compute diffraction through an aperture.

$\sqrt{(x-x_0)^2+(y-y_0)^2+z^2}$. An alternative way to exploit the linearity of this system

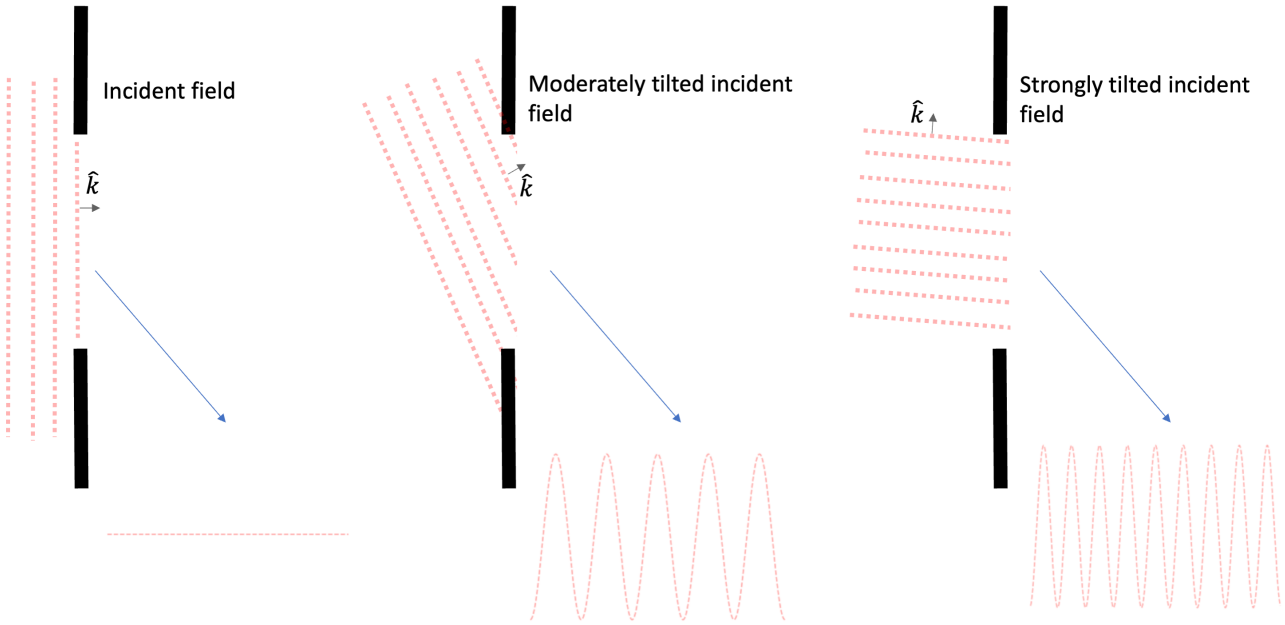


Figure B.2. Figure demonstrating visually how tilting an incident plane wave and evaluating on an aperture can be equivalently viewed as a field with variable frequency present on the aperture.

is to expand the incident field into a series of tilted plane waves (as opposed to spherical) which have a well defined direction of propagation. This can be done by Fourier transforming the incident field with respect to x and y as

$$\mathcal{A}_0(k_x, k_y) = F[E(x, y, 0)], \tag{B.2}$$

where $\mathcal{A}_0$ is referred to as the *angular spectrum* of the input fields. The plane defined by $(x, y, 0)$ is fixed in space, and the original wavelength of light determines the frequency of the incident wave. There is an equivalence between higher spatial frequencies and tilted plane waves as shown in Figure B.2. If a plane wave of a fixed frequency is tilted and

then evaluated on an aperture it is equivalent to just having a higher frequency incident wave. The equivalent tilted plane wave would have a propagation direction which satisfies $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. A plane wave propagating with direction (k_x, k_y, k_z) between plane $z_0 = 0$ and $z_0 = z$ will simply pick up a phase term of the form $\exp(ik_z z)$ meaning the propagated field can be computed as the superposition of all the incident plane waves multiplied by the phase factor

$$E(x, y, z) = \int_{R^2} dk_x dk_y \mathcal{A}_0(k_x, k_y) \exp(ik_z z + k_x x + k_y y) \tag{B.3}$$

which is the inverse Fourier transform

$$E(x, y, z) = F^{-1}[\mathcal{A}_0(k_x, k_y) \exp(izk_z)] = F^{-1}[\mathcal{A}_0(k_x, k_y) \exp(iz\sqrt{k^2 - k_x^2 - k_y^2})]. \tag{B.4}$$

In the case of a digital holographic system we know we have a collimated laser beam as our input electric field. This is known to have a very small angular spectrum such that the first order Taylor expansion $k_z \approx k - \frac{k_x^2 + k_y^2}{2k}$ holds. Plugging this in we see that

$$E(x, y, z) = \exp(ikz) F^{-1}[F[E_0] \exp(iz \frac{-k_x^2 - k_y^2}{2k})]. \tag{B.5}$$

Turning our perspective over to computation we denote the FFT operator as $\boldsymbol{F}$ and the propagation operator as we will denote as $\boldsymbol{A}$ which is defined as

$$\boldsymbol{A}(z) := \exp\left(iz\left(K^2 - k_x^2 - k_y^2\right)\right) circ(\frac{k_x^2 + k_y^2}{\left(K^2 - k_x^2 - k_y^2\right)}), \tag{B.6}$$

where

$$circ(x) = \mathbb{I}_{||x||<1} \tag{B.7}$$

From this we can define the ASM operator as

$$\text{ASM}(E_0, z) = \boldsymbol{F}^{-1}\boldsymbol{A}(z)\boldsymbol{F}E_0. \tag{B.8}$$

B.2. Intensity Holograms

Consider an 'object' electric field denoted $O(x, y, z)$ and a reference field which is a plane wave denoted by $R(x, y, z) = \exp(-i\boldsymbol{k}_r \cdot \boldsymbol{x})$. The goal of holography is to encode the object field in a medium as an interference pattern and then reconstruct or *playback* the field later. The interference pattern on plane $z = 0$ is given by

$$H_I(x, y) = ||O + R||^2 = ||O||^2 + ||R||^2 + O^*R + R^*O \tag{B.9}$$

where $O = O(x, y, 0)$ and $R = R(x, y, 0)$. Assuming that this reference pattern in encoded in the transmission of a media the media can be then be illuminated with a second plane wave R' yielding the electric field

$$R'H_I = R'(||O||^2 + ||R||^2) + O^*RR' + R'R^*O. \tag{B.10}$$

Where $R'RO = exp(i(k - k_r) \cdot x)O$ and $R'RO^* = exp(i(k + k_r) \cdot x)O^*$ which are two twin scaled copies of the original object field evaluated at the plane of the interference pattern. To see what the total optical field looks like in space we can simply apply the ASM kernel and propagate the optical field to a new depth z. Ignoring the first two 'DC terms' we

have

$$ASM(H_I, z) - DC = \boldsymbol{F}^{-1}\boldsymbol{A}\boldsymbol{F}[RR'O^*] + \boldsymbol{F}^{-1}\boldsymbol{A}\boldsymbol{F}[R'R^*O] \tag{B.11}$$

where $DC = R'(||O||^2 + ||R||^2)$. The convolution theorem gives

$$\boldsymbol{F}[RR'O^*] = \boldsymbol{F}[RR'] * \boldsymbol{F}[O^*] = \delta(k - k_r) * \boldsymbol{F}[O^*] = \boldsymbol{F}[O^*](k - k_r) \tag{B.12}$$

$$\boldsymbol{F}[R'R^*O] = \boldsymbol{F}[R^*R'] * \boldsymbol{F}[O] = \delta(k + k_r) * \boldsymbol{F}[O] = \boldsymbol{F}[O](k + k_r) \tag{B.13}$$

which allows us to see that

$$ASM(R'H_I, z) - DC = (RR')O(x, y, z) + (R^*R')O^*(x, y, z) \tag{B.14}$$

which are just two twin copies of the original object field scaled by the interference pattern of the reference and playback waves.

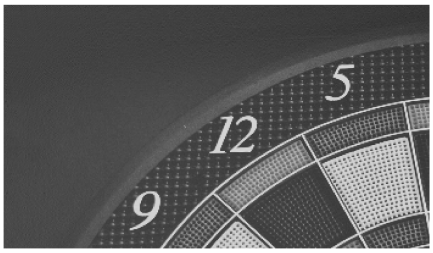
B.3. TI DLP Phase Level Table

The following are the provided phase levels for the TI DLP.

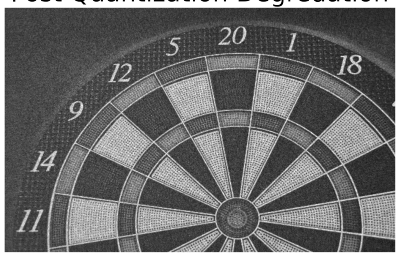
Phase Level	Percent Delay	Phase Delay
1	0	-3.1416
2	0.0100	-3.0788
3	0.0205	-3.0128
4	0.0422	-2.8764
5	0.0560	-2.7897
6	0.0727	-2.6848
7	0.1131	-2.4310
8	0.1734	-2.0521
9	0.3426	-0.9890
10	0.3707	-0.8124
11	0.4228	-0.4851
12	0.4916	-0.0528
13	0.5994	0.6245
14	0.6671	1.0499
15	0.7970	1.8661
16	0.9375	2.7489
17	1	3.1416

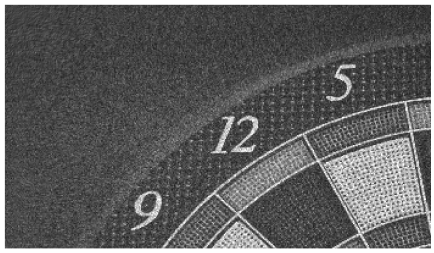
B.4. Example Results

Full Bit Depth 2 Frame SGD



Post Quantization Degredation





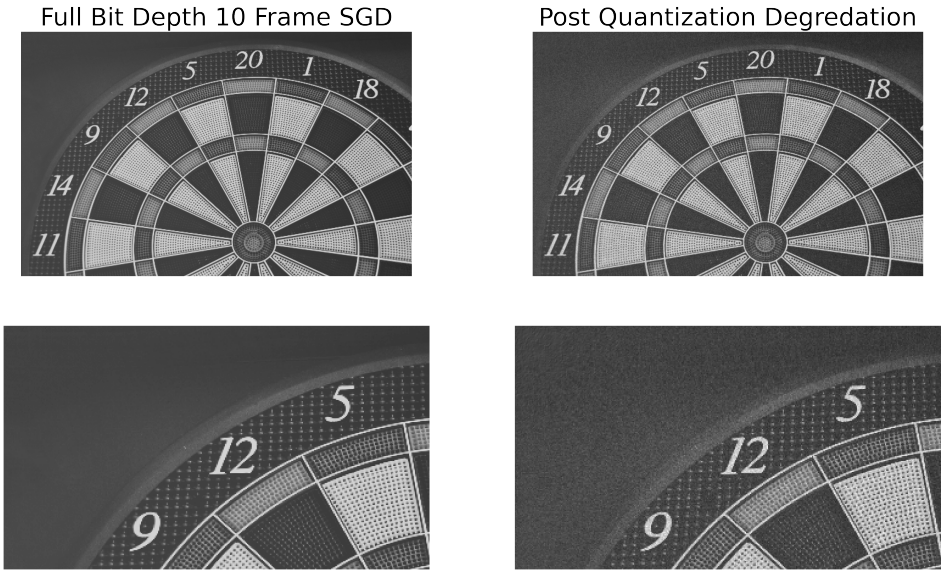
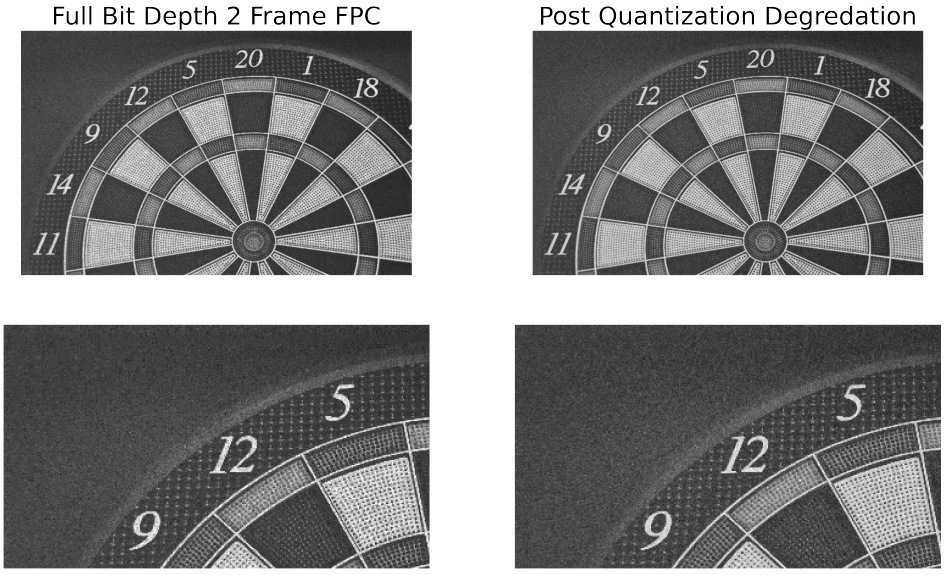


Figure B.3. Results for 2 and 10 frame SGD before and after quantization.



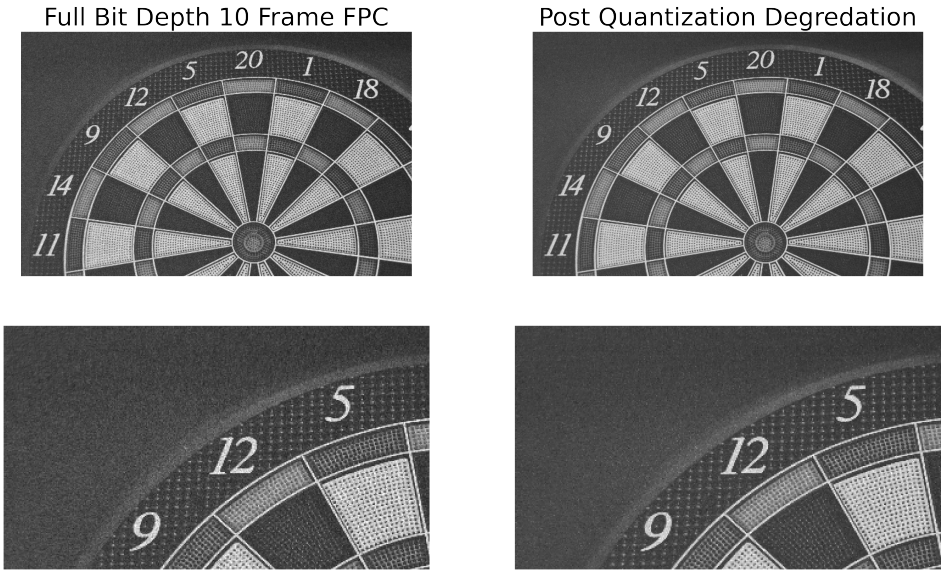
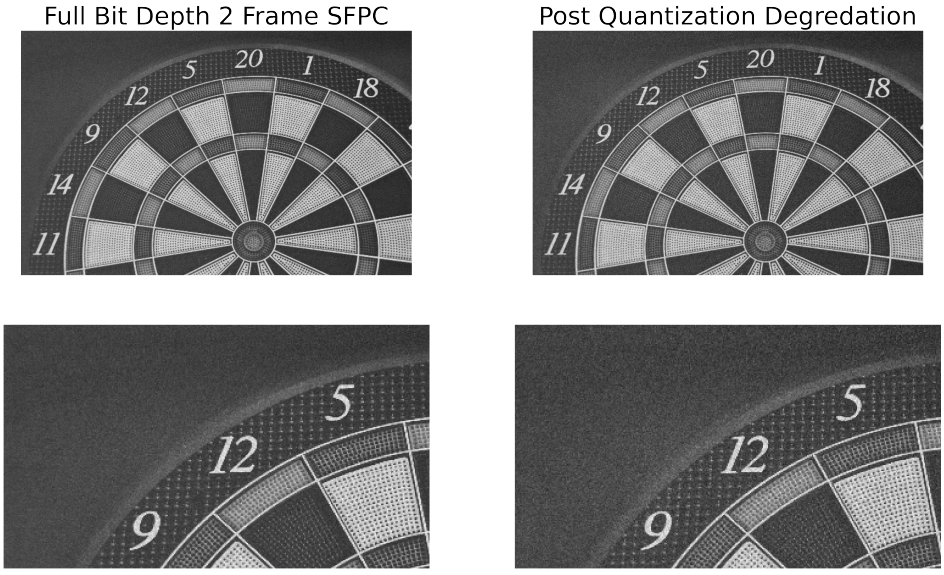


Figure B.4. Results for 2 and 10 frame FPC before and after quantization.



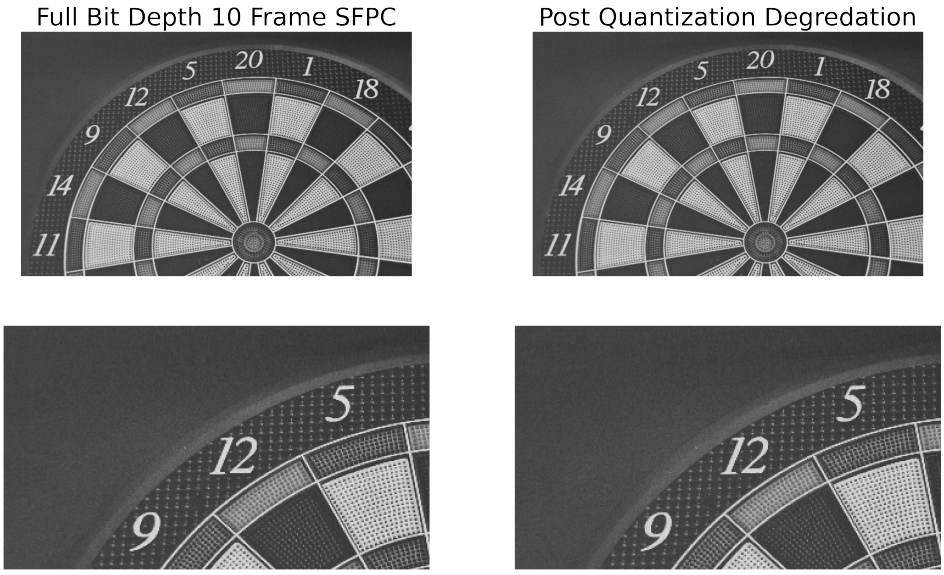
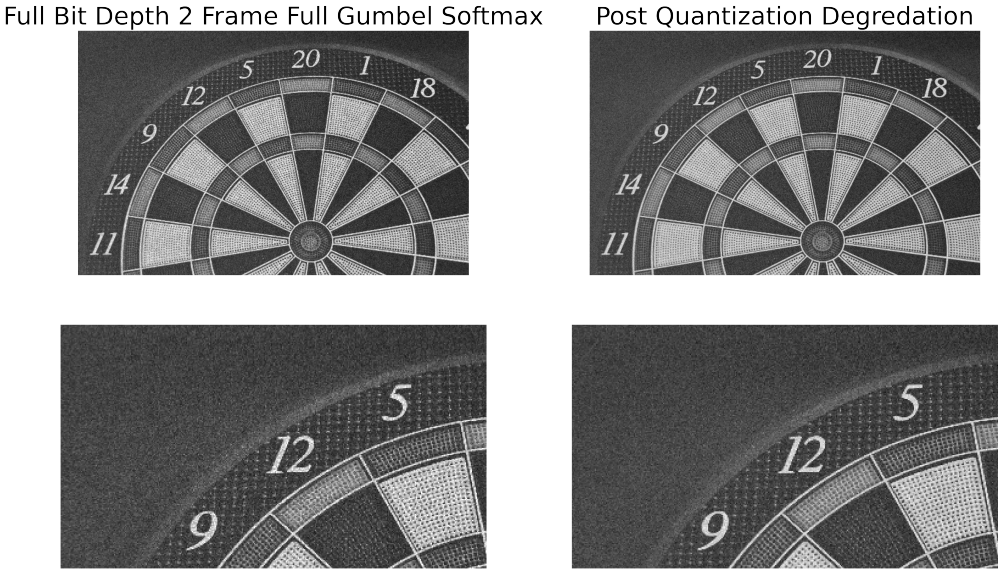


Figure B.5. Results for 2 and 10 frame SFPC before and after quantization.



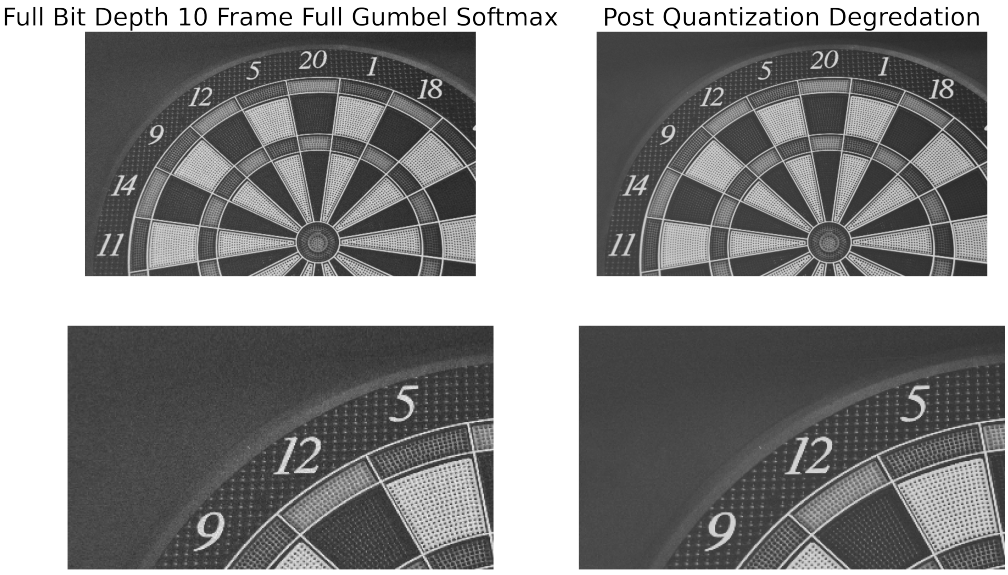
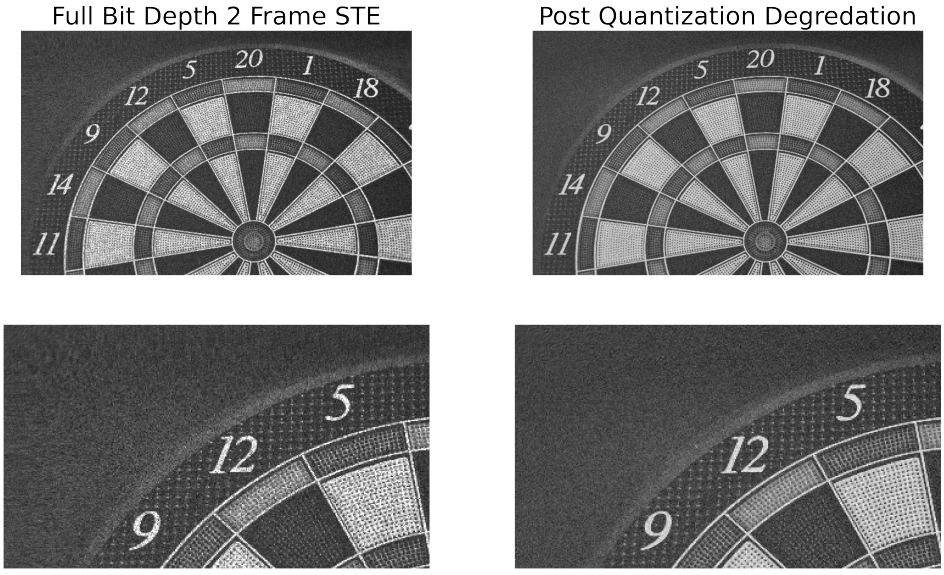


Figure B.6. Results for 2 and 10 frame Gubell Softmax before and after quantization.



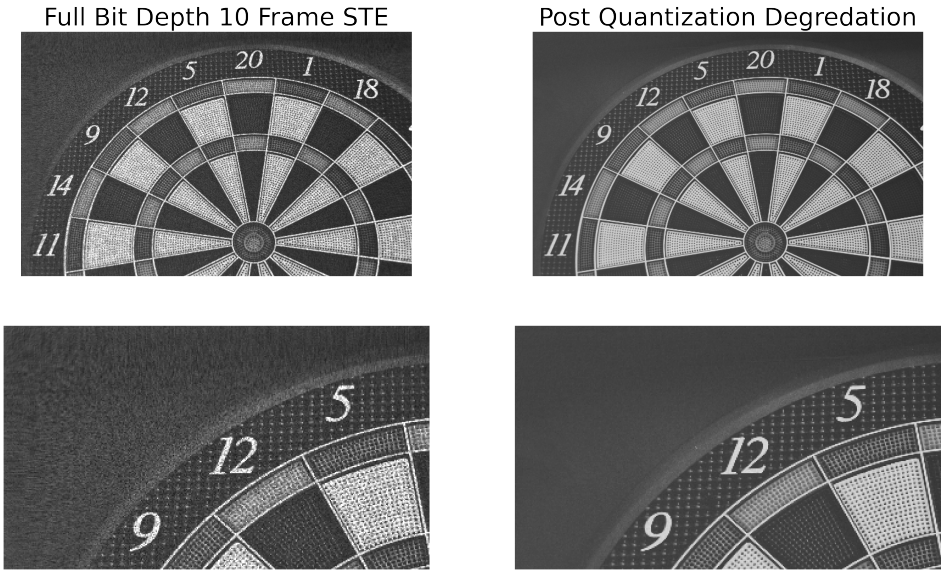
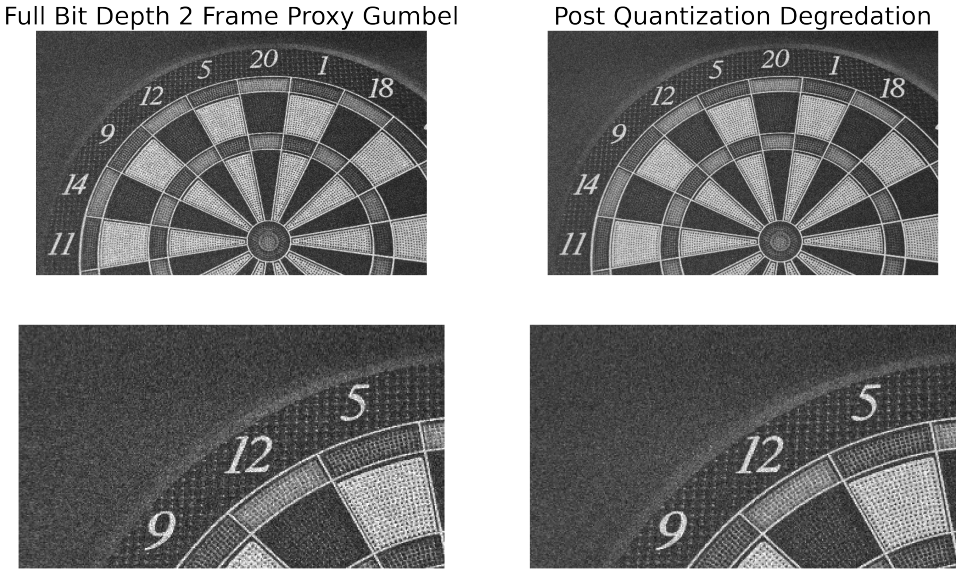


Figure B.7. Results for 2 and 10 frame STE before and after quantization.



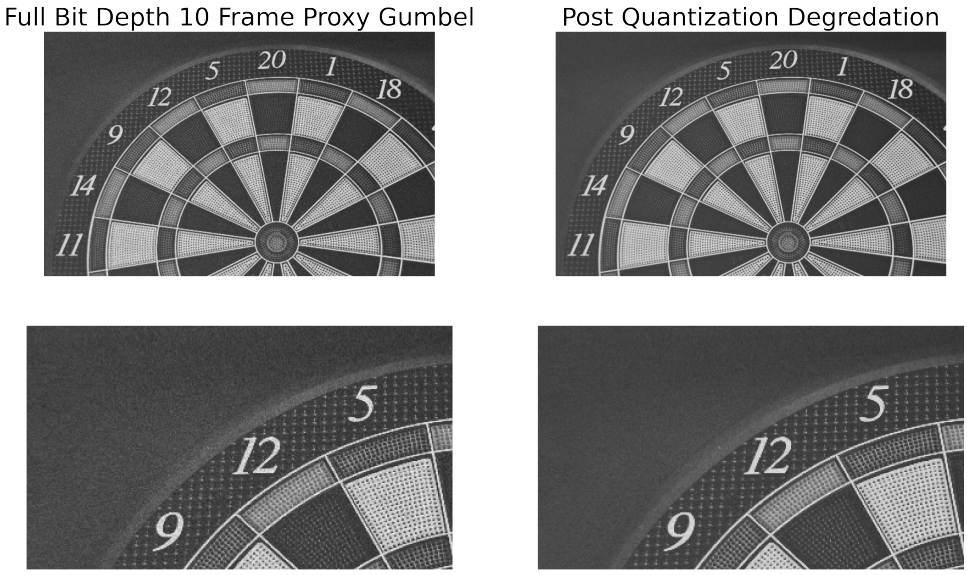


Figure B.8. Results for 2 and 10 frame Proxy Gumbel before and after quantization.

APPENDIX C

Appendix C: Psuedo-code and Algorithms

C.1. XRF Delayering Incorporating RIS

Inputs:

- XRF: XRF elemental map
- R: RIS data cube
- N_c, N_x, N_y: Number of clusters and pixels in x and y directions
- $Shifts$: Pixel shift directions and distance

Output:

- S, B:Two Delayered XRF elemental Maps (Surface and below Surface)

Register R to XRF image:

$R \leftarrow \text{Sift}(R; XRF)$

Cluster R:

$R \leftarrow \text{K\_means}(R; N_c)$

Construct Surface Image Stack:

$S \leftarrow 0$

for $k \leq N_c$ **do**

  $m_k = \text{mean}_{ij}(XRF \ * \ \mathbb{I}_{i,j \in \ cluster \ k})$

  for $i, j \leq N_X, N_y$ **do**

    $S_{ij} = S_{ij} + m_k \ * \ \mathbb{I}_{i,j \in \ cluster \ k}$

  end

end

for $sh \ \in \ \text{len(Shifts)}$ **do**

  $(s_i, s_j) = Shifts(sh)$

  $S^{sh}_{ij} = S_{i+s_i, j+s_j}$

end

Compute Subsurface Intensity:

$B_{ij} \leftarrow min_{sh}(XRF_{ij} - S^{sh}_{ij})$

Return: S, B

Algorithm 2: Psuedo-code for the Surface and Subsurface XRF Delayering algorithm

C.2. Holography Algorithms

C.2.1. A note on automatic differentiation

In this section we discuss the implementations of the algorithms laid out in chapter 3. As is quickly becoming standard, all algorithms were implemented using a scientific computing package which has automatic differentiation capabilities, specifically Pytorch. In languages with automatic-differentiation, when a new elementary function is written it must be written with both the function value called the forward operation as well as it's derivative evaluated at the input called the backward function. The forward value is returned and the backwards value is stored in memory. When a series of these elementary functions are called in sequence, the precise order of the function calls as well as their gradient values are stored in a graph structure called the *computational graph*. Since the vast majority of functions are computed using elementary functions the backward function does not need to be explicitly defined unless an inexact gradient is needed for some reason such as in the case of proxy gradient methods.

When doing optimization for parameter recovery, a loss function is computed. This loss function is a mapping from the parameters to a positive valued scalar function which describes how well the model agrees with the data. Normally, the goal is to minimize this loss function using a gradient based method to find the variables which fit the data best. When the loss function is computed in a language which has automatic differentiation capabilities, the gradients of the loss function can be efficiently computer by multiplying together the stored backwards values all the way down the computational graph. This will, under most circumstances have the same computational complexity as the forward

pass making it a computationally cheap method. All psuedo-code in this appendix is written assuming the gradients are being computed in a modern autograd language such as pytorch. In this psuedocode we adopt the convention that '.backward()' backpropagates the gradients through the computational graph.

C.2.2. SGD

Inputs:

- z, λ, dx: Propagation distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern stack, quantization level look up table
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max SGD iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Propagation Object:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$
$K \leftarrow 1/\lambda$
$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$
def $ASM(x; H)$:
$forward(x)$: **return** $conv(x, A)$

while it $< max\_its$ or $er > tol$ **do**
$Field^f = Input\_Field \ * \ \exp(-1jP^f)$
Intensity_Output $= \sum_{f \leq N_f} ||ASM(Field^f, A)||^2$
Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$
$grad =$ Loss.backward()
$P^f_{ij} \leftarrow P^f_{ij} + lr \ grad^f_{ij}$
$it = it + 1$
end

Wrap Phase Pattern and Quantize:

$P^f_{ij} = atan(Imag(P^f_{ij})/Real(P^f_{ij}))$
$P^f_{ij} \leftarrow min_{q \in \text{ Look up table}}(P^f_{ij} - q)^2$
Return: P

Algorithm 3: Psuedo-code for the Stochastic Gradient Descent Algorithm for Computational Holography.

C.2.3. STE

The straight through estimation algorithm looks exactly the same as the SGD algorithm defined in ap. C.2.2 with the addition of quantization and which has a backwards function which lets the gradients pass through unchanged.

Inputs:

- z, λ, dx: Propagation Distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern, quantization level look up table
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max STE iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Objects:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$
$K \leftarrow 1/\lambda$
$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$
def $ASM(x; A)$:
  $forward(x)$: **return** $conv(x, A)$
def $diff\_quant(x; LUT)$:
  $forward(x)$: **return** $min_{q\in \text{ Look up table}}((x_{ij}) - q)^2$
  $backward(x, grad\_in)$: **return** $grad\_in$
while it $< max\_its$ or $er > tol$ **do**
  $P_{quant} = diff\_quant(P, LUT)$ $Field^f = Input\_Field * \exp(-1jP^f_{quant})$
  Intensity_Output $= \sum_{f\leq N_f} ||ASM(Field^f, A)||^2$
  Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$
  $grad =$ Loss.backward()
  $P^f_{ij} \leftarrow P^f_{ij} + lr\ grad^f_{ij}$
  $it = it + 1$
end
Wrap Phase Pattern and Quantize:
$P^f_{ij} = atan(Imag(P^f_{ij})/Real(P^f_{ij}))$
$P^f_{ij} \leftarrow min_{q\in \text{ Look up table}}(P^f_{ij} - q)^2$
Return: P

Algorithm 4: Psuedo-code for the Straight Through Estimation Algorithm for Computational Holography.

C.2.4. Framewise PC

The Framewise PC algorithm is a version of SGD where some frames are quantized and frozen during the optimization process allowing quantization information to flow into

the problem. In this algorithm superscripts refer to the frame. The derivative of the quantization operator used here is zero so the derivative for the entire phase pattern stack will only be non-zero for the frames which are unquantized.

Inputs:

- z, λ, dx: Propagation Distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern, quantization level look up table
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max FPC iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Objects:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$
$K \leftarrow 1/\lambda$
$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$
def $ASM(x; A)$:
  $forward(x)$: **return** $conv(x, A)$
def $quant(x; LUT)$:
  $forward(x)$: **return** $min_{q\in \text{ Look up table}}((P_{ij}) - q)^2$
while it $< max\_its$ or $er > tol$ **do**
  for $f < N_f$ **do**
    $P^{fr\neq f}_{quant} \leftarrow quant(P^{fr}; LUT)$
    $P^{f}_{quant} \leftarrow P^{f}$
    $Field^f r = Input\_Field \ * \ \exp(-1jP^f)$
    Intensity_Output $= \sum_{fr\leq N_f} ||ASM(Field^f r, A)||^2$
    Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$
    $grad =$ Loss.backward()
    $P^f_{ij} \leftarrow P^f_{ij} + lr \ grad^f_{ij}$
  end
  $it = it + 1$
end
Wrap Phase Pattern and Quantize:
$P^f_{ij} = atan(Imag(P^f_{ij})/Real(P^f_{ij}))$
$P^f_{ij} \leftarrow min_{q\in \text{ Look up table}}(P^f_{ij} - q)^2$
Return: P

Algorithm 5: Psuedo-code for the Framewise Predictor Corrector Algorithm for Computational Holography.

C.2.5. Psuedo-code for the Single Frame Predictor Corrector Algorithm for Computational Holography.

The SFPC algorithm is a version of SGD where odd and even rows are updated independently allow for partial quantization.

Inputs:

- z, λ, dx: Propagation Distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern, quantization level look up table
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max SFPC iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Objects:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$

$K \leftarrow 1/\lambda$

$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$

def $ASM(x; A)$:

$forward(x)$: **return** $conv(x, A)$

def $quant(x; LUT)$:

$forward(x)$: **return** $min_{q\in \text{ Look up table}}((P_{ij}) - q)^2$

while it $< max\_its$ or $er > tol$ **do**

Odd update

$(P_{quant})_{i,j \ s.t. \ i \ odd} \leftarrow P_{i,j}$

$(P_{quant})_{i,j \ s.t. \ i \ even} \leftarrow quant(P_{i,j}; LUT)$

$Field^{fr} = Input\_Field \ * \ \exp(-1jP^{fr})$

Intensity_Output $= \sum_{fr\leq N_f} ||ASM(Field^{fr}, A)||^2$

Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$

$grad =$ Loss.backward()

$P^{fr \ s.t. \ fr \ odd} \leftarrow P_{ij}^{fr} + lr \ grad_{ij}^{fr}$

Even update

$(P_{quant})_{i,j \ s.t. \ i \ even} \leftarrow P_{i,j}$

$(P_{quant})_{i,j \ s.t. \ i \ odd} \leftarrow quant(P_{i,j}; LUT)$

$Field^{fr} = Input\_Field \ * \ \exp(-1jP^{fr})$

Intensity_Output $= \sum_{fr\leq N_f} ||ASM(Field^{fr}, A)||^2$

Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$

$grad =$ Loss.backward()

$P^{fr \ s.t. \ fr \ odd} \leftarrow P_{ij}^{fr} + lr \ grad_{ij}^{fr}$

$it = it + 1$

end

Wrap Phase Pattern and Quantize:
$P_{ij}^f = atan(Imag(P_{ij}^f)/Real(P_{ij}^f))$
$P_{ij}^f \leftarrow min_{q\in \text{ Look up table}}(P_{ij}^f - q)^2$
Return: P

Algorithm 6: Psuedo-code for the Single Frame Predictor Corrector Algorithm for Computational Holography.

C.2.6. Gumbel-Softmax

The Gumbel-Softmax method relies on a technique known as categorical reparameterization to efficiently sample the space of SLM patterns to find an optima. As the space is sampled the control pattern which controls the mean of the sampling process converges to a suitable SLM pattern to be displayed.

Inputs:

- z, λ, dx: Propagation Distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern, quantization level look up table
- w, t_0, t_{rate}: Sigmoid width , Starting temperature, Annealing rate
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max GS iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Objects:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$
$K \leftarrow 1/\lambda$, $t \leftarrow t_0$
$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$
def $ASM(x; A)$:
  $forward(x)$: **return** $conv(x, A)$
def $gum\_quant(x; LUT, w, t)$:
  $forward(x)$:
    $x_j \leftarrow Sigmoid(w(x_j - LUT_j))$
    $x_j \leftarrow x_j(1 - x_j) + Gumbel(0, 1)$
    $x \leftarrow Softmax(x, t)$
    $x_{quant} = \sum_j LUT_j x_j$
    Return: x_{quant}

while it $< max\_its$ or $er > tol$ **do**
  $P_{quant} = gum\_quant(P_{quant}; LUT, w, t)$ $Field^f = Input\_Field * \exp(-1jP^f)$
  Intensity_Output $= \sum_{f \le N_f} ||ASM(Field^f, A)||^2$
  Loss $=$ (Intensity_Output $- T)^2/N_x/N_y$
  $grad =$ Loss.backward()
  $P^f_{ij} \leftarrow P^f_{ij} + lr\ grad^f_{ij}$ $t \leftarrow t_0 \exp\left(-\frac{it}{max\_its}\log(t_{rate})\right)$
  $it = it + 1$
end

Wrap Phase Pattern and Quantize:

$P^f_{ij} = atan(Imag(P^f_{ij})/Real(P^f_{ij}))$
$P^f_{ij} \leftarrow min_{q \in \text{ Look up table}}(P^f_{ij} - q)^2$
Return: P

Algorithm 7: Psuedo-code for the Gumbel-Softmax Algorithm for Computational Holography

C.2.7. Proxy Gumbel

The Proxy Gumbel method is a fusion of the random gradient from the GS method with the deterministic quantization of a standard STE method. For more complex proxy methods it is necessary to edit the computational graph directly to independently alter either the gradient or the function value at that point used to compute the gradient. In this algorithm two forward models are calculated using different quantization methods. The proxy gradient can be computed with back propagation by replacing the function value but not the stored gradient information. in the syntax of pytorch we can access this by using the ".data" field in the tensor.

Inputs:

- z, λ, dx: Propagation Distance, Wavelength, Pixel size
- N_f, N_x, N_y: Number of frames and pixels in x and y directions (size of P)
- T, P, LUT: Desired output intensity, initial phase pattern, quantization level look up table
- w, t_0, t_{rate}: Sigmoid width , Starting temperature, Annealing rate
- $Input\_Field$: Desired output intensity
- tol, max_its, lr: Error tolerance, Max PGS iterations, Learning rate

Output:

- P: SLM phase delay pattern to display image

Define Objects:

$kx, ky \leftarrow meshgrid(linspace(0, 2\pi/dx), linspace(0, 2\pi/dx))$
$K \leftarrow 1/\lambda$, $t \leftarrow t_0$
$A \leftarrow \exp\left(1iz\frac{-(kx^2+ky^2)}{K}\right) * circ(\frac{kx^2+ky^2}{K^2-kx^2+ky^2})$
def $ASM(x; A)$:
  $forward(x)$: **return** $conv(x, A)$
def $gum\_quant(x; LUT, w, t)$:
  $forward(x)$:
    $x_j \leftarrow Sigmoid(w(x_j - LUT_j))$
    $x_j \leftarrow x_j(1 - x_j) + Gumbel(0, 1)$
    $x \leftarrow Softmax(x, t)$
    $x_{quant} = \sum_j LUT_j x_j$
    Return: x_{quant}
def $quant(x; LUT)$:
  $forward(x)$: **return** $min_{q\in \text{ Look up table}}((x_{ij}) - q)^2$
while it $< max\_its$ or $er > tol$ **do**
  $P_{gum\_quant} = gum\_quant(P; LUT, w, t)$
  $P_{quant} = quant(P; LUT)$
  $Field^f_{gum\_quant} = Input\_Field * \exp(-1jP^f_{gum\_quant})$
  $Field^f = Input\_Field * \exp(-1jP^f_{quant})$
  $\text{Intensity\_Output}_{gum\_quant} = \sum_{f\le N_f} ||ASM(Field^f, A)||^2$
  $\text{Intensity\_Output} = \sum_{f\le N_f} ||ASM(Field^f, A)||^2$
  $\text{Intensity\_Output}_{gum\_quant}.data = \text{Intensity\_Output}.data$
  $\text{Loss} = (\text{Intensity\_Output}_{gum\_quant} - T)^2/N_x/N_y$
  $grad = \text{Loss.backward()}$
  $P^f_{ij} \leftarrow P^f_{ij} + lr\ grad^f_{ij}$
  $t \leftarrow t_0 \exp\left(-\frac{it}{max\_its}\log(t_{rate})\right)$
  $it = it + 1$
end

Wrap Phase Pattern and Quantize:
$P_{ij}^f = atan(Imag(P_{ij}^f)/Real(P_{ij}^f))$
$P_{ij}^f \leftarrow min_{q \in \text{ Look up table}}(P_{ij}^f - q)^2$
Return: P

Algorithm 8: Psuedo-code for the Proxy Gumbel Algorithm for Computational Holography.