Physics-based Computational Modeling of Human Skin using Machine Learning for Physiological Parameter Estimation

by

Saurabh Vyas

A thesis submitted to The Johns Hopkins University in conformity with the requirements for the degree of Master of Science.

Baltimore, Maryland

May, 2014

Abstract

The skin in the largest organ in the human body and often subject to the greatest exposure to outside elements throughout one's lifetime. Current data by the World Health Organization suggests that more than 10 people die each hour worldwide due to skin related conditions. Many of these conditions include cancers, such as melanoma, which are growths that originate in the epidermis and if left untreated can spread throughout the body, reducing the chances of survival to less than 1%. If these tumors are detected during the early stages, the chances of survival are over 99%. Unfortunately, there only exist coarse diagnostic metrics, such as evaluations of color, texture, boundaries, and asymmetry, which are not sufficient for early detection of these cancers.

In order to develop a screening technology, we require a non-invasive means of measuring the various biological components that make up the layers of the skin, i.e., melanosome concentration, collagen concentration and blood oxygen saturation, amongst others. The temporal analysis of changes in these components can serve as a critical tool in diagnosing the progression of these malignant cancers and in

understanding the pathophysiology of cancerous tumors. Quantitative knowledge of these parameters can also be useful in applications such as wound assessment, drug delivery, and point-of-care diagnostics, amongst others.

From a systems level perspective, we seek to develop a non-invasive, non-ionizing, and rapid technology that exposes an afflicted area on the skin to light, measures the amount of light that is reflected, transmitted, and/or absorbed, and using this information infers the concentration of each of the materials that make up the skin. Naturally, this inference would require *a priori* knowledge about the relationship between reflected light and concentration of biological materials. This is the goal of this thesis, the development of a computational model that relates the concentration of biological skin materials to a light reflectance measurement from the surface of the skin. This light reflectance measurement is obtained using hyperspectral imaging (HSI) or reflectance spectrometry. HSI allows for imaging well beyond the visible (VIS) region of the electromagnetic spectrum; past the near-infrared (NIR) and through the short wave infrared (SWIR). HSI allows us to obtain a reflectance measurement for each wavelength (band) spanning from 400 nm (VIS) to 1800 nm (SWIR). Imaging past the VIS can capture characteristic absorptions and other physiological makers typically exhibited by skin components outside the VIS region.

In this thesis, we developed a method to estimate human skin parameters, such as melanosome concentration, collagen concentration, oxygen saturation, skin thickness, and blood volume, using hyperspectral radiometric measurements (signatures)

obtained from *in vivo* skin. We developed a computational model based on *Kubelka-Munk* theory and the Fresnel equations. This model generates a forward mapping (a transformation) between skin parameters and a corresponding HSI reflectance spectra. This is a complex model, and not invertible. Therefore, we used machine learning based regression to generate the inverse mapping (the inverse transformation) between skin parameters and hyperspectral signatures. This yields a transformation (i.e., an inverse transformation) between the skin parameter vector space and the HSI signature vector space. Simply put, using a reflectance signature from a patch of skin, we can estimate the concentration of the biological materials that make up that patch of skin.

Another challenge in the field has been that of obtaining ground truth. Methods to estimate skin parameters have been developed by several other studies, but no group has yet to compare their method to actual ground truth. Therefore, there is no direct way to assess the accuracy of the parameter estimation method. A major reason for this has to do with the practical difficulty associated with obtaining this ground truth; it involves biopsies and further biochemical analysis by a pathologist. For some parameters (e.g., melanosome concentration) it is unclear how one would proceed with determining the true concentration. For one skin parameter of dermatological interest, epidermal and dermal thickness, we developed a methodology based on Ultrasound imaging (US) to obtain a proxy ground truth against which to benchmark our machine learning method. For the first time, this provided a direct

validation of the performance of the estimation methodology.

We tested our methods using synthetic and *in vivo* skin signatures obtained in the VIS through the SWIR domains from 24 patients of both genders and Caucasian, Asian, and African American ethnicities acquired under IRB approval at the Johns Hopkins Hospital. Performance validation showed promising results: good agreement with the ground truth (average absolute error of 0.05±10e-3 percent) and well-established physiological precepts, as well as strong agreement with the gold standard obtained from Ultrasound imaging (mean error of 0.09±0.05 mm). Our early results suggested that our methods have potential use in the characterization of skin abnormalities and in non-invasive pre-screening of malignant skin cancers.

Thesis Committee:

Professor Philippe M. Burlina

Professor Jeffrey H. Siewerdsen

Professor Jon Meyerle

Acknowledgments

The most important lesson I learned as a Master's student is that research takes time; it's rarely easy or straightforward. I owe much of this lesson, and the success I've had with my research to my advisor Dr. Philippe Burlina. I want to thank him for always being right next door (literally) to answer any question or doubt I may have, or to help brainstorm a solution to a particular problem. Most importantly, I want to express my deepest gratitude to Phil for being one of the nicest and most approachable people I have ever had the pleasure of working with.

I also want to thank everyone in my section at APL whom I have had the opportunity to work with over the past few years – Amit, Daniel, Ryan, Nathan, Dean – their insights, comments, and excitement about science and technology, has made working at APL something I look forward to everyday.

This thesis would not have been possible without the help of Dr. Luis Garza, and Dr. Sewon Kang. I'm grateful to them for allowing me to shadow them in their clinic and interact with their patients at the Johns Hopkins Hospital, Department of Dermatology. Not only did this yield the unique datasets we analyzed in this thesis,

but also gave a humanizing perspective to the work we were trying to do, and the impact it could someday have.

I also want to thank my thesis committee, and the various faculty members and peers at Hopkins whom I have had the opportunity to work and grow with over the past six years.

I can never forget the wonderful friendships I have made at Hopkins; our late night discussions about seemingly unrelated things has often led to some deeper insight into not only my work, but also other aspects of my life.

Last, but never the least, I want to thank my parents for their unwavering support of my academic endeavors – I wouldn't be where I am without their love, support, and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The American Academy of Dermatology reported approximately 114,900 new cases of melanoma in the year 2010 alone of which approximately 8,700 resulted in death [5–7]. The rate of other common cancers has dropped in the last decade; however, the rate of melanoma is rising faster than that of the seven most common cancers [8, 9]. Though melanoma only accounts for less than 5% of skin cancer cases per year, it accounts for over 75% of skin cancer related deaths [8, 10, 11]. In fact, it is estimated that one person dies of melanoma every hour [5]. Melanoma is a malignant tumor of the melanocytes—cells responsible for producing the light-absorbing pigment called melanin. Melanin is responsible for skin-color and also helps protect against UV-light from the sun. The early stages of melanoma are characterized by proliferation of melanocytes between the outer layer (epidermis) and the layer beneath (dermis). If detected at this stage, the melanocytic tumor has a depth of less

(a)

(b)

Figure 1.1: Photographs of (a) a late stage melanoma; (b) melanocytic and non-melanocytic tumors. Courtesy of National Cancer Institute (AV-8500-3850) [1].

than 1 mm and can be excised with only local surgery and often has a very good prognosis. In many cases, however, the cancer is not detected at this stage and can spread around the body through lymph vessels leading to distant metastasis—spreading to parts of the body remote from the tumor's point of origin. Such deeper melanomas prompt sentinel lymph node studies to rule out the possibility of distant spread and have lower predicted survival. Currently, there exists no blood test for melanoma, and hence it can easily be missed in an annual physical exam during the early stages of the disease. The most common diagnostic for melanoma remains a visual examination by a health care professional [12]. This is a challenging task, even for seasoned dermatologists; Figure 1.1 (b) shows several examples of cancerous and

benign melanomas. Without any quantitative information, it is challenging for a dermatologist to accurately separate the cancerous tumors in Figure 1.1 (b), from the benign ones.

It is therefore no surprise that significant research and development is underway to design efficient and accurate non-invasive diagnostics tools, which can detect early signs of tissue abnormalities responsible for melanoma. The most sophisticated of such emerging technologies include: digital dermoscopy, videodermoscopy, and reflectance confocal microscopy [12]. These methodologies offer unique advantages but suffer from limitations due to traditional single band imaging. This poses a compromise between an accurate diagnosis with high spatial resolution versus degree of invasiveness and cost considerations. Such reasons, coupled with a steep learning curve and lack of easy integration, have prevented these technologies from being fully adapted into a large-scale clinical setting.

In the following sections, we present a novel approach that has the potential to aid in the development of a non-invasive technology for the pre-screening of cancers such as melanoma. Our methodology is a combination of a physics-based model of human skin, machine learning regression, and hyperspectral imaging. Each of these are detailed in the following sections. We also present a brief review of prior work in this research area. Finally, we conclude this chapter with a brief summary of the major objectives of this thesis.

1.1 Skin Optics

Skin optics is a research area rich in applications involving diagnosing, characterizing, and understanding the properties of tissues and organs. For over three decades, the area has attracted many research groups that have targeted a host of different applications. Several studies have attempted the task of developing tools to estimate the biological parameters that make up the layers of skin. This is an important task because quantitative knowledge of these parameters can be invaluable in applications such as medical diagnostics, wound-care, drug-delivery, and skin aging, amongst others. Many studies have cited [13–15] that skin parameters can be used in a longitudinal study to trace the growth and spread of skin cancers. Such longitudinal studies can provide guidance for clinicians to develop metrics to catch cancers such as melanoma during the early stages of the disease. These studies can also be used to understand the pathophysiology of malignant tumors, which could ultimately help in the development of better methods for drug delivery.

While this area has been investigated for decades, there is yet to be a robust technology that can aid in pre-screening for cancerous lesions. Some limiting factors have included: (1) cost of imaging systems, (2) the acquisition time, (3) expertise needed to operate the technologies, and (4) computational complexity of image analysis algorithms, amongst others. Recent advances in imaging technologies, coupled with next generation processing hardware and software has now mitigated many of these challenges. We posit that these changes have made computer aided diagnostics

a tractable reality, one which if translated into clinical workflow, can have an impact on patient care and outcomes.

1.2 Hyperspectral Imaging

Imaging sensors, or cameras, have become quite ubiquitous in modern day consumer electronics. One would be hard pressed to find a current cell phone without an integrated camera. Given the proliferation of these cameras, along with a parallel increase in computing power, there has been a boom in the image processing and computer vision communities to design methods to analyze and manipulate images and videos. Popular applications of this work have included Adobe Photoshop, Gimp, and the Xbox Kinect, amongst several others. However these are examples of image processing on data depicting visible light. There are also a number of applications in analyzing images acquired from other parts of the electromagnetic (EM) spectrum. As will be discussed later in this section, analyzing data from other parts of the EM spectrum is a key component of the work in this thesis.

1.2.1 Red-Green-Blue Imaging Principles

When we think of a camera we immediately think of a cell phone camera, or a digital point-and-shoot. Such cameras operate on a traditional RGB or Red-Green-Blue model to render a depiction of visible light. Typically an image is an $M \times N$

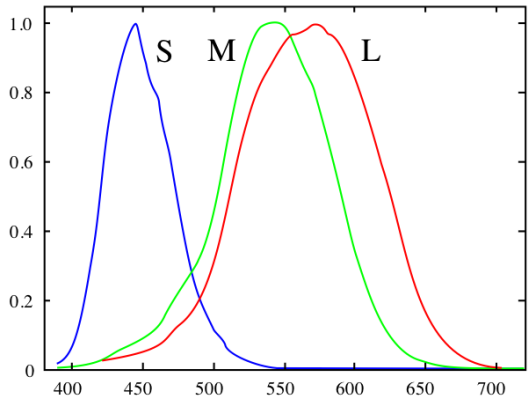


Figure 1.2: Normalized response for each type of cone as a function of wavelength. This picture was taken from [2] under the GNU Free Documentation License.

matrix or array, where each element (i.e., pixel) has an intensity (i.e., gray value) associated with it. An RGB image (i.e., a color image), as is found in a typical camera, is an $M \times N \times 3$ matrix, where the third dimension is known as a channel, or band, and refers to the red, green, and blue channel respectively. Therefore, each pixel has three values associated with it; a red channel value, a blue channel value, and a green channel value. Moreover each of these channels corresponds to a specific range of wavelengths, as opposed to a single wavelength per channel. This is an important distinction that will differentiate RGB imaging from hyperspectral imaging.

This RGB model provides a robust depiction of visible light and color because it was designed with the human physiology in mind. The human eye has light-sensitive photoreceptor cells called cones that are responsible for interpreting color. These cones come in three varieties; "L" cones respond to long wavelengths, "M" cones

respond to medium wavelengths, and "S" cones respond to short wavelengths. Figure 1.2 depicts the normalized response of each type of cone as a function of wavelength. Therefore, in the context of an RGB camera, we can think of each channel (RGB) as a response of each type of cone (LMS). For example, we know that each pixel of an RGB image has three values associated with it (i.e., R, G, and B). Each one of these values is obtained by integrating the response curves (the red curve for the red channel, the green curve for the green channel, and the blue curve for the blue channel) similar to the ones given in Figure 1.2. Therefore, each of the three values in the pixel of an RGB image, does not correspond to a particular wavelength of light, but rather an integral over a range of wavelengths.

1.2.2 Multi- and Hyper-spectral Imaging

Hyperspectral imaging (HSI), unlike RGB, (a) extends well past the visible region of the EM spectrum, and (b) divides the spectrum into a number of thinly spaced bands (i.e., wavelengths). To fully understand this phenomenon, we start with a discussion of what is known as multispectral imaging. Recall that if we had a grayscale image, each pixel would have a single value associated with it. If we had a color image, each pixel would be a vector, where each component would be an integrated response over the red, green, and blue channel (where a channel is a range of wavelengths). This type of image is called a multispectral image because each pixel has more than one spectral component, i.e., more than one wavelength of information. Imagine now

if each channel of an RGB image were allowed to have an additional channel, for a total of six channels. One way to achieve this would be to have each of the R, G, and B channels to only be an integral for half the range of wavelengths represented by that channel. In other words, each channel would only average half of each of the curves in Figure 1.2. Each pixel would then be a six-dimensional vector. We can now repeat this process, and in the limit every single wavelength would be individually represented, so that there is no need for averaging. This is known as hyperspectral imaging. Of course, HSI doesn't require every single wavelength to be individually represented, but the bands are often spaced with a very small step size (typically no more than 1-2 nm). Put another way, the number of bands in the image can be one way to differentiate a multispectral image from a hyperspectral image, though this is not set in stone. Figure 1.3 is a visual example of the difference between a HSI and a multispectral image.

1.2.3 HSI in Skin Optics

An HSI is often referred to as a cube because the data is a K-D matrix of the form $M \times N \times K$, where M and N are the spatial dimensions (x-y) and K is the number of bands. A typical plot for a HSI is shown on the right side of Figure 1.4. These plots are often referred to as HSI radiometric measurements, or signatures. It can then be inferred that these signatures are the true strength of this imaging modality. Notice in Figure 1.4 that the signature for each of the materials commonly found in nature

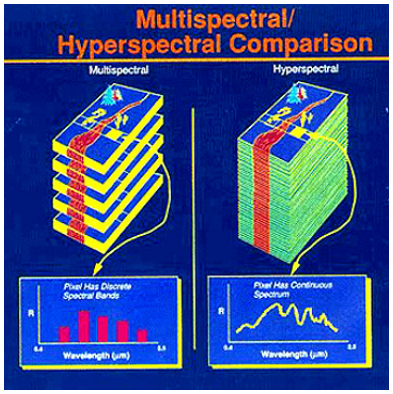


Figure 1.3: Difference between a hyperspectral and a multispectral image. This picture was taken from [3] courtesy of Dr. Nicholas M. Short, Sr. at NASA GSFC.

can be uniquely distinguished by their HSI signature. Suppose we only had an RGB image of a pixel of water; it would be very difficult to uniquely distinguish it from the sky or some other blue-ish material. Finally, note that each of the signatures is a depiction of a single pixel. In practice we would obtain 2-3 neighboring signatures of the same material and average them together (as a means to overcome the effects of noise).

A HSI signature is a function of wavelength, and is often expressed as a normalized relectance, i.e., the amount of light that was reflected from the material. Note that HSI extends past the visible region (approx. 750 nm) and through the NIR and the SWIR (past 2000 nm). Imaging at these high wavelengths allow for deeper penetration, and can often express the underlying characteristics of the material being imaged. For example, notice in Figure 1.4 the peaks and dips are very specific and characteristic of the materials being imaged.

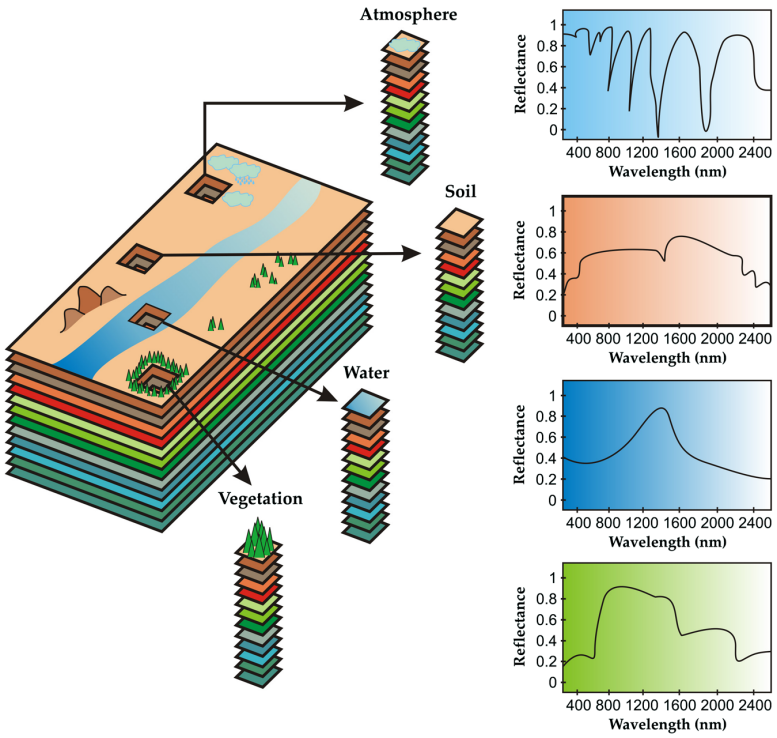


Figure 1.4: The HSI signature for different materials found in nature. This picture was taken from Javier Plaza (jplaza@unex.es) [4].

Given the nature of HSI to allow for differentiation of materials, it has typically been used for remote sensing and spectroscopic applications. However, HSI has recently received increased attention for this ability to detect and classify anomalous areas in a wide variety of materials, including the human skin [16]. A HSI system is able to measure specific spectral signatures based on the pigmentation and the color of human skin in the VIS through the SWIR (short wave infrared) regions of the EM spectrum [16]. Again referring to Figure 1.4, suppose our pixel was at the border of two materials (e.g., water and soil). We would then expect the estimated

signature to be a mixture of two different signatures. If there existed some *a priori* information about how the materials were mixed, or if we had a model for the estimated signature, which was parametrized by the concentration of soil and water, then we could recover the individual concentration of each material. The latter is the approach used in this work. HSI allows us to obtain signatures from patches of human skin; these patches are mixtures of biological materials (e.g., melanin, collagen, blood, etc.). This thesis develops a computational model of human skin, i.e., a model that expresses the relationship between the measured signature and the concentration of the biological materials that make up that signature. This model then allows us to estimate the concentration of each of the biological skin materials from the estimated HSI signature. An example of an HSI signature obtained from a patch of human skin is provided in Figure 1.5 for illustration purposes. Note that for our application, we are not interested in full images, but only a single pixel obtained from a patch of skin of clinical interest. This speeds up the acquisition and computation time to nearly real-time.

1.3 Computational Modeling

Computational modeling is the field of developing mathematical theories and simulations to study and analyze the behavior of physical, and in our case, physiological phenomenon. If we were interested in studying the mechanism of a biological phe-

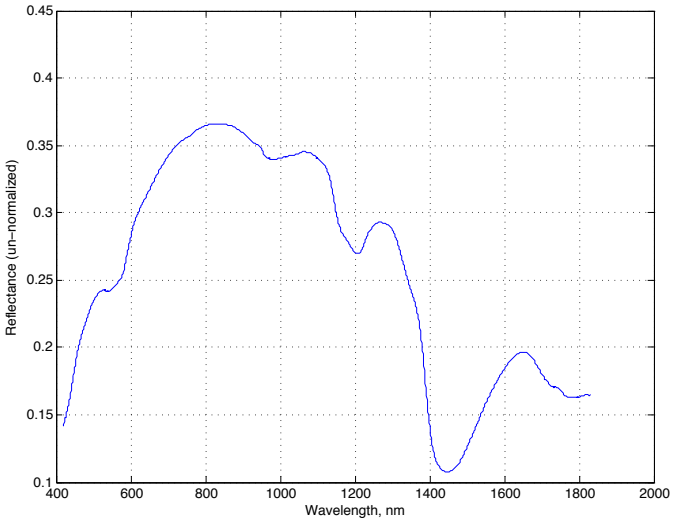


Figure 1.5: The HSI signature for a patch of human skin. This signature was obtained *in vivo* at the Johns Hopkins Hospital under an IRB approved study.

nomenon, then we would need to ensure that even if the model was deemed accurate, the results were not merely phenomenological. However, in our case, we don't seek to ascertain a mechanism, but merely relate an observed phenomenon to some underlying parameters. In other words, in a clinic we observe a measured HSI reflectance signature. We merely seek a mathematical function that completely explains that measurement in terms of our own defined parameters (i.e., the skin parameters), and no other parameters.

In this thesis we developed this model in two stages: (1) the forward model, and (2) the inverse model. Strictly speaking the inverse model yields the solution we seek, but we will motivate the need for the forward model later in this section. We start with the forward model. The terminology "forward" refers to finding a mapping from

the skin parameter vector space (i.e., each skin parameter is a scalar quantity, for example, collagen concentration, but when all parameters are taken together we get a vector – one dimension for each parameter), to the HSI vector space (each dimension corresponding to a spectral band). The technical details of the forward model are discussed in Chapter 2.

The task of estimating the skin parameters from the HSI signature was accomplished using the inverse model. This model was developed using a machine learning based regression methodology. In particular we used supervised learning; we sought to infer a mathematical function that explained a set of labeled data. For example, suppose we had a moderately large dataset of HSI signatures collected from various individuals who spanned various genders, ages, and ethnicities, i.e., variety in their biological skin components (e.g., large range of melanin, collagen, etc. concentrations). Further suppose that for each of these HSI signatures, we knew the ground truth, i.e., for each signature we had the exact corresponding parameter vector (concentration of each of the parameters). Then, we could train a supervised learning methodology that would take this dataset of labeled data and infer a mathematical function that would allow us to go from spectral space to parameter space. Then, if presented with an unlabeled HSI signature, we could use the inferred mapping to estimate the underlying parameters. This is exactly the framework for supervised learning, and this labeled dataset is often denoted as a "training dataset."

In practice, it is very difficult, if not impossible, to measure each of the skin

biological parameters. Therefore it is not possible to instantiate such a training dataset using *in vivo* data. Suppose we had a method that allowed us to create a proxy or a synthetic dataset that had HSI skin signatures and their parameters, where the relationship between the two was highly accurate, and the signatures were an accurate depiction of *in vivo* signatures. This proxy dataset could then be used as an appropriate training dataset. This is exactly the reason why we developed the forward model; it allowed for a principled way of estimating a highly accurate relationship between HSI signatures and skin parameters, which could then serve as a training dataset for the inverse model. The technical details of our machine learning based inverse model and the motivation for using machine learning for this task are described in Chapter 3.

1.4 Prior Work

One potential taxonomy of research in skin optics is as follows: (1) studies focused on developing robust biophysical models of human skin, (2) studies focused on developing methods to estimate the biological components of skin, (3) studies focused on using computational models, coupled with imaging modalities, for medical diagnostics. In this section, we review some of these studies, and also highlight our own contribution.

The first group of studies deals with developing models of how light interacts with

human skin. The goal of these studies is to create models that can represent skin and all of its optical and biochemical properties using simply the measured spectral power distribution. This power distribution is given by the measured reflectance and transmittance [17]. One of the earliest such models was developed by Kubelka and Munk who related the reflectance spectra of paint to its absorption and scattering coefficients [16, 18]. The Kubelka-Munk (K-M) theory based model uses the absorption and the scattering coefficients as inputs for the energy transport equations to describe the transfer of radiation in scattering media (such as human skin). The K-M model is also referred to as the "Flux Model" because it incorporates two fluxes (diffuse upwards and downwards fluxes). Since the original work by Kubelka and Munk, the model has been improved and optimized by several groups. Van Germet et al. [19] expanded the K-M model to account for tissue scattering behavior. Tuchin et al. [20] and Yoon et al. [21] built on this work and incorporated four and seven fluxes, respectively, to account for the radiation scattering. Meglinski et al. [7] developed a hybrid K-M model that uses the Fresnel equations to speed up the model output. Nunez et al. [16] modified this model by incorporating optical parameters coupled with *in vivo* and *ex vivo* measurements. They reported better performance for their model compared with other K-M based models [16]. Therefore, in this thesis, we employed their hybrid K-M model.

Several groups have also developed variants of the K-M model such as the diffusion theory model, which uses the Boltzmann photon transport equation, the absorption

and scattering coefficients, and a phase function [22]. Other variants include the radiative transport model [23, 24], and the Monte Carlo methods model [25] however these are typically used for laser applications. In addition, Baranoski et al. [17] commented that for these models, the comparisons between modeled and measured data are seldom provided. Nonetheless, these models have been used by several groups [26] for a host of biomedical applications. In contrast to the K-M variant models, Hanrahan and Kreuger [27] (expanded by Li et al. [28]) developed a scattering model, i.e. the H-K model, which models skin as two-layers, epidermis and dermis. Finally, Stam et al. developed the discrete-ordinate model, i.e. the D-O model, which has been used to simulate the scattering behavior of human skin [29]. The D-O model treats skin as a single layer with homogeneous optical properties and index of refraction. While each model has advantages and disadvantages, the K-M based models offer the greatest compromise between computational efficiency and accuracy [17].

The second group of studies focuses on developing methods to estimate human skin parameters. In particular, several groups have used variants of the K-M model in order to estimate skin parameters. Our study best fits into this category. The following studies are the closest to ours, in terms of methodology and validation. Cotton, Claridge, Preece et al. [30, 31] employed a K-M theory based model to estimate the melanin concentration in skin. They also developed [32,33] an inverse model based on finding a set of optimal image filters to minimize the error between a mapping from

color images to skin parameters. Doi and Tominaga [34] have also used a K-M model, coupled with a least squares method, to fit measured and estimated reflectance spectra in order to estimate skin pigments (e.g. melanin, etc). Pilon et al. [35, 36] developed a semi-empirical model for diffuse reflectance of two-layered media by approximating the solution of the Radiative Transfer Equation (RTE). They then estimated a set of skin parameters using this optical two-layer model and an inverse method based on least squares minimization. We compare our methods, results, and validation to these studies in Chapter 4.

In addition, there also exists work by Anderson et al. [37] and Wan et al. [38] who used the K-M model to estimate the reflectance of dermal and epidermal tissues in vivo (but the physical skin parameters). Miyake et al. [39–41] estimated skin chromophores using Independent Component Analysis, and by matching simulated reflectances to the estimated reflectances at each pixel of a multi-spectral image. Moncrieff et al. [42] created a model of skin and tissue coloration by finding the spectral composition of light remitted from skin parameters. They used this model to find a mapping between color images and pigmented skin lesions. Some groups, such as Alander, Kaartinen, Leonardi, Zerubia, et al. [43–47], have also used hyperspectral and multispectral imaging, often coupled with K-M model variants and/or inverse methods, to estimate skin chromophore concentrations and classify skin pigmentation.

The last group of studies apply computational models and medical imaging modalities to medical diagnostic applications. In particular, these studies apply the methods

from the previous two groups of studies to a clinical setting. For example, Pilon et al. [48] collected hyperspectral oximetry data from 54 *in vivo* subjects with some degree of foot ulcers. They then used their algorithms to classify the subjects into two groups: one, whose ulcers healed within 24 weeks, and second, whose ulcers did not heal within 24 weeks. They based their classification on the estimated concentrations of oxyhemoglobin, de-oxyhemoglobin, and oxygen saturation. Similar studies by other groups include classifying skin lesions from images [49], investigating skin alterations in diabetes patients [50], and assessing hemodynamic changes in skin, post-burn [51], amongst several others.

A subset of the methods and results presented in this thesis have been published by our group in peer-reviewed conference proceedings and journal articles [52–55].

1.5 Thesis Objectives

Aim 1. Develop a physics-based forward model to transform biological skin parameters into an estimated hyperspectral signature. The human skin is made up of various biological materials. We want to create a model that relates each of these biological materials to the measured hyperspectral signature, as would be obtained in a clinical setting. We use principles of optics and the interaction of light with multi-layered interfaces to model all paths light can take when incident upon the surface of such an interface, i.e., the skin. These principles allow us to esti-

mate how much light we can expect to be reflected off the surface of the skin based on the concentration of the underlying materials. Therefore, we can create a "forward" model that allows us to estimate a reflectance measurement, given the concentration of each of the biological skin components.

Aim 2. Develop a machine learning based regression methodology to invert the forward model. In a clinical setting we are presented with the inverse problem, i.e., we are not given biological materials and asked to estimate a reflectance measurement, but rather we are presented with a reflectance measurement, and we desire the concentration of the underlying parameters. If our forward model were invertible, then this would be trivial. Unfortunately it is not invertible, and therefore we take a statistical approach to invert the model. We use principles of machine learning, in particular support vector regression, and k-nearest neighbors, to develop a regressor that can take a HSI measurement and recover the underlying biological parameters. These are supervised learning methods, and therefore we use our forward model to generate a dataset of synthetic labeled signatures for training, and test on both new synthetic signatures, and *in vivo* signatures obtained under IRB at the Johns Hopkins Hospital.

Aim 3. Develop a novel performance evaluation methodology using a clinical Ultrasound system. A challenge in skin optics for nearly 30 years has

been developing non-invasive tools for performance validation. Suppose we develop a non-invasive method to estimate the thickness of human skin. How would we go about validating if our estimate was correct? It would require one to excise that patch of skin through a biopsy and measure it under a microscope. The burden is even greater for parameters such as collagen and melanosomes, which would require extensive biochemical analysis. In this thesis we have used a number of performance validation schemes (e.g., physiological precept analysis), as will be discussed in later sections. However, for skin thickness, we are able to obtain a "gold standard" by collecting Ultrasound (US) scans of the exact same patch of skin as the HSI signature. US has sufficient resolution to see the boundary between skin, fat, and bone, and hence with the aid of a dermatologist we can measure the thickness of the skin in the Ultrasound image. This is then compared to our machine learning based estimate, and a measure of true performance can be established.

Chapter 2

Physics-based Forward Modeling

We seek to develop a model that allows us to estimate a measured reflectance HSI signature based on just the biological parameters that make up the layers of the human skin. This is a physics-based model, meaning that the mathematical relationship between parameters and reflectance is derived using first principles from optics and photonics.

We start by describing how human skin is modeled in this thesis, which includes underlying assumptions and simplifications to make the model tractable. Next we detail mathematical notation for the forward model that is used for the rest of this thesis. Finally, the last two sections of this thesis review the fundamental optical equations and theories needed to model the path of light as it enters the skin, and is subsequently reflected back.

2.1 Human Skin Model

The model of human skin presented in this thesis was based on the models derived by Meglinski et al. [7] and later modified by Nunez et al. [16]. Skin was treated as an $N = 10$-layered Lambertian material, which allowed for uniform bidirectional reflection. The 10 layers are defined as follows: Layers 1 - 5 represent the strata of human skin; these layers are: stratum corneum, stratum lucidum, stratum granulosum, stratum spinosum, and stratum basale. Layers 6 - 9 are the dermises; these layers are: papillary dermis, upper blood net dermis, reticular dermis, and deep blood net dermis. Layer 10 is the subcutaneous tissue layer, and it is assumed to be infinitely thick, allowing the model to ignore the transmittance through it, as it is zero for a layer with infinite thickness.

Skin is composed of seven biological parameters. In this work, we estimated six such parameters. These parameters were: melanosome concentration (p_m), collagen concentration (p_c), O_2-saturation (p_{bo}), blood volume (p_{bl}), layer thickness, (p_{dt}), and subcutaneous reflectance (p_{sr}). The last parameter, water volume (p_{wl}), was assumed known and hence not estimated. This was a reasonable assumption because the original model detailed by Meglinski et al. [7] was not dependent on water level. Put another way, while the volume of water in certain layers of skin can change as a function of pathology or old age, the net effect on the measured reflectance is marginal, largely due to the fact that the change due to age and even pathology is minimal. Table 2.1 details each of the parameters, and the valid normal (i.e., healthy)

Table 2.1: Biological Parameter Detailed Descriptions and Physiological Ranges.

Parameter	**Description**	**Range (per Vol.)**
p_m (%)	Melanosome fraction by volume in the epidermis	*0.80% - 43%*
p_c (%)	Collagen fraction by volume in the reticular dermis	*15% - 30%*
p_{bo} (%)	Percentage of oxygenation of hemoglobin in blood	*70% - 100%*
p_{bl} (%)	Percentage of blood by volume in the dermis	*0.25% - 2.00%*
p_{dt} (mm)	Thickness of each layer measured in mm	*1.20 - 3.00mm*
p_{wl} (%)	Percentage of water by volume in each layer of skin	*5, 20, 20, 20, 20, 50, 60, 70, 70*
p_{sr}	Scale factor for the subcutaneous reflection	*0.40 - 0.65*

physiological range for each parameter. The percentage of water in each layer of skin, as described by Nunez et al. [16] is also included in the Table.

In order to make the model tractable, the following underlying assumptions (as described by Nunez et al. [16]) were made. First, each layer was assumed to have similar optical properties, and homogeneous absorption and scattering coefficients. This means that the concentration of each parameter is the same for all the layers. The parameter p_{sr} is an internal model parameter that doesn't have any physiological meaning. Therefore, it is omitted from further analysis. Therefore, we were estimating five parameters in this study (omitting p_{sr}). Next, each layer had a particular

water percentage. These water percentages were assumed known and not estimated. They were kept constant for each layer based on work by Meglinski et al. [7]; these are tabulated in Table 2.1. Finally, blood was assumed to be uniformly distributed in the dermis layers (rather than in differing concentrations for each layer) and zero in the strata. While some of these constraints may not be consistent with real human skin, Nunez et al. demonstrated minimal modeling error despite these underlying simplifications [16].

2.2 Forward Model

The proposed method was based on a physics-based forward model that describes the reflectance spectra of human skin based on physiological optical parameters that make up its layers.

The forward model describes a method of modeling the reflectance spectra of each layer based on the knowledge of each layer's thickness and the optical properties of its constituent components [6, 16]. In general, the forward mapping can be described as follows:

$$F : \mathbf{p} \rightarrow \mathbf{s} = f(\mathbf{p}) \tag{2.1}$$

where,

$$\mathbf{p} = \Big[p_m, p_c, p_{bo}, p_{wl}, p_{bl}, p_{dt}, p_{sr} \Big]^T$$

is a vector containing the skin parameters (see Table 2.1), and

$$\mathbf{s} = \left[\lambda_{N_A}, \ldots, \lambda_{N_B}\right]^T$$

represents the corresponding hyperspectral signature vector.

2.3 Kubelka-Munk Theory

The relationship described by equation (2.1) is based on a set of analytical models [6,16] that describe the transmission, $t_n(\lambda)$, and reflection, $r_n(\lambda)$, of light at a specific wavelength λ in a layer of (biological) material, where n denotes the layer number. The reflection and transmission are computed using the Kubelka-Munk equations, given by:

$$t_n(\lambda) = \frac{4\beta_n(\lambda)}{(1+\beta_n(\lambda))^2 e^{K_n(\lambda)d_n} - (1-\beta_n(\lambda))^2 e^{-K_n(\lambda)d_n}}$$

$$r_n(\lambda) = \frac{(1-\beta_n(\lambda))^2 (e^{K_n(\lambda)d_n} - e^{-K_n(\lambda)d_n})}{(1+\beta_n(\lambda))^2 e^{K_n(\lambda)d_n} - (1-\beta_n(\lambda))^2 e^{-K_n(\lambda)d_n}} \tag{2.2}$$

with the following parameters tied to the absorption and scattering properties of the biological materials:

$$\beta_n(\lambda) = \sqrt{\frac{A_n(\lambda)}{A_n(\lambda) + 2S_n(\lambda)}}$$

$$K_n(\lambda) = 2\sqrt{A_n(\lambda)(A_n(\lambda) + 2S_n(\lambda))} \tag{2.3}$$

where d_n denotes the thickness of layer n, and is equal to p_{dt}. The coefficients, $A_n(\lambda)$ and $S_n(\lambda)$ are based on the absorption, $a_n(\lambda)$, and the scattering, $s_n(\lambda)$, coefficients and are given by:

$$A_n(\lambda) = \frac{a_n(\lambda)}{\frac{1}{2} + \frac{1}{4}\left(1 - \frac{s_n(\lambda)}{(s_n(\lambda)+a_n(\lambda))}\right)}$$

$$S_n(\lambda) = \frac{s_n(\lambda)}{\frac{4}{3} + \frac{38}{45}\left(1 - \frac{s_n(\lambda)}{(s_n(\lambda)+a_n(\lambda))}\right)} \tag{2.4}$$

The coefficient $a_n(\lambda)$ is computed separately for each strata (layers 1-5) and dermis (layers 6-9), and is based on the components of $\mathbf{p}$ as seen by:

$$a_{n_{1-5}} = p_m a_m + p_c a_c + p_{wl} a_{wl} + a_{ce}$$

$$a_{n_{6-9}} = p_c a_c + (p_{wl} + 0.9 p_{bl}) a_{wl} + p_{bl}[p_{bo} a_{ohb} + (1 - p_{bo}) a_{dhb} + a_{ce} + a_{bil}(\lambda)] \tag{2.5}$$

$$s_n = \frac{30477(\lambda)^{-1.283}}{0.3} p_c a_c$$

where a_m, a_c, and a_{wl}, are the absorption profiles of the biological materials (melanosome, collagen, water) contained in $\mathbf{p}$ based on empirically derived values tabulated in [16]. The coefficients a_{ce}, a_{ohb}, a_{dhb}, and a_{bil} are the absorption profiles of betacarotene, oxygenated hemoglobin, de-oxygenated hemoglobin, and bilirubin, whose absorption profiles are also included in [16].

2.4 Light Transport Model

In addition to using the Kubelka-Munk equations, we also used the Fresnel equation, as detailed by Nunez et al. [16]. Since human skin has uniform bidirectional reflection, Nunez et al. assumed that light incident on the skin surface is always normal to the surface. This allowed us to use the Fresnel equation to describe the amount of reflection that is normal to the interface separating the skin from air. The equation is based on the tabulated indices of refraction for air and the stratum corneum (skin Layer 1), given by η_{air}, and η_{sc} respectively. The Fresnel reflection FR is then given by:

$$FR = \left(\frac{\eta_{sc} - \eta_{air}}{\eta_{sc} + \eta_{air}}\right)^2 \tag{2.6}$$

The reflectance path of light was modeled in the following fashion. For every layer, light can take one of four paths; it is either:

1. absorbed,
2. scatters out of the top of the layer,
3. scatters out of the bottom of the layer, or
4. doesn't scatter, and continues along its path.

Therefore, the reflectance and transmittance between any two interfaces (layers) are infinite sums dependent on what path light takes between those two interfaces.

In addition, light being reflected off the surface of the skin is actually made up of light coming from each path leaving a skin layer, as well as the Fresnel reflection. In other words, the total fraction of light leaving each layer i is the product of:

i. the Fresnel transmittance,

ii. transmittance of all the layers it had to go through in order to reach layer i,

iii. the reflectance of the layer i (the model assumes an infinitely thick bottom layer, and therefore has pure reflectance),

iv. the transmittance of all the layers it must once again traverse in order to reach the top, and finally

v. the Fresnel transmittance.

Note that the transmittance from layer 1 to layer i is the same as the transmittance from layer i to layer 1.

Using this methodology, the the total reflectance and transmittance between any two interfaces is the sum of all reflectance paths (see steps 1-4 above) and the transmittance and reflectance it accrued (see steps i-v above). This is given by:

$$\begin{aligned} R_j &= r_i + t_i^2 r_j \sum_{m=1}^{\infty} (r_i r_j)^m \\ T_j &= t_i t_j \sum_{m=1}^{\infty} (r_i r_j)^m \end{aligned} \tag{2.7}$$

where R_j, and $T_j = 1 - R_j$, are the total reflectance and transmittance between interface i and j. The other variables are computed using equations (2.2) - (2.5) from

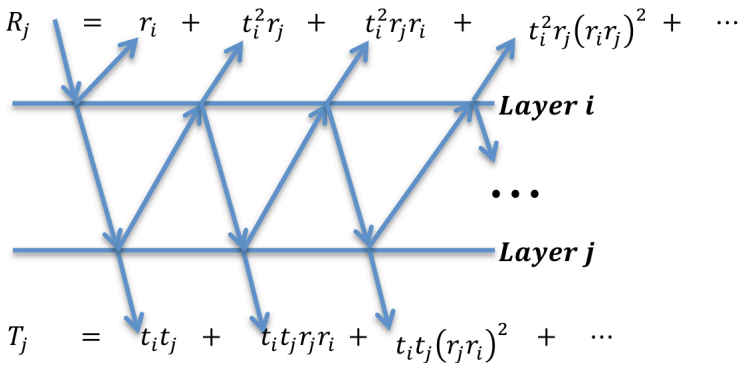


Figure 2.1: A graphical representation of Equation 2.7; the light transport model for possible paths of light between any two interfaces (layers).

above. A graphical version of this, expanded from [16], is presented in Figure 2.1.

The final step is to iterate the quantities in Equation 2.7 for all optical layers 1 through $N = n + 1$. This is analogous to the procedure presented in Figure 2.1 and Equation 2.7, except adapted for all N layers of the skin, until the overall reflectance R_N (and transmittance T_N) is computed. The closed form solution is given by:

$$\begin{aligned} R_{n+1} &= R_n + \frac{T_n^2 r_{n+1}}{1 - R_n r_{n+1}} \\ T_{n+1} &= \frac{T_n t_{n+1}}{1 - R_n r_{n+1}} \end{aligned} \tag{2.8}$$

Using this methodology, the reflectance spectra of human skin, given by $\mathbf{s}$, was generated in the UV-VIS through the SWIR regions of the electromagnetic spectrum based on the physiological parameters $\mathbf{p}$.

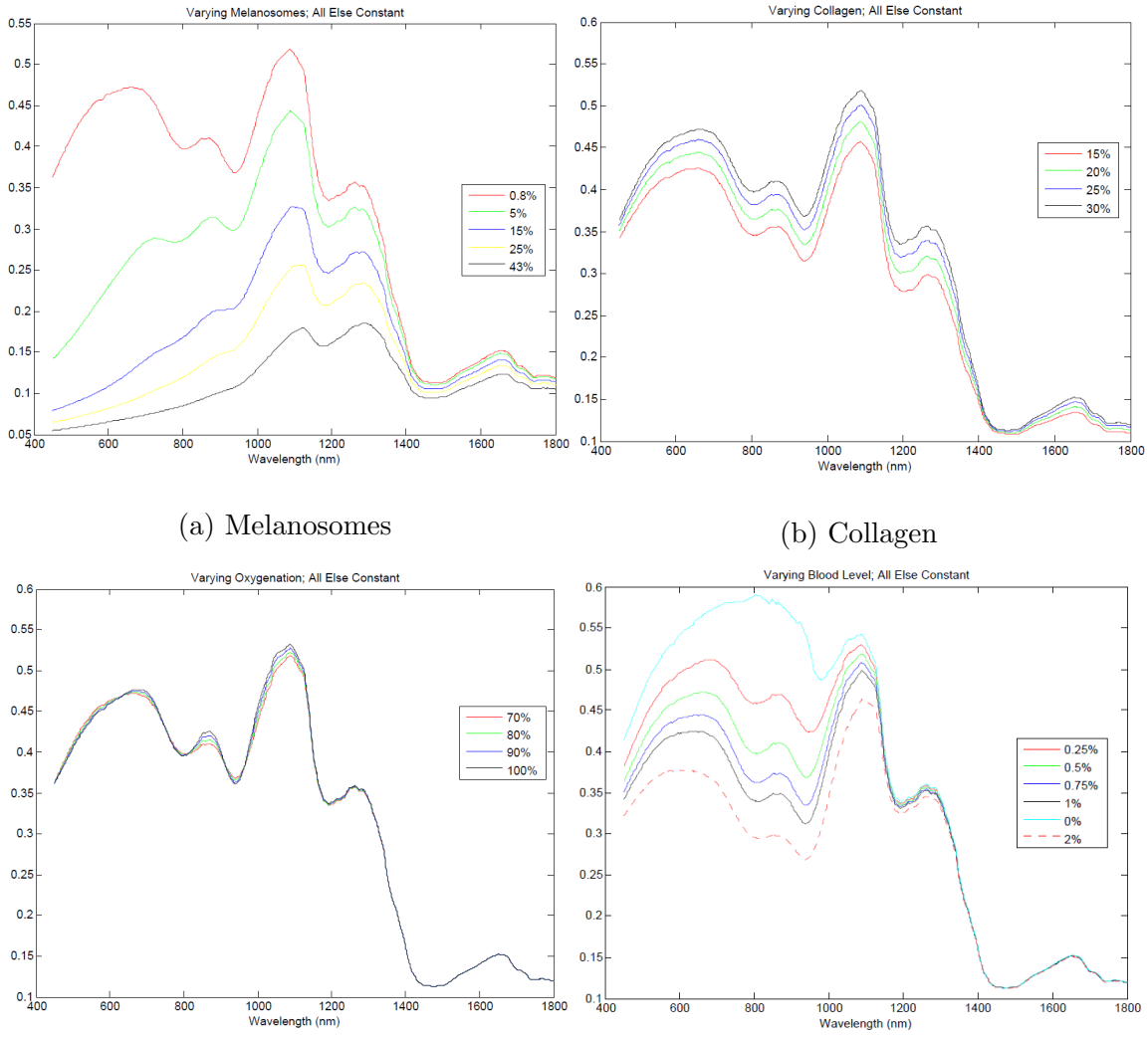
2.5 Results

To understand the effect of each physiological parameter, we plotted Equation (2.8) as a function of a single physiological parameter. Specifically, we assigned each component of $\mathbf{p}$ a value that was in the center of the physiological range. Next, we plotted the expected reflectance signature while manipulating a single component of $\mathbf{p}$. For example, after initializing every component of $\mathbf{p}$ we plotted Equation (2.8) as a function of varying the melanosome concentration from 0.8% to 43% (the full range of p_m). This was repeated for collagen, oxygen saturation, and blood volume.

Figure 2.2 demonstrates the result of our experiment. Note that "initializing" the value of each parameter to be in the center of the physiological range may not actually be physiological, i.e., while each component is independently physiological, all taken together may not be physiological. Nonetheless, this *synthetic* experiment still shows the effect that varying a single parameter can have on the shape and magnitude of a HSI signature, if all other components are held constant.

Figure 2.2 (a) revealed that manipulating the melanosome concentration affects the shape of the HSI signature in the VIS region. In particular, the characteristic "w" shape around 800 nm was degraded as the concentration of melanin was increased. Figure 2.2 (b) revealed that for collagen, there was no loss in shape, but rather the magnitude of the signature was changed. Figure 2.2 (c) demonstrated that varying the oxygen saturation had almost no impact on the HSI signature; this issue is later explored in greater detail in Chapter 4. Finally, Figure 2.2 (d) showed a similar

pattern for blood volume as was observed for melanin, however, it seemed that the characteristic "w" was largely retained independent of concentration. It was also observed that the changes in shape and magnitude were more pronounced in the 400 - 1100 nm region than the 1100 - 1800 nm range.



(a) Melanosomes

(b) Collagen

(c) Oxygen Saturation

(d) Blood Volume

Figure 2.2: Impact on the shape and magnitude of the estimated HSI signature as a function of manipulating a single physiological parameter.

Chapter 3

Machine Learning based Inverse Modeling

In Chapter 2, we described a physics-based model that maps the physiological parameters of skin to its observed reflectance spectrum. This section presents two methods to compute the inverse map, answering the question: "*given a reflectance spectrum of human skin, what are its underlying physiological parameters?*".

This problem can be expressed mathematically as follows:

$$G : \mathbf{s} \rightarrow \mathbf{p} = g(\mathbf{s}) \tag{3.1}$$

which is the inverse of Equation (2.1). We adopted two machine learning regression algorithms to estimate the inverse mapping function $g(\cdot)$: (A) support vector regression (SVR) [56, 57] and (B) k-nearest neighbors based regression (k-NN) [58].

The use of regression is motivated by the availability of a large number of training

spectra $\mathbf{s}$ and their associated physiological parameters $\mathbf{p}$. It leverages the abundance of this data to yield an estimate of the parameters of $g(\cdot)$ with minimal error and low variance. In contrast to other studies to estimate skin parameters, we utilized machine learning methods because they provide the following benefits: *Non-parametric*: the functional form of $g(\cdot)$ is not known and it cannot be analytically computed. The learning paradigm is data-driven and does not require any *a priori* knowledge about the functional form of $g(\cdot)$. *Generalizable*: these methods avoid over-fitting and hence offer good results to unseen data when compared to traditional methods. *Kernels*: these methods can exploit the "kernel trick" to use linear regression to accurately model non-linear functions. Since one cannot assume linearity for $g(\cdot)$, machine-learning regression can be a powerful tool for learning the non-linear function that maps the physiological parameters to their observed reflectance spectrum.

3.1 Support Vector Machine Regression

For regression, we first used Support Vector Machine Regression (SVR), a now established machine learning approach, with several open source implementations (e.g. [59]) in C++ and Matlab. The remainder of this section gives a high level motivation of the SVR technique and its implementation. The full mathematical details are left as a reference to Vapnik [56] and Smola [60].

To solve the regression problem in a way that can be approached using SVR,

we decoupled the function $g(\cdot)$, which maps a spectral vector $\mathbf{s}$ to a skin biological parameter vector $\mathbf{p}$, into a set of five (one for each component of $\mathbf{p}$, excluding p_{wl}, which was known, and hence kept constant) scalar regression subproblems where:

$$h : \mathbf{s} \rightarrow p = h(\mathbf{s}) \tag{3.2}$$

where p is one of the scalar parameters (components of $\mathbf{p}$) we wish to estimate (e.g. collagen level, melanosome level, etc.), and h is the scalar regression function. Each of the scalar parameters were then individually estimated in the same fashion.

Because the approach is data driven, we started by considering the training dataset that was generated using the forward model described in Chapter 2. This consisted of parameters vectors $\mathbf{p}_i$ and their associated spectral vectors $\mathbf{s}_i$. We considered the set of pairs including one spectral vector and one biological scalar parameter we wish to estimate:

$$\{(\mathbf{s}_1, p_1), ..., (\mathbf{s}_n, p_n)\} \tag{3.3}$$

In the simple linear case, the function h can take a functional form given by:

$$h(\mathbf{s}) = \sum_{i=1}^{l} \alpha_i \langle \mathbf{s}, \mathbf{w}_i \rangle + \ b \tag{3.4}$$

where $\langle \mathbf{s}, \mathbf{w} \rangle$ represents the linear dot product. The regression function h approximates the output parameter p as a weighted linear combination of the input spectral vector $\mathbf{s}$, dot product with a set of spectral vectors $\mathbf{w}_i$, with an added offset b. The goal is to find a small $\mathbf{w}$. This is achieved by minimizing the norm, i.e. $||\mathbf{w}||^2 = \langle \mathbf{w}, \mathbf{w} \rangle$ [60]. This problem can then be written as a constrained optimization problem

given by:

$$
\begin{aligned}
&minimize \quad \frac{1}{2}||\mathbf{w}||^2 + \ C\sum_{i=1}^{l}(\zeta_i + \zeta_i^*) \\
&subject\ to \begin{cases} p_i - \langle \mathbf{w}, \mathbf{s}_i \rangle - \ b \ \leq \ \epsilon + \zeta_i \\ \langle \mathbf{w}, \mathbf{s}_i \rangle + \ b - p_i \ \leq \ \epsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \ \geq \ 0 \end{cases}
\end{aligned} \tag{3.5}
$$

where ζ_i and $\zeta_i{}^*$ are slack variables to account for infeasible constraints on the problem as per the soft margin loss function. C is a strictly positive constant, and it accounts for the degree to which errors larger than ϵ are tolerated. This is accounted for by optimizing the soft margin loss setting. The specific function, the ϵ-insensitive loss function, $|\xi|_e$ is given by:

$$
|\xi|_e : \begin{cases} 0 \quad if\ |\xi| \leq \epsilon \\ |\xi| - \epsilon \quad otherwise \end{cases} \tag{3.6}
$$

In this case (as is common in most cases), the dimensionality of $\mathbf{w}$ is much higher than the number of observations, therefore, the optimization problem posed in Equation (3.5) can be solved with much more ease in its dual formulation. As done in [56], a dualization method formulated by R. Fletcher [61] using Lagrangian multipliers is implemented. The first task is to construct a Lagrange function from the primal objective function and its corresponding constraints. As per [62–64] this function has a saddle point with respect to the primal and dual variables at the solution. The

Lagrangian function is then given by:

$$\begin{aligned} L := \frac{1}{2}||\mathbf{w}||^2 + \ & C\sum_{i=1}^{l}(\zeta_i + \zeta_i^*) + \sum_{i=1}^{l}(\eta_i\zeta_i + \eta_i^*\zeta_i^*) \\ & - \sum_{i=1}^{l}\alpha_i(\epsilon + \zeta_i - p_i + K(\mathbf{w}, \mathbf{s}_i) + \ b) \\ & - \sum_{i=1}^{l}\alpha_i^*(\epsilon + \zeta_i^* + p_i - K(\mathbf{w}, \mathbf{s}_i) - \ b) \end{aligned} \tag{3.7}$$

where L is the Lagrangian and α_i, $\alpha_i{}^*$, η_i, and $\eta_i{}^*$ are the Langrangian multipliers. The Lagrangian multipliers in Equation (3.7) have to be greater than or equal to zero as it is a constraint posed by the optimization problem. Furthermore, as a consequence of the saddle point condition, the partial derivative of L with respect to each of the primal variables ($\mathbf{w}$, b, η_i, and $\eta_i{}^*$), is zero. This is seen by the following:

$$\begin{aligned} \frac{\partial L}{\partial b} &= \sum_{i=1}^{l}(\alpha_i^* - \alpha_i) = 0 \\ \frac{\partial L}{\partial \mathbf{w}} &= \mathbf{w} - \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\mathbf{s}_i = 0 \\ \frac{\partial L}{\partial \zeta_i^*} &= C - \alpha_i^* - \eta_i^* = 0 \end{aligned} \tag{3.8}$$

Equation (3.8) can then be substituted into equation (3.5), and after some algebraic manipulations, and evaluations of the derivatives, the new dual optimization problem can hence be written as:

$$maximize \begin{cases} -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\mathbf{s}_i - \mathbf{s}_j) \\ \\ -\epsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l}(\alpha_i + \alpha_i^*)p_i \end{cases} \tag{3.9}$$

$$subject\ to \quad \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0; \alpha_i, \alpha_i^* \in [0, C]$$

In Equation (3.9) the dual variables have been eliminated, and the "support vector expansion," can then be written as:

$$\mathbf{w} = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\mathbf{s}_i \tag{3.10}$$

This analysis is typically used for the linear case of SVR; in this work, the non-linear case needs to be used. However, the analysis and methodology is largely similar. Kernel methods are used in order to account for non-linearities. Therefore, the linear dot product is converted into a kernel dot product, given by $\mathrm{K}\langle\mathbf{w}, \mathbf{s}\rangle$. Then, the process proceeds in the same manner, and from analogy, it arrives at the following result for the support vector expansion:

$$g(\mathbf{s}) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) \cdot K\langle\mathbf{s}, \mathbf{s}_i^*\rangle + \ b \tag{3.11}$$

The complexity of the function's representation by support vectors, therefore, only depends upon the number of support vectors, and not the dimensionality of the input space, $\mathbf{X}$. Finally, b can be computed using the Karush-Kuhn-Tucker (KKT)

conditions [65, 66]. They state that the product between the dual variables and the constraints must go to zero. This can be seen more formally by the following:

$$\begin{aligned} \alpha_i(\epsilon + \zeta_i - p_i + (\mathbf{w}, \mathbf{s}_i) + b) &= 0 \\ \alpha_i^*(\epsilon + \zeta_i^* - p_i + (\mathbf{w}, \mathbf{s}_i) + b) &= 0 \\ (C - \alpha_i)\zeta_i &= 0 \\ (C - \alpha_i^*)\zeta_i^* &= 0 \end{aligned} \tag{3.12}$$

Based on these constraints, there can never be a set of dual variables that are both non-zero at the same time. Therefore, the following conditions are imposed on b:

$$\begin{aligned} b &\geq max\{-\epsilon + p_i - (\mathbf{w}, \mathbf{s}_i) | \alpha_i < C \; or \; \alpha_i^* > 0\} \\ b &\leq min\{-\epsilon + p_i - (\mathbf{w}, \mathbf{s}_i) | \alpha_i > 0 \; or \; \alpha_i^* < C\} \end{aligned} \tag{3.13}$$

A more formal treatment of choosing an appropriate b is detailed in a technical report by S. S. Keerthi et al. [67].

In this manner, each parameter p_i, from Table 2.1, was independently regressed from the other parameters. Therefore, each p_i is a scalar, where it is one of the five components of $\mathbf{p}$ being estimated. This methodology was used in this work in order to first compute, $g(\mathbf{s})$, as given by Equation (3.11) and ultimately the desired inverse mapping G, as detailed in Equation (3.1).

In programmatic terms, the regression machine uses the following steps:

1. Train the machine to obtain α_i, α_i^*, $\mathbf{s}_i^*$, and b

2. Form the regression by first taking the dot product between $\mathbf{s}$ and each support vector $\mathbf{s}_i$

3. These are then scaled by the weights α_i^* and α_i

4. A linear combination is taken

5. Finally offset by the constant b.

This is conveniently implemented in Matlab and numerous Open Source implementations such as LibSVM [59] or OpenCV [68]. A more rigorous formulation of support vector regression is provided in literature [60, 69, 70].

3.2 k-Nearest Neighbors

This learning algorithm was used in order to classify objects based on a "majority vote" system. Much like the SVR approach, this algorithm started with a database of input training sets, known as the feature space, as presented in Equation (3.3). The k-NN algorithm regresses on a new testing set based on the closest training examples it finds in the feature space.

In this work, the feature space consisted of approximately 300,000 reflectance spectra generated as per Equation (2.1). The testing set consist of *in vivo* hyperspectral signatures that obey Equations (2.1) and (3.1). The goal was to compute g from Equation (3.1). As is typical for hyperspectral signatures, the k-NN algorithm

needs to compare the shape of each hyperspectral signature from the testing set to the signatures from the training set. This was done by computing the inner product between the two signatures, and can be written as the spectral angle give by:

$$\Phi = \cos^{-1}\left(\frac{\mathbf{s}_{testing} \cdot \mathbf{s}_{training}}{||\mathbf{s}_{testing}||\ ||\mathbf{s}_{training}||}\right) \tag{3.14}$$

where $\mathbf{s}_{training}$ is a computationally generated hyperspectral signature from the training set contained within the feature space, $\mathbf{s}_{testing}$ is the signature obtained in vivo that is being classified and $||\cdot||$ denotes the Euclidean norm. Once we found the "closest" signature from the testing dataset, we assigned the corresponding parameters of the closest signature to the testing signature. Naturally, this method relies heavily on the correspondence between training parameters and signature being highly accurate.

Alternatively, experiments were also performed using two other versions of the k-NN algorithm: (1) the closest neighbor was found by computing the Euclidean distance between spectra and (2) the training and testing datasets were first whitened, and then spectral angle is used to compute the nearest neighbors. A more mathematically rigorous treatment of k-NN regression is included in literature [58, 71].

3.2.1 Whitening

This section reviews the theory of whitening and how it was accomplished in this thesis. Whitening an observed dataset, $\bar{X}$, refers to the process of making its components uncorrelated and their variances equal to unity. This can also be expressed as

making its covariance matrix equal to the identity matrix, i.e.:

$$E\{\bar{X}\bar{X}^T\} = I$$

Whitening the observed data essentially converts the covariance from an ellipse to a sphere (having all dimensions be the same length, i.e. uniform). Mathematically, all the diagonal entires (eigenvalues) are made equal. Since whitening is a linear operation, a linear transformation by using the eigenvalue-decomposition (EVD) of the covariance matrix can be found. This can be expressed as:

$$E\{\bar{X}\bar{X}^T\} = VDV^T$$

where, V is an orthogonal matrix of eigenvectors and D is the diagonal matrix of its eigenvalues. The final step of the whitening is then:

$$\bar{X} = VD^{-1/2}V^T\bar{X} \tag{3.15}$$

Whitening is often used as a pre-processing step in the HSI literature before performing any type of analysis [72]. We used whitening for the k-NN algorithm, but we did not use it for SVR; we anecdotally found whitening to perform poorly with the SVR algorithm. This needs to be validated before any significant conclusions can be drawn.

The next chapter serves as a results section for this chapter; the performance of both SVR and k-NN is evaluated and discussed for both synthetic and *in vivo* data.

Chapter 4

Experiments

In this work, the goal was to compute the inverse mapping G, and hence create a model that can infer the underlying constitutive physiological parameters of human skin from hyperspectral signatures. As described in Chapters 2 and 3, we used a machine learning based regression methodology to accomplish this goal. Furthermore, as described earlier, we used a physics-based forward model to help instantiate a training dataset, which was then used for the training phase of our supervised learning methods.

This chapter is dedicated to evaluating the performance of our forward and our inverse method. The evaluations were done through two sets of experiments. The first set involved performance validation through synthetic experiments. The second set involved performance validation through *in vivo* experiments. In both sets, a skin reflectance model, which was generated according to equations (2.2) - (2.8), was

used. The skin reflectance model (labeled the *training dataset* for the remainder of this thesis) was generated as follows.

Following the physics-based model described in equation (2.1), approximately $N \approx 300,000$ exemplars of all seven parameters described by $\mathbf{p}$ were uniformly sampled on a grid (with a finitely small step size of approximately 0.05%) and distributed along their entire physiological domain. These ranges were detailed in Table 2.1. An equal number of samples were generated for each p_i. The forward model F was used to generate a dataset of hyperspectral signatures corresponding to each set of parameters $\mathbf{p}$. The water level was kept constant for each layer as detailed in [16]. This dataset is denoted as:

$$Training\ Dataset : \{(\mathbf{p}_i^T, \mathbf{s}_i^T); i = 1,\ \ldots,\ N\} \tag{4.1}$$

This dataset was used for both sets of experiments. If the utility of the forward model is to be believed, then this is a valid labeled dataset, which can be used for learning. We put this assumption to the test when performing our *in vivo* experiments.

4.1 Synthetic Experiments

In the first set of experiments, synthetic data was used to test the accuracy of the inverse model G. Therefore, a new testing *synthetic dataset* was generated by using the forward model F and a set of K = 50,000 biological parameters that were not contained within the training dataset. Each set of K = 50,000 parameters was

generated by randomly sampling along the parameters' physiological range. If a parameter was found to be too close to one already contained within the training dataset, it was discarded and a new one was generated. This ensured that there was no overlap between the training and testing datasets. The training set and the newly instantiated testing set were then used to perform the synthetic experiments. The synthetic testing dataset is denoted as:

$$\textit{Synthetic Dataset} : \{(\mathbf{p}_j^S, \mathbf{s}_j^S); j = 1, \ \ldots, \ K\} \tag{4.2}$$

4.1.1 Methods

Tuples from the training dataset (4.1), $\{(\mathbf{p}_i^T, \mathbf{s}_i^T)\}$, were used to train the SVR. The SVR algorithm employed here is implemented using the NTU SVM Library [56, 59] in Matlab. This produced the trained SVR model and its associated support vectors. Then, $\{\mathbf{s}_j^S\}$, as well as the SVR model was used to estimate $\{\mathbf{p}_j^S\}$. In these experiments, the values of $\{\mathbf{s}_j^S\}$ and $\{\mathbf{p}_j^S\}$ were already known, and therefore, they served as the ground truth values. However, by using the SVR and only $\{\mathbf{s}_j^S\}$, according to G, the *estimated* value of $\{\mathbf{p}_j^S\}$, given by, $\{\hat{\mathbf{p}}_j^S\}$, could be computed. The estimated values, $\{\hat{\mathbf{p}}_j^S\}$, could then be compared with the ground truth values, $\{\mathbf{p}_j^S\}$, in order to compute an error associated with the inverse mapping G. For visualization purposes, Figure 4.1 shows a block diagram outlining this experimental protocol.

The same experiment was repeated using the three flavors of the k-NN regression algorithm. For each $\mathbf{s}_j^S$, the k-nearest neighbors (in this case, the one neighbor in the

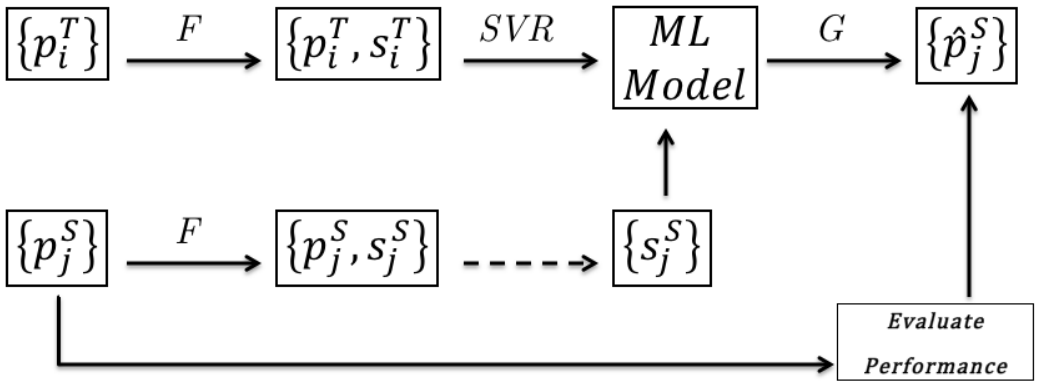


Figure 4.1: A graphical representation of the synthetic experimental protocol.

training dataset that had the smallest spectral angle, Φ, with each $\mathbf{s}_j^S$) were found. The estimated parameters, $\{\hat{\mathbf{p}}_j^S\}$, were the parameters corresponding to the nearest neighbor spectra, i.e. $\{\mathbf{p}_i^T\}$. Once again, since the ground truth values of $\{\hat{\mathbf{p}}_j^S\}$ were already available (given by $\{\mathbf{p}_j^S\}$), an error associated with the regression could be computed.

4.1.2 Results

In order to evaluate the performance, we computed the average absolute errors (AAE), for all four algorithms, along with the standard deviations (Std. Dev.) associated with the estimated biological parameter. These results are provided in Table 4.1. The average absolute error was computed as follows:

$$Average\ Absolute\ Errors \equiv \frac{\sum_{i=1}^{K} |\hat{\mathbf{p}}_j^S - \mathbf{p}_j^S|}{K};\ \ K = 50,000$$

Table 4.1: Average Absolute Errors (AAE) and the Standard Deviation (Std. Dev.) associated with Biological Parameter Estimation for synthetic experiments

Parameter	**SVR**		***k*-NN (Spectral Angle)**	
	AAE	Std. Dev.	AAE	Std. Dev.
p_m (%)	0.3379	5.0e-03	0.2387	6.8e-03
p_c (%)	0.4057	2.7e-03	0.2919	5.5e-03
p_{bo} (%)	4.1200	1.4e-01	2.8947	3.5e-01
p_{bl} (%)	0.0269	2.1e-03	0.0171	2.6e-03
p_{dt} (%)	0.0482	1.1e-03	0.0141	6.2e-03

Parameter	***k*-NN (Euclid. Dist.)**		***k*-NN (Whitened)**	
	AAE	Std. Dev.	AAE	Std. Dev.
p_m (%)	0.2566	4.0e-03	0.2427	9.6e-02
p_c (%)	0.4096	1.3e-03	0.2871	6.9e-03
p_{bo} (%)	5.7891	2.1e-01	2.6231	2.2e-01
p_{bl} (%)	0.0313	1.7e-03	0.0259	1.5e-03
p_{dt} (%)	0.0282	2.2e-03	0.0018	9.2e-03

4.2 In Vivo Experiments

The *in vivo* experiments were performed using a dataset obtained from *in vivo* hyperspectral imaging of 24 individuals of both genders and Caucasian ($n = 10$),

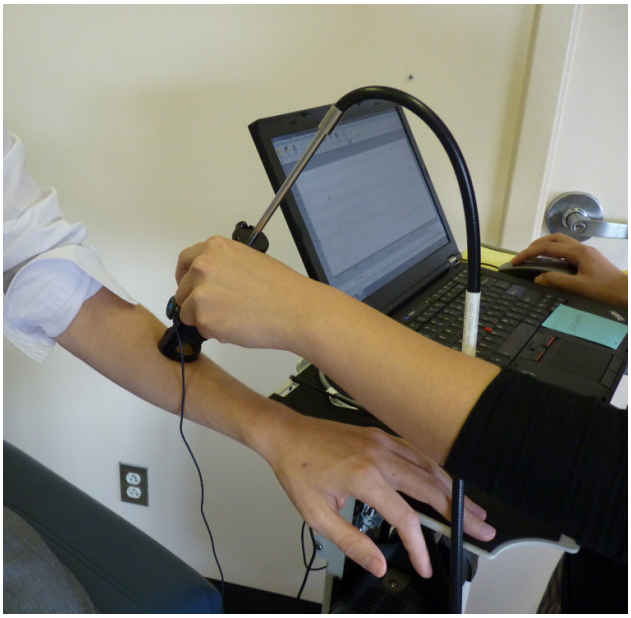


Figure 4.2: Hand-held Spectroradiometer probe, and positioning in the clinic. Photograph courtesy of Johns Hopkins Hospital, Department of Dermatology, CTREP.

Asian ($n = 8$), and African American ($n = 6$) ethnicities. The data was obtained at Johns Hopkins Hospital, Department of Dermatology, under protocols approved by the Institutional Review Board (IRB). All patients gave informed consent.

Hyperspectral signatures were obtained from each of the 24 individuals using the Analytical Spectral Devices (ASD), Inc., (Boulder, CO, USA) FieldSpec 3 Portable Spectroradiometer. The spectroradiometer had a hand-held probe which was positioned at a perpendicular angle to the skin, with the enclosed lens at a height of 5 cm from the skin. The spectroradiometer had a lens diameter of 10 mm, a field of view of 25 degrees, a 100 ms scanning time, and a built-in illumination source. Figure 4.2 is a photograph of the Spectroradiometer probe, and demonstrates the approximate

positioning of the probe during each scan.

The spectroradiometer has two detectors, one containing a 512-element Si photodiode array (for imaging up to 1000 nm) and the other detector contains two graded index InGaAs photodiodes (for imaging beyond 1000 nm). The instrument is calibrated using a panel whose reflectance is known; the amount of light captured by the instrument is correlated with the reflectance of the panel for each wavelength. The instrument is re-calibrated after each measurement. The spectra were obtained from 450 nm to 1800 nm blue with a 1 nm step size (bandwidth). The *in vivo* dataset obtained from this IRB is denoted as:

$$In\ Vivo\ Dataset : \{\mathbf{s}_k^R; k = 1,\ \ldots,\ M\} \tag{4.3}$$

4.2.1 Methods

This dataset was compiled by *in vivo* hyperspectral imaging of 24 individuals. Approximately ten hyperspectral signatures were collected from each individual, from five anatomical locations on their bodies (two signatures from each location), to sum to a grand total of $M = 241$ signatures in the *in vivo* dataset. These five locations included: the back, the palm, the cheek, the dorsal forearm (DF) and the upper inner arm (UIA).

The major difference between the *in vivo* experiments and the synthetic experiments is the ground truth. In the synthetic experiments, the estimated parameters $\{\hat{\mathbf{p}}_j^S\}$ were compared to the ground truth parameters $\{\mathbf{p}_i^T\}$, in order to assess the

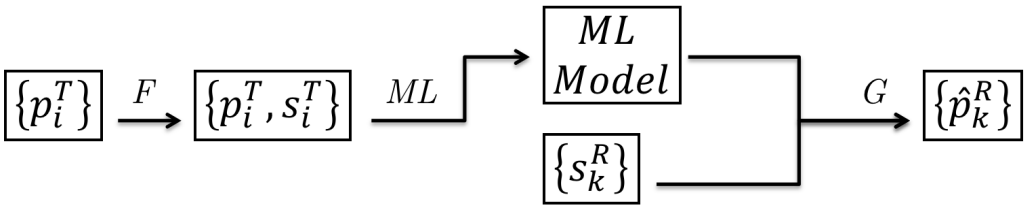


Figure 4.3: A graphical representation of the *in vivo* experimental protocol.

performance of the algorithms. The ground truth targets $\{\mathbf{p}_k^R\}$ were not available for the *in vivo dataset*. However, we still performed performance validations for the *in vivo* dataset. The experiments using SVR and k-NN were repeated in the same manner as done in Section 4.1. Again, for illustration purposes, refer to Figure 4.3. Notice that the targets $\{\mathbf{p}_k^R\}$ were not available, therefore the same analysis as in Section 4.1 cannot be done.

4.2.2 Results

The estimated parameters $\{\hat{\mathbf{p}}_k^R\}$ are summarized in Table 4.2 as a function of anatomical location, and ethnicity. In addition, a bar graph representation of these results is provided in Figure 4.4. We analyze these results using physiological precepts in the discussion section of this chapter. Note that the estimates for the skin thickness parameter, i.e., p_{dt}, were intentionally left out of this Table and Figure. This parameter was analyzed separately using a novel ground truthing methodology; the results and discussions are provided in Chapter 5. Therefore in this section we

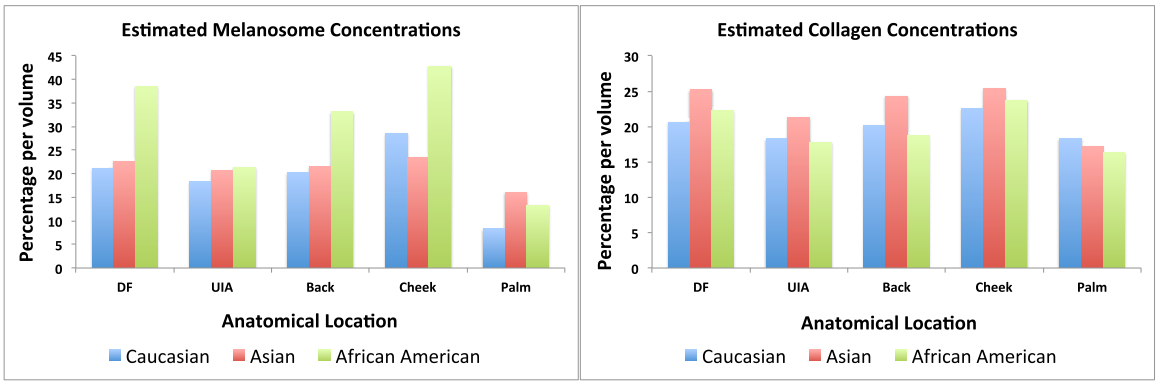
were only concerned with four parameters.

Table 4.2 and Figure 4.4 represent the value of each of the parameters estimated by the SVR and k-NN algorithms; these results are presented in the parameter space. Another way to visualize and analyze both the forward model and the inverse model is to take this output of the inverse model (the parameters) and use the forward model to convert them back into a HSI signature. We can then compare the signatures directly. Figure 4.5 shows a plot of three such signatures plotted alongside their corresponding *in vivo* signatures; we plot one example per ethnicity.

We performed additional error analysis in the HSI signature vector space; we reported error bounds between the "estimated" and the ground truth (i.e., the *in vivo* spectra). Table 4.3 reports the spectral angle, the Root Mean Square Error (RMSE), and the standard deviation between the *in vivo* measured spectra and the inverse model estimated spectra for each of the 24 patients. In Table 4.4 we built on this analysis and report the spectral angle error (measured in radians) as a function of anatomical location for each of the 24 patients.

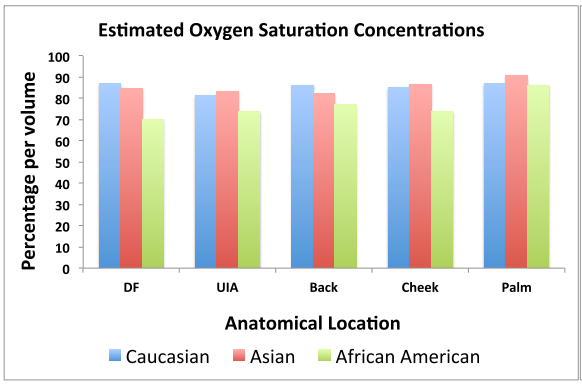
4.3 Discussion

In the past 30 years, several studies have attempted to solve the inverse problem of estimating skin parameters from spectral data. The validation of estimated skin parameters is a very difficult task because, in many cases, it is currently impossible

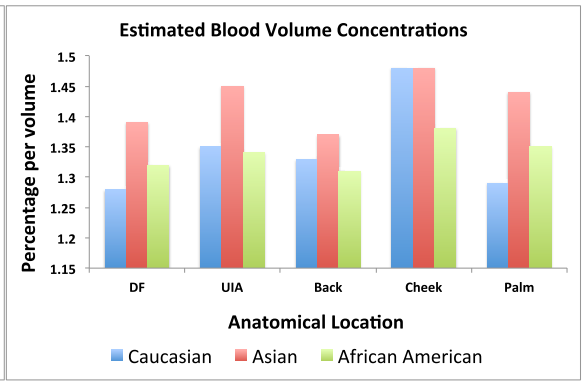


(a) Melanosomes

(b) Collagen



(c) Oxygen Saturation



(d) Blood Volume

Figure 4.4: Bar graph representations of the estimated (a) melanosome concentration, (b) collagen concentration, (c) oxygen saturation concentration, and (d) blood volume concentration, as a function of both ethnicity and anatomical location.

Table 4.2: Estimated skin parameters (percentage by volume) as a function of ethnicity and anatomical location (Dorsal Forearm (DF), Upper Inner Arm (UIA), Back, Cheek, Palm). The colors were used to aid in referencing specific values from the Table in the discussion section.

	Caucasian					
	DF	UIA	Back	Cheek	Palm	*Average*
$p_m(\%)$	21.2	18.3	20.2	28.5	8.3	19.3
$p_c(\%)$	20.6	18.3	20.2	22.6	18.3	20.0
$p_{bv}(\%)$	1.28	1.35	1.33	1.48	1.29	1.35
$p_{os}(\%)$	86.9	81.3	86.3	85.0	87.0	85.3
	Asian					
	DF	UIA	Back	Cheek	Palm	*Average*
$p_m(\%)$	22.6	20.7	21.5	23.5	16.1	20.9
$p_c(\%)$	25.2	21.3	24.3	25.4	17.2	22.7
$p_{bv}(\%)$	1.39	1.45	1.37	1.48	1.44	1.43
$p_{os}(\%)$	84.6	83.2	82.2	86.6	90.8	85.5
	African American					
	DF	UIA	Back	Cheek	Palm	*Average*
$p_m(\%)$	38.5	21.4	33.1	42.8	13.3	29.8
$p_c(\%)$	22.3	17.8	18.7	23.7	16.3	19.8
$p_{bv}(\%)$	1.32	1.34	1.31	1.38	1.35	1.34
$p_{os}(\%)$	70.0	73.9	76.9	73.3	86.3	76.1

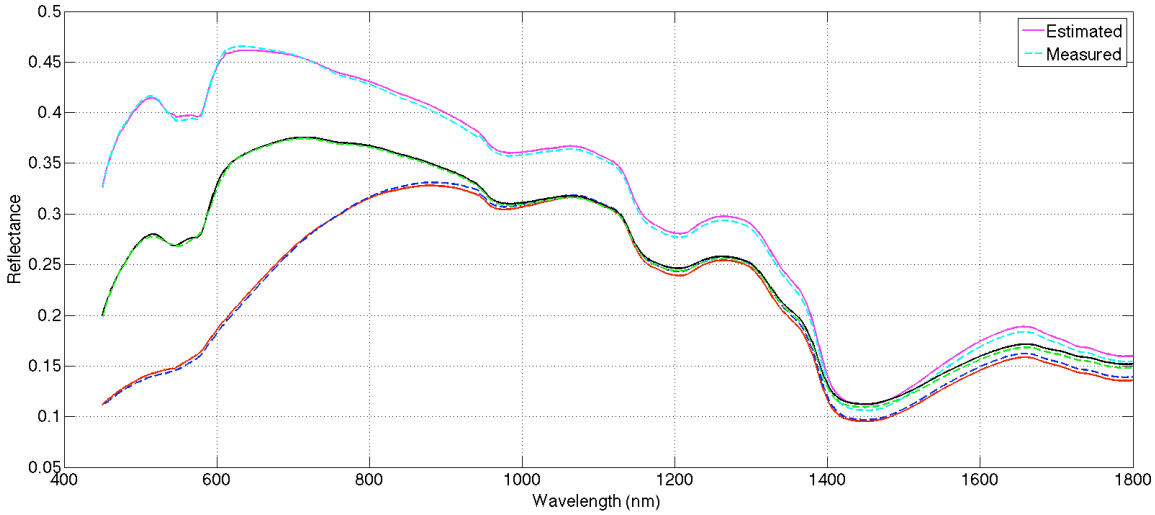


Figure 4.5: Plotted, as examples, are three of the 241 in vivo signatures along with their estimated signatures using the inverse method (SVR). The estimated parameters for each example are: $[p_m, p_c, p_{bo}, p_{bl}] = [9\%, 24\%, 82\%, 1.4\%]$ for the top trace, [17%, 19%, 83%, 0.86%] for the middle trace, and [29%, 31%, 79%, 0.68%] for the bottom trace. Each trace corresponds to a different ethnicity.

Table 4.3: Root Mean Square Error (RMSE) and Spectral Angle Error (Φ) in Radians

Pt.	**Eth.**	Φ	**RMSE**	**SD**
7	Cau.	0.135	0.036	0.002
14	Cau.	0.138	0.062	0.003
15	Cau.	0.139	0.034	0.006
17	Cau.	0.140	0.041	0.006
18	Cau.	0.140	0.041	0.002
19	Cau.	0.140	0.043	0.002
20	Cau.	0.142	0.034	0.003
21	Cau.	0.142	0.036	0.002
22	Cau.	0.143	0.066	0.002
23	Cau.	0.144	0.037	0.004

Pt.	**Eth.**	Φ	**RMSE**	**SD**
2	Asn.	0.131	0.048	0.003
3	Asn.	0.132	0.076	0.008
4	Asn.	0.133	0.072	0.007
5	Asn.	0.133	0.055	0.005
6	Asn.	0.134	0.086	0.009
10	Asn.	0.137	0.036	0.002
13	Asn.	0.138	0.044	0.003
24	Asn.	0.144	0.063	0.002

Pt.	**Eth.**	Φ	**RMSE**	**SD**
1	Afr. Am.	0.129	0.109	0.012
8	Afr. Am.	0.135	0.046	0.003
9	Afr. Am.	0.136	0.062	0.006
11	Afr. Am.	0.137	0.056	0.005
12	Afr. Am.	0.137	0.040	0.003
16	Afr. Am.	0.140	0.062	0.002
Average	n/a	0.137	0.0445	0.0097

Table 4.4: Spectral Angle Error (in Radians) for Anatomical Locations

Pt.	**Eth.**	**DF**	**UIA**	**Back**	**Cheek**	**Palm**
7	Cau.	0.130	0.131	0.137	0.134	0.138
14	Cau.	0.134	0.136	0.141	0.135	0.142
15	Cau.	0.135	0.137	0.143	0.135	0.142
17	Cau.	0.135	0.139	0.144	0.135	0.142
18	Cau.	0.136	0.139	0.145	0.136	0.143
19	Cau.	0.136	0.140	0.145	0.137	0.144
20	Cau.	0.137	0.141	0.145	0.137	0.145
21	Cau.	0.137	0.141	0.146	0.139	0.145
22	Cau.	0.138	0.141	0.147	0.139	0.147
23	Cau.	0.138	0.141	0.147	0.139	0.149

Pt.	**Eth.**	**DF**	**UIA**	**Back**	**Cheek**	**Palm**
1	Afr. Am.	0.129	0.125	0.134	0.129	0.123
8	Afr. Am.	0.131	0.134	0.138	0.134	0.138
9	Afr. Am.	0.131	0.135	0.138	0.134	0.138
11	Afr. Am.	0.132	0.135	0.139	0.134	0.141
12	Afr. Am.	0.133	0.135	0.139	0.135	0.142
16	Afr. Am.	0.135	0.139	0.143	0.135	0.142

Pt.	**Eth.**	**DF**	**UIA**	**Back**	**Cheek**	**Palm**
2	Asn.	0.129	0.127	0.135	0.129	0.133
3	Asn.	0.130	0.129	0.136	0.130	0.134
4	Asn.	0.130	0.129	0.136	0.130	0.134
5	Asn.	0.130	0.130	0.136	0.131	0.137
6	Asn.	0.130	0.130	0.137	0.133	0.137
10	Asn.	0.132	0.135	0.139	0.134	0.140
13	Asn.	0.133	0.136	0.140	0.135	0.142
24	Asn.	0.138	0.142	0.149	0.140	0.150
Avg.		0.1338	0.1358	0.1412	0.1350	0.1407

to obtain the ground truth. We will have more to say about this in Chapter 5. For the case of synthetic skin signatures, the task is simpler, because the ground truth is readily available. It is apparent from Table 4.1, given the results, that the inverse mapping performed as desired. The task for *in vivo* signatures is much more difficult because the ground truth is no longer available. Many studies have presented novel inverse methods, but to the best of our knowledge, no studies have augmented their inverse methods with biopsies to ascertain ground truth. A potential validation scheme has often been to check the estimated parameters against well-established physiological precepts. This scheme arose from exploring physiological features that are specific to human skin. For example, as noted by Zonois et al.

and Nunez et al. [16, 73], melanin is directly responsible for skin color. Therefore, we expect African American subjects to have a larger melanosome concentration than Asian or Caucasian subjects. In a similar manner, there are other physiological precepts that can be checked to ensure that they fit within the realm of physiological plausibility. While this does not guarantee the estimates are accurate, it does offer a good performance validation criteria.

Following the precept that skin color correlates with melanosome concentration, it was observed in the first row of Table 4.2 and the green bars in Figure 4.4 that for each anatomical location, the relative concentration of p_m was the largest for African Americans (t-test p-value 0.0105) than for Asians or Caucasians (see purple cells in Table 4.2). Furthermore, we would expect higher p_m concentration in the dorsal forearm and the cheeks, than the palm or the back because these areas are naturally exposed to greater sunlight and hence are inherently more tan. It is clear from Figure 4.4 and Table 4.2 (gray cells) that for caucasian patients, the average p_m concentration was highest in the cheeks at 28.5% and the DF at 21.2%, similarly for Asian and African Americans it was highest in cheeks and the DF at 23.5% and 22.6% and 42.8% and 38.5% respectively. It is also noteworthy that comparing only cheeks and DF, the relative p_m concentrations were still highest for African Americans followed by Asians and then Caucasian patients. Physiology dictates that we should expect lower p_m concentrations in the palm than the other four areas imaged. It is clear in Figure 4.4 and Table 4.2 (cyan cells) that the lowest p_m concentration was in

the palm than any other area. There are a number of studies that estimate p_m, and our results are consistent with many of those studies. As such, Zhai et al. [74] reported p_m concentration of approximately 15% for Caucasians. Our estimate of approximately 18% was in agreement, where the 3% deviation is negligible, considering the variability in skin tone (e.g. tanning, etc.) amongst individuals for each ethnicity.

Typically, collagen concentration is higher in the cheeks and the dorsal forearm than other anatomical locations [74, 75]. It can be seen in Figure 4.4 and Table 4.2 (green cells) that the largest collagen concentration for each ethnicity was in the cheeks followed by the DF (t-test p-value 0.0491). There is no pattern for collagen based on ethnicity. We noted that the relative collagen concentration between ethnicities remained constant (t-test p-value 0.0334). All patients involved in this study were healthy, so we would expect oxygen saturation concentrations to be greater than 70%. Table 4.2 (red cells) shows that the relative oxygen saturation concentrations were above 70%. Mortimer et al. [76], Pilon et al. [36] and Tuchin et al. [20] have noted that blood volume varies based on anatomical location. It is expected that the blood volume is larger in the cheeks than the DF or the UIA. These precepts are consistent with what was observed in Table 4.2 (orange cells) and Figure 4.4 (t-test p-value 0.0217). In summary, the concentrations for all estimated parameters should be physiologically meaningful, i.e. within the acceptable physiological range as outlined in Table 2.1. This requirement was satisfied as seen by comparing Tables 2.1 and 4.2. While these results were encouraging (and statistically significant), it is important to

note that while all efforts were made to collect hyperspectral data from macroscopically homogeneous skin, a minority of patients had freckles and other types of hyper pigmentation (benign nevi, etc). This prevented data collection from a completely homogeneous patch of skin. It is possible that these pigmentations contribute some degree of error to our results.

Another potential method to validate the estimated parameters is to test the inverse-forward modeling loop itself. In other words, the measured ground truth spectra can be compared to the estimated spectra. While this does not guarantee that the estimated parameters are accurate, it provides some encouraging evidence that the inverse-forward modeling loop works. In this regard, the better the forward model, the higher the probability that the estimated parameters are accurate. Figure 4.5 shows that the estimated spectra and the ground truth spectra were in good agreement with each other. This performance validation metric was further quantified in Tables 4.4 and 4.5, which provide the spectral angles and RMSE for each patient and each anatomical location. While no claims can be made about the exact accuracy of the estimated parameters, the performance validation done through physiological precept analysis and comparisons of the measured and estimated spectra, provided encouraging evidence that the estimated parameters fall within the realm of physiological plausibility.

An important consideration in this study was that we fixed the the water percentage in the forward model. The percentage of water in each layer of skin can vary based

on anatomical location, age, health, etc. In this study, we chose the values tabulated by Meglinski [7] and Nunez [16], which were obtained based on weighted population averages for these particular anatomical locations. It appears that this choice still led to reasonable errors on the modeled signature (see Table 4.4 and Figure 4.4) as well as the estimated underlying parameters for our synthetic experiments (see Table 2). However, taking into account varying water percentages and performing sensitivity analysis are important endeavors which we intend to pursue in future studies. It must be noted that our approach can be extended to estimate the water percentage of each of the nine layers of skin. Indeed our inverse methods based on machine learning can trivially extend to handle these extra parameters.

In analyzing our results, it can be seen in Table 4.1 that oxygen saturation concentrations exhibited larger errors than the other parameters. A classic problem that was discussed recently by Nishidate et al. [78] has been that of deoxyhemoglobin overestimation in regions with a high concentration of melanosomes. The results in this study seemed to echo these findings, especially in Table 4.1 when comparing the oxygen saturation in African Americans for high melanin regions, and even so in Asian and Caucasian subjects. These trends were consistent even when compared to non-machine learning studies, such as Pilon et al. [36] or others discussed above. Similarly, this may explain why the errors in Table 4.1 for oxygen saturation were larger since the synthetic experiments were composed of a wide range of melanosome concentrations. In this regard, our study can serve as a springboard for further in-

vestigating this issue. In particular, as a baseline, our synthetic experiments could be redone, where the error in oxygen saturation is plotted as a function of increasing melanosome concentration in the forward model. This is potential future work.

While we cannot compare our results directly to other studies because it is not a one-to-one comparison, i.e. we estimated different parameters, on a different dataset, using a different methodology, we can however compare our methodologies and validation metrics. Here we provide a summary of this comparison for a few closely related studies. Pilon et al. [36] used a two-layer optical model and simulate reflectance from a semi-empirical model based on an approximation of the RTE using Monte Carlo methods, and an inverse method based on least squares minimization. They used a subset of the same physiological precepts as presented here to judge their performance. The estimated blood level (ranging from approximately 1% - 2% by volume in the dermis) was in agreement with this study. While the melanin concentration was not provided as a fraction per volume, if average density for melanin is assumed [77], the estimates were consistent (ranging from approximately 20% - 40% by volume in the epidermis). Claridge, Preece et al. [30, 32, 33] used a physics-based skin coloring model to find a mapping between digital images to histological parameters. Furthermore, they used spectral filter optimization to reduce error. They perform validation by computing errors associated with their RGB and optimal filters. This is not too different from comparing modeled and estimated signatures, as was done in this study. Effectively, the ground truth is being compared with its estimate in another vector space (rather

than in the parameter space). Claridge et al. also employ a modified K-M model with four layers (in contrast to the ten layers used in this study). The physical values of the parameters were not reported, so no further comparisons can be made. Finally, Doi and Tominaga [34] also used a K-M based model and a least squares approach to minimize the error between modeled and estimated spectra. Their analysis is similar to the metric used here to compare estimated spectra to the ground truth. They, however, used least squares to find weights for each parameter, rather than reporting their physical values. Therefore, a direct comparison is not possible. In summary, we showed that the skin parameters from *in vivo* patients can be non-invasively estimated using our machine learning methods. We reported acceptable accuracy as per our validation metrics and the metrics used by aforementioned other studies.

Finally, an important consideration before such a method can be translated into clinical use is the computational complexity. This study employed a single point spectroscopic system. For this system, our model and algorithm implemented in Matlab R2012b (Natick, MA, USA) running on an Intel i3-3225 CPU with 8GB DDR3-1600 RAM, took approximately 30 seconds to estimate the parameters for each signature. It is important to note that the code was not optimized to take advantage of the parallel architecture of modern CPUs. In the future, we would like to employ a hyperspectral imager. While this would mean that each pixel would require 30 seconds to analyze, one could implement this system into C++, or some other compiled language so that the computation time is reduced by a factor of 10-

20 based on the number of threads on the CPU (a current inexpensive CPU has 8 threads). GPU-based methods are also becoming wide-spread, and hence is also a potential avenue for development.

Chapter 5

Ground Truth via Ultrasound Imaging

In Chapter 4 we described how our machine learning method was used to estimate skin parameters. We showed very good agreement with the ground truth in our synthetic experiments (error of 0.05±10e-3%), and found similar agreement in the *in vivo* experiments when comparing to physiological precepts (p-values < 0.05). We discussed that a major challenge in the skin optics community has been the task of validating automated methods for skin parameter estimation. To our knowledge no dataset currently exists that contain both HSI signatures and corresponding ground truth parameters. Physiological precept analysis remains the only accepted metric to judge plausibility of automated methods.

There exist a number of dermatological applications that require accurate esti-

mates for skin parameters. One of the most common such parameters is the skin thickness; this is directly linked to pathology, and is crucial information for many drug delivery applications, or in wound care applications where non-contact estimation is critical. The current method for estimating thickness requires one to excise a sample of skin through a biopsy and then study it under a microscope. This is clearly not a good option when this needs to be repeated for multiple sites, or on multiple occasions. Non-invasively estimating skin thickness is very challenging because skin varies a great deal from one anatomical location to another. It also varies based on ethnicity, environmental factors, and age, amongst others. While many researchers have proposed methods for estimating skin thickness, they seldom have access to any gold standard. In such cases the best that can be done is comparisons to a database to test for physiological plausibility. Skin thickness for each layer can change on the micron scale. Therefore comparing to a database with a range of values cannot truly benchmark the methods. In this thesis, we discuss a novel method using Ultrasound imaging to validate the estimation of skin thickness. This method is discussed next.

5.1 Ultrasound Imaging

Ultrasound (US) refers to sound waves with frequencies in the Megahertz (MHz) range. A typical medical ultrasound imaging system operates in the 4-10 MHz range. US imaging, or ultrasonography, at a high level, works by generating a sound wave

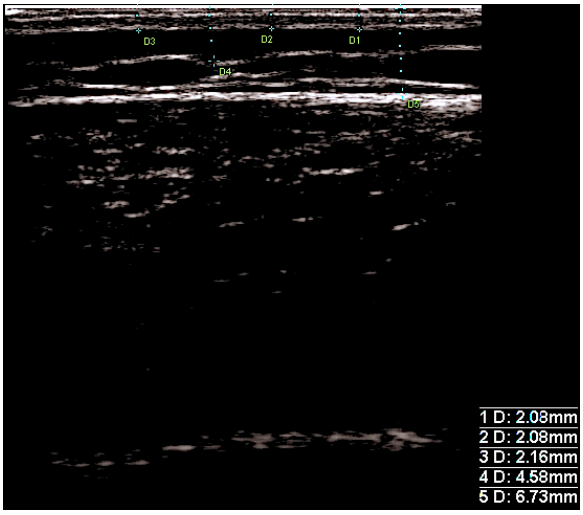


Figure 5.1: An ultrasound scan acquired from the dorsal forearm at the Johns Hopkins University, Department of Dermatology under IRB protocol.

using piezoelectric elements encased within a transducer. These sound waves enter the body and are selectively attenuated by tissues, bones, etc., and return back to the transducer in the form of echoes. These echoes are captured by the transducer in the form of vibrations, which are then converted into electrical impulses and then into a digital image. This process can occur in 2D, 3D, or 4D (volume + time).

One of the many advantages of US include the fact that it is real time; an operator can adjust the probe dynamically and scan different areas based on what he/she would like to image. More importantly, US is quite robust at imaging soft tissues, muscles, and bones, and in particular it is very good at delineating the boundary between these tissue types. This is quite important in our application because it makes it easy to delineate the boundary between skin, and muscle. Figure 5.1 is an example of an US

scan acquired *in vivo* of a human dorsal forearm. Notice the clear delineations between different skin layers and tissue types. Therefore, given an US scan of an anatomical location, a trained physician can clearly indicate and measure the thickness of skin or some other anatomical feature (as shown for example in Figure 5.1).

5.2 Experimental Methodology

In Equation (4.1) we defined a training dataset that was used for our machine learning algorithm. We also defined in Equation (4.3) a testing *in vivo* dataset of HSI signatures that was acquired under an IRB protocol at the Johns Hopkins Hospital from 24 patients. As implied in the previous section, under an amended IRB protocol, we also acquired another dataset of US scans from the same anatomical locations as the dataset in Equation (4.3). Therefore we now have a corresponding dataset to that of Equation (4.3). This dataset is called the Validation Dataset, and is given by:

$$Validation\ Dataset : \{US_j; j = 1,\ \ldots,\ M = 24\} \tag{5.1}$$

In order to use the Validation Dataset, we asked the physician on our team to annotate the US sequences, marking and computing the thickness. This can then effectively be used as gold standard because skin is visible in the sequences, and a medical expert can use it to ascertain the thickness by delineating the boundary between skin and subcutaneous fat. An example of such an annotation is included in Figure 5.1.

In this work, only one physician was asked to perform the delineations and hence

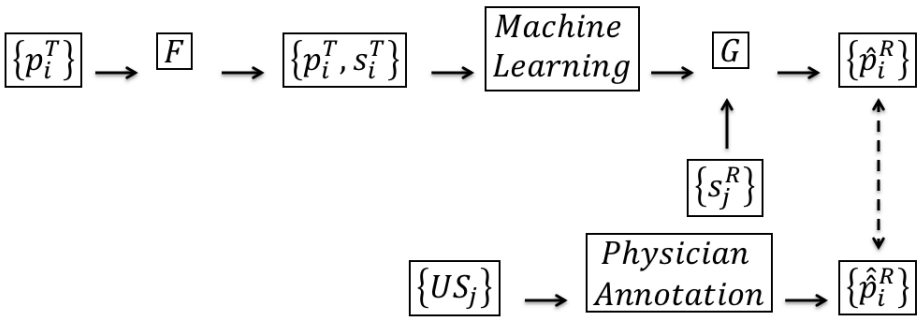


Figure 5.2: Flowchart of the experimental set-up.

no intra-observer error can be determined. However, the physician repeated the annotations three times for each sequence. We computed the intra-observer error as a standard deviation between the three estimates, which had a value of 10e-4. Given this result, the physician annotation was assumed to be the average of all three trials. Finally, a flowchart of our methodology is provided in Figure 5.2.

5.3 Results

As error analysis, we computed the mean error (Average), absolute error (A.E.), and the standard deviation (SD) between the machine learning estimated thicknesses, $\hat{\mathbf{p}}_i^R$, and the gold standard physician estimated thicknesses, $\hat{\hat{\mathbf{p}}}_i^R$. The results (total thickness) for each sample are tabulated in Table 5.1.

Note that for each sample, the resultant error and deviation were quite small. This suggests that the method is capable of automatically estimating the thickness for a

variety of skin types and anatomical locations with good accuracy. Furthermore, we would like to highlight the efficacy of our methods by citing the fact the the 5th-percentile, 95th-percentile, and standard deviations were quite small (mean error and deviations of 0.09±0.10 mm); this suggests that the method is robust despite differences in age, skin type, or ethnicity.

5.4 Discussion

A natural question one could pose against this method is why perform the machine learning or the modeling if US can solve the problem. The answer is:

1. Our modeling methodology not only estimated skin thickness, but also all the other skin parameters (e.g., melanin, collagen, etc.), therefore unless a physician is only interested in one skin parameter, US is not the ideal solution;

2. Under pathology it is rarely the case that only one parameter is affected, for example in melanoma, melanin, blood volume, and dermal thickness are all affected, therefore only analyzing one parameter is not sufficient;

3. Finally, in cases of burns or wounds, the thickness of skin is a critical component needed to better decide how much and where to deliver skin-based drugs. In such cases, contacting the skin with US gel and the probe is not desirable as it could lead to complications and patient discomfort.

Table 5.1: The estimated skin thickness and the statistics.

Pt.	Est.	Ref.	A.E.	Pt.	Est.	Ref.	A.E.
1	2.21mm	2.08mm	0.13mm	13	1.65mm	1.49mm	0.16mm
2	1.55mm	1.67mm	0.12mm	14	1.75mm	1.78mm	0.03mm
3	3.77mm	3.73mm	0.04mm	15	6.22mm	6.14mm	0.08mm
4	2.01mm	1.91mm	0.10mm	16	4.62mm	4.65mm	0.03mm
5	3.55mm	3.65mm	0.10mm	17	2.57mm	2.44mm	0.13mm
6	3.40mm	3.41mm	0.01mm	18	2.84mm	2.78mm	0.06mm
7	2.60mm	2.49mm	0.11mm	19	4.37mm	4.27mm	0.10mm
8	7.86mm	7.80mm	0.14mm	20	5.81mm	5.69mm	0.12mm
9	6.83mm	6.97mm	0.14mm	21	7.49mm	7.48mm	0.01mm
10	5.99mm	5.98mm	0.01mm	22	2.54mm	2.48mm	0.06mm
11	5.45mm	5.31mm	0.14mm	23	3.54mm	3.44mm	0.10mm
12	5.14mm	5.23mm	0.09mm	24	3.84mm	3.79mm	0.05mm

Statistic	Est.	Ref.	A.E.
Mean	4.07mm	4.02mm	0.09mm
Median	3.66mm	3.69mm	0.10mm
St. Dev.	1.90mm	1.91mm	0.05mm
5% Value	1.62mm	1.61mm	0.01mm
95% Value	7.60mm	7.55mm	0.15mm

Chapter 6

Conclusion

6.1 Thesis contributions

Skin related conditions such as melanoma affects the lives of hundreds of thousands of people worldwide. Most of these cases can be treated with minimal local surgery, and often yield a good prognosis. Unfortunately, in most cases the tumor is not detected early, and hence metastasizes and spreads throughout the body, reducing the chances of survival quite significantly. Therefore, there is a need for a robust screening technology that can help detect these cancers early. This thesis focused on the development of a physics-based model of skin coupled with hyperspectral imaging, and machine learning to help that goal along. The developmental contributions of this thesis were three fold:

1. A physics-based forward model to map skin parameters to an expected multi-

band (hyperspectral) reflectance signature,

2. A machine learning based inverse model to regress a hyperspectral signature into constitutive skin parameters, and

3. An Ultrasound-based performance evaluation methodology to test model efficacy for a clinically relevant task.

Physics-based Forward Model

In order to recover the skin components from a hyperspectral reflectance signature, we needed to have a model for that reflectance measurement parameterized by the skin components. This was accomplished by using *Kubelka-Munk* Theory, and a light transport model, which allowed us to express all possible paths light could take when it was incident on a multi-layered surface (such as human skin). These set of theories model light interactions, and can express absorbances, and transmissions based on the optical properties of the materials the light interacts with. Therefore, we can mathematically describe the amount of light we expect to reflect from the surface of the skin. This reflected light is akin to the reflectance measurement reported by a HSI sensor. We showed in this thesis, e.g., Figure 4.5, that we were able to develop such a physics-based model, which could approximate *in vivo* signatures acquired in a clinic, with reliable agreement.

Machine Learning Inverse Model

In a clinical setting we were interested in recovering the skin components from a corresponding hyperspectral signature. This can be posed as a regression problem, and we used machine learning to solve it. The forward model allowed us to generate a vast *labeled* dataset of tuples of skin parameters and hyperspectral signatures. This allowed us to proceed with a supervised learning framework. We trained a support vector regression machine using the labeled tuples obtained from the forward model. This in-effect instantiated the inverse model; in other words a mapping from spectral space to parameter space. We tested this model on both synthetic and *in vivo* signatures, and found good agreement with synthetic ground truth, and well established physiological precepts, as discussed in Section 4.3.

Ultrasound Validation

Many studies have attempted to develop methods to estimate *in vivo* skin parameters, but most suffer from the inability to evaluate their methods against ground truth. The ground truth in this context refers to the exact concentrations of the biological materials in the sample of skin being imaged. One way to obtain this ground truth would be through a biopsy of a skin sample, and then examination of the sample under a microscope to ascertain the true value of each biological material. However this is quite invasive, and hence not the best course of action.

One skin parameter of interest in dermatology is the thickness of the skin. This parameter is critical for drug delivery applications, as well as diagnostically for conditions such as bed sores. We proposed the use of a clinical Ultrasound system to measure the thickness of skin. Ultrasound is robust at delineating the boundaries between different tissue types, for example, skin and bone. We demonstrated that a dermatologist was able to effectively use Ultrasound to mark the delineation of skin and bone on an image, and since the depth of penetration is known, we obtained the thickness of the skin. We could then compare the estimates from Ultrasound, to those we obtained from our modeling methodology. We demonstrated that our method was quite robust at estimating the skin thickness as shown in Table 5.1.

6.2 Future work

Many studies have suggested that the temporal analysis of biological skin parameters, as could be done in a longitudinal study, is critical in detecting the onset of various skin related cancers. While we demonstrated that the skin parameters could be estimated reliably, there is much work that still needs to be done before a screening technology can be translated into a clinical setting.

1. First, we demonstrated the performance of our methods on healthy patients. While this is an early step in the right direction, we now need to evaluate our methods on patients with skin abnormalities. We need to ensure that the

presence of certain disorders does not bias the estimation.

2. In parallel to evaluating our methods on patients with skin conditions, we could also start longitudinal studies. A potential way to do this would be through translation into clinics that perform bi/annual physical exams. For example, we could evaluate how the concentration of melanin changes over time. In the summer the skin is exposed to the sun a great deal more than it is during the winter; some studies cite that there is a measurable biochemical change, which we posit could be detected with our methodology. This is of course speculative, and would need to be investigated.

3. Through our current IRB protocol, we could start collecting data from current skin cancer patients. This could be an opportunity to obtain biopsies from these patients. These biopsies could help us obtain "true" ground truth, against which we could compare our methods.

4. While using Ultrasound is a good way to compare the estimate of skin thickness, Optical Coherence Tomography (OCT), has also been recently used in dermatology applications. OCT has very high spatial resolution, and could hence provide a more robust estimate than Ultrasound. Figure 6.1 shows a typical OCT scan acquired from the dorsal hand.

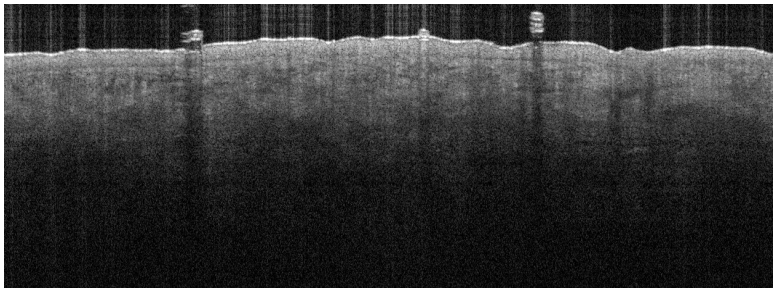


Figure 6.1: An OCT scan acquired from the dorsal hand. Photograph courtesy of Dr. Jon Meyerle from the Johns Hopkins University, Dept. of Dermatology.

6.3 Summary

We were able to demonstrate in this thesis that the biological components that make up the layers of the skin can be estimated with high accuracy. This was done both through synthetic experiments, and *in vivo* data collected from 24 Caucasian, Asian, and African American patients, under IRB at the Johns Hopkins Hospital. We found that our results were not biased based on age, sex, or ethnicity (skin type). We were also able to propose the use of clinical Ultrasound for a proxy ground truth to evaluate the performance of skin thickness estimation. We believe our work in this thesis is both a valuable and a necessary first step towards developing a skin cancer screening technology.

Bibliography

[1] Wikipedia, "Wikipedia: Melanoma," 2014. [Online]. Available: http://en.wikipedia.org/wiki/Melanoma

[2] ——, "Wikipedia: Cones," 2014. [Online]. Available: http://en.wikipedia.org/wiki/Cones

[3] ——, "Wikipedia: Hyperspectral imaging," 2014. [Online]. Available: http://en.wikipedia.org/wiki/Hyperspectral

[4] J. Plaza, A. J. Plaza, and C. Barra, "Multi-channel morphological profiles for classification of hyperspectral images using support vector machines," *Sensors*, vol. 9, no. 1, pp. 196–218, 2009.

[5] "American cancer society: Cancer facts and figures," 2010.

[6] I. Meglinski and S. Matcher, "Modeling of skin reflectance spectra," in *Proceedings of SPIE*, vol. 4241, 2001, p. 78.

[7] ——, "Quantitative assessment of skin layers absorption and skin reflectance

spectra simulation in the visible and near-infrared spectral regions," *Physiological measurement*, 2002.

[8] E. Linos, S. Swetter, M. Cockburn, G. Colditz, and C. Clarke, "Increasing burden of melanoma in the united states," *Journal of Investigative Dermatology*, vol. 129, no. 7, pp. 1666–1674, 2009.

[9] M. Xing, "Braf mutation in papillary thyroid microcarcinoma: the promise of better risk management," *Annals of surgical oncology*, vol. 16, no. 4, pp. 801–803, 2009.

[10] H. Rogers, M. Weinstock, A. Harris, M. Hinckley, S. Feldman, A. Fleischer, and B. Coldiron, "Incidence estimate of nonmelanoma skin cancer in the united states, 2006," *Archives of dermatology*, 2010.

[11] "National center for chronic disease prevention and health promotion: The burden of skin cancer," 2010.

[12] E. Wurm and H. Soyer, "Scanning for melanoma," *Volume 33 Number 5 An Independent Review, OCTOBER 2010.*, vol. 5, p. 150.

[13] G. Zonios, A. Dimou, I. Bassukas, D. Galaris, A. Tsolakidis, and E. Kaxiras, "Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection," *Journal of biomedical optics*, vol. 13, p. 014017, 2008.

[14] D. Dicker, J. Lerner, P. Van Belle, S. Barth, M. Herlyn, D. Elder, and W. El-Deiry, "Differentiation of normal skin and melanoma using high resolution hyperspectral imaging," *Cancer biology & therapy*, 2006.

[15] M. Van Gemert, S. Jacques, H. Sterenborg, and W. Star, "Skin optics," *Biomedical Engineering, IEEE Transactions on*, 1989.

[16] A. Nunez, "A physical model of human skin and its application for search and rescue," DTIC Document, Tech. Rep., 2009.

[17] G. Baranoski and A. Krishnaswamy, "Light interaction with human skin: from believable images to predictable models," in *ACM SIGGRAPH ASIA 2008 courses*. ACM, 2008, p. 27.

[18] P. Kubelka and F. Munk, "An article on optics of paint layers," *Z. Tech. Physik*, vol. 12, pp. 593–601, 1931.

[19] M. Van Gemert and W. Star, "Relations between the kubelka-munk and the transport equation models for anisotropic scattering," *Lasers Life Sci.*

[20] V. Tuchin, S. Utz, and I. Yaroslavsky, "Tissue optics, light distribution, and spectroscopy," *Optical Engineering*, 1994.

[21] G. Yoon, "Absorption and scattering of laser light in biological media-mathematical modeling and methods for determining the optical properties," Texas Univ., Austin, TX (USA), Tech. Rep., 1988.

[22] A. Ishimaru, *Wave propagation and scattering in random media.* Wiley-IEEE Press, 1999, vol. 12.

[23] S. Prahl, "Light transport in tissue," Ph.D. dissertation, University of Texas at Austin, 1988.

[24] H. Van de Hulst, "Multiple light scattering: Tables," *Formulas, and Applications (Academic, New York, 1980)*, vol. 1, pp. 477–492, 1980.

[25] D. Churmakov, I. Meglinski, S. Piletsky, and D. Greenhalgh, "Analysis of skin tissues spatial fluorescence distribution by the monte carlo simulation," *Journal of Physics D: Applied Physics*, 2003.

[26] M. Shimada, Y. Yamada, M. Itoh, and T. Yatagai, "Melanin and blood concentration in a human skin model studied by multiple regression analysis: assessment by monte carlo simulation," *Physics in Medicine and Biology*, vol. 46, p. 2397, 2001.

[27] P. Hanrahan and W. Krueger, "Reflection from layered surfaces due to subsurface scattering," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques.* ACM, 1993.

[28] C. So-Ling and L. Li, "A multi-layered reflection model of natural human skin," in *Computer Graphics International 2001. Proceedings.* IEEE, 2001, pp. 249–256.

[29] J. Stam, "An illumination model for a skin layer bounded by rough surfaces," in *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*. Springer-Verlag, 2001, pp. 39–52.

[30] S. Cotton, "A noninvasive skin imaging system," *School of Computer Science Research Reports - University of Birmingham CSR*, 1997.

[31] S. Cotton and E. Claridge, "Developing a predictive model of human skin colouring," in *Proceedings of SPIE Medical Imaging*, 1996.

[32] E. Claridge and S. Preece, "An inverse method for the recovery of tissue parameters from colour images," in *Information Processing in Medical Imaging*, 2003.

[33] S. Preece and E. Claridge, "Spectral filter optimization for the recovery of parameters which describe human skin," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2004.

[34] M. Doi and S. Tominaga, "Spectral estimation of human skin color using the kubelka-munk theory," in *Proceedings of SPIE*, 2003.

[35] D. Yudovsky and L. Pilon, "Rapid and accurate estimation of blood saturation, melanin content, and epidermis thickness from spectral diffuse reflectance," *Applied optics*, 2010.

[36] ——, "Retrieving skin properties from in vivo spectral reflectance measurements," *Journal of biophotonics*, 2011.

[37] R. Anderson and J. Parrish, "The optics of human skin," *Journal of Investigative Dermatology*, vol. 77, no. 1, pp. 13–19, 1981.

[38] S. Wan, R. Anderson, and J. Parrish, "Analytical modeling for the optical properties of the skin with in vitro and in vivo applications," *Photochemistry and Photobiology*, vol. 34, no. 4, pp. 493–499, 1981.

[39] N. Tsumura, H. Haneishi, and Y. Miyake, "Independent-component analysis of skin color image," *JOSA A*, 1999.

[40] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake, "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin," in *ACM Transactions on Graphics (TOG)*, vol. 22, no. 3. ACM, 2003, pp. 770–779.

[41] N. Tsumura, M. Kawabuchi, H. Haneishi, and Y. Miyake, "Mapping pigmentation in human skin from a multi-channel visible spectrum image by inverse optical scattering technique," *Journal of Imaging Science and Technology*, vol. 45, no. 5, pp. 444–450, 2001.

[42] E. Claridge, S. Cotton, P. Hall, and M. Moncrieff, "From colour to tissue histol-

ogy: Physics-based interpretation of images of pigmented skin lesions," *Medical Image Analysis*, vol. 7, no. 4, pp. 489–502, 2003.

[43] I. Kaartinen, P. Välisuo, V. Bochko, J. Alander, and H. Kuokkanen, "How to assess scar hypertrophya comparison of subjective scales and spectrocutometry: A new objective method," *Wound Repair and Regeneration*, 2011.

[44] L. Leonardi, M. Sowa, J. Payette, and H. Mantsch, "Near-infrared spectroscopy and imaging: a new approach to assess burn injuries," *American Clinical Laboratory*, vol. 19, no. 8, pp. 20–22, 2000.

[45] J. Payette, E. Kohlenberg, L. Leonardi, A. Pabbies, P. Kerr, K. Liu, and M. Sowa, "Assessment of skin flaps using optically based methods for measuring blood flow and oxygenation," *Plastic and reconstructive surgery*, vol. 115, no. 2, p. 539, 2005.

[46] S. Prigent, X. Descombes, D. Zugaj, P. Martel, and J. Zerubia, "Multi-spectral image analysis for skin pigmentation classification," 2010.

[47] S. Prigent, X. Descombes, D. Zugaj, and J. Zerubia, "Spectral analysis and unsupervised svm classification for skin hyper-pigmentation classification," 2010.

[48] D. Yudovsky, A. Nouvong, and L. Pilon, "Hyperspectral imaging in diabetic foot wound care," *Journal of Diabetes Science and Technology*, vol. 4, no. 5, p. 1099, 2010.

[49] S. Dreiseitl, L. Ohno-Machado, H. Kittler, S. Vinterbo, H. Billhardt, and

M. Binder, "A comparison of machine learning methods for the diagnosis of pigmented skin lesions," *Journal of Biomedical Informatics*, 2001.

[50] J. Nyström, B. Lindholm-Sethson, L. Stenberg, S. Ollmar, J. Eriksson, and P. Geladi, "Combined near-infrared spectroscopy and multifrequency bio-impedance investigation of skin alterations in diabetes patients based on multivariate analyses," *Medical and Biological Engineering and Computing*, vol. 41, no. 3, pp. 324–329, 2003.

[51] M. Sowa, L. Leonardi, J. Payette, J. Fish, and H. Mantsch, "Near infrared spectroscopic assessment of hemodynamic changes in the early post-burn period," *Burns*, vol. 27, no. 3, pp. 241–249, 2001.

[52] S. Vyas, H. Van Nguyen, P. Burlina, A. Banerjee, L. Garza, and R. Chellappa, "Computational modeling of skin reflectance spectra for biological parameter estimation through machine learning," in *Proceedings of SPIE*, vol. 8390, 2012, p. 83901B.

[53] S. Vyas, A. Banerjee, L. Garza, S. Kang, and P. Burlina, "Hyperspectral signature analysis of skin parameters," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2013, pp. 867 002–867 002.

[54] S. Vyas, A. Banerjee, and P. Burlina, "Machine learning methods for in vivo skin parameter estimation," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. IEEE, 2013, pp. 524–525.

[55] ——, "Estimating physiological skin parameters from hyperspectral signatures," *Journal of biomedical optics*, vol. 18, no. 5, pp. 057 008–057 008, 2013.

[56] V. Vapnik, S. Golowich, and A. Smola, "Support vector method for function approximation, regression estimation, and signal processing," in *Advances in Neural Information Processing Systems 9*, 1996.

[57] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[58] A. Navot, L. Shpigelman, N. Tishby, and E. Vaadia, "Nearest neighbor based feature selection for regression and its application to neural activity."

[59] C.-C. Chang and C.-J. Lin, "Libsvm : a library for support vector machines. acm transactions on intelligent systems and technology," 2011.

[60] A. Smola and B. Scholkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[61] R. Fletcher, "Practical methods of optimization."

[62] O. Mangasarian, *Nonlinear programming*. Society for Industrial Mathematics, 1994, vol. 10.

[63] G. McCormick, "Nonlinear programming: Theory, algorithms, and applications." *John Wiley & Sons, Inc., 605 3rd Ave., New York, 1983, 464*, 1983.

[64] R. J. Vanderbei, "Loqo users manualversion 3.10."

[65] W. Karush, "Minima of functions of several variables with inequalities as side constraints," *Master's thesis, Department of Mathematics, University of Chicago*, 1939.

[66] H. Kuhn and A. Tucker, "Nonlinear programming," in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, vol. 5. California, 1951.

[67] S. Keerthi, S. Shevade, C. Bhattacharyya, and K. Murthy, "Improvements to platt's smo algorithm for svm classifier design," *Neural Computation*, vol. 13, no. 3, pp. 637–649, 2001.

[68] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[69] R. Collobert and S. Bengio, "Svmtorch: Support vector machines for large-scale regression problems," *The Journal of Machine Learning Research*, vol. 1, pp. 143–160, 2001.

[70] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, 1998.

[71] M. Tommola, M. Tynkknyen, J. Lemmetty, P. Harstela, and L. Sikanen, "Estimating the characteristics of a marked stand using k-nearest-neighbour regression," *International Journal of Forest Engineering*, vol. 10, no. 2, 1999.

[72] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 3, pp. 608–619, 2004.

[73] G. Zonios, J. Bykowski, and N. Kollias, "Skin melanin, hemoglobin, and light scattering properties can be quantitatively assessed in vivo using diffuse reflectance spectroscopy," *Journal of Investigative Dermatology*, vol. 117, no. 6, pp. 1452–1457, 2001.

[74] H. Zhai, K. Wilhelm, and H. Maibach, *Dermatotoxicology*. Informa Healthcare, 2007.

[75] C. Lovell, K. Smolenski, V. Duance, N. Light, S. Young, and M. Dyson, "Type i and iii collagen content and fibre distribution in normal human skin during ageing," *British Journal of Dermatology*, 1987.

[76] R. Kelly, R. Pearse, R. Bull, J. Leveque, J. de Rigal, and P. Mortimer, "The effects of aging on the cutaneous microvasculature," *Journal of the American Academy of Dermatology*, vol. 33, no. 5, pp. 749–756, 1995.

[77] T. Dwyer, G. Prota, L. Blizzard, R. Ashbolt, and M. Vincensi, "Melanin density and melanin type predict melanocytic naevi in 19-20 year olds of northern european ancestry," *Melanoma Research*, 2000.

[78] I. Nishidate, N. Tanaka, T. Kawase, T. Maeda, T. Yuasa, Y. Aizu, T. Yuasa, and

BIBLIOGRAPHY

K. Niizeki, "Noninvasive imaging of human skin hemodynamics using a digital red-green-blue camera," *Journal of biomedical optics*, vol. 16, no. 8, pp. 086 012–086 012, 2011.

Vita

Saurabh Vyas received B.S. degrees in Biomedical Engineering and Electrical Engineering from the Johns Hopkins University in May 2012, and enrolled in the Biomedical Engineering M.S.E. program at the Johns Hopkins University in August 2012. Since July 2012, he has also been a Research Engineer at the Johns Hopkins University, Applied Physics Laboratory. His research focuses on computer vision and machine learning, in particular their application to biomedical modeling, and medical image analysis. His work has been published in several proceedings, including *IEEE*, *SPIE*, and *MICCAI*, as well as articles in *Journal of Biomedical Optics*, and *Ultrasound in Medicine & Biology*.

Starting in September 2014, Saurabh will begin working towards his Ph.D. in Bioengineering at Stanford University, focusing on computational neuroscience and systems neural engineering.