Multispectral image analysis in laparoscopy – A machine learning approach to live perfusion monitoring

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

von der KIT-Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Sebastian Josef Wirkert

aus Freyung

Tag der mündlichen Prüfung:	18.01.2018
Erster Gutachter:	Prof. Dr.-Ing. Rüdiger Dillmann
Zweiter Gutachter:	Prof. Dr.-Ing. Lena Maier-Hein

ACKNOWLEDGEMENTS

THESIS - about four years of my time bent by the people around me. First and foremost I thank Prof. Dr. Lena Maier-Hein, you gave me the best guidance imaginable throughout these years. One of the many things I learned from you is that smart people don't know everything - but are not afraid to ask, work hard and learn. Many thanks to Prof. Dr. Rüdiger Dillmann for waking my interest in machine learning as an undergraduate and supervising my thesis at the KIT. This work wouldn't have been possible without the generous funding provided by the ERC starting grant COMBIOSCOPY, a sincere thanks to the European Union.

A lone computer scientist on a interdisciplinary topic would have been a lost cause. I'd like to thank the London connection: Dan Elson, Neil Clancy and Dan Stoyanov for kickstarting me with physics knowledge. A big help was the "inside" knowledge from camera developers, specifically Jerome Baron, Julien Pichette, Murali Jayapala and Andy Lambrechts from imec and Bartlomiej Grychtol and Nikolaos Deliolanis from Fraunhofer PAMB. A thanks to Matthias Lambertz from Wolf who provided us with much needed hardware right from the start and to Storz for generously providing us with the tower. A medical topic of course needs medical doctors. Special thanks go to Hannes Kenngott, Peter Sauer, Edgar dos Santos, Benjamin Mayer, Peter Full, Dogu Teber and Tobias Simpfendörfer. What you guys do - save lives at daytime and science at night - you are real-life superheros!

Here's to my colleagues. First to Anant Vemuri, with whom I spent countless hours discussing, programming and experimenting. To Justin Iszatt, my first Bachelor student, without your work everything afterwards would not have been possible. To the students I supervised, Yan Zhang, Adyasha Maharana, Sara Moccia, Silvia Seidlitz and Leonardo Ayala: I am proud of your work and hope I did you justice. To the Biophotonics crew, mainly Esther Wild, Thomas Kirchner and Janek Gröhl for clearing things up when I got stuck. To Papa Navi aka Alexander Seitel, Keno März and Alfred Franz, you were like a swiss army knife. To Sebastian Brechtel and Sean Scherer, thanks for making my thesis more readable. To the grand ol' daddy of MBI, Prof. Dr. Hans-Peter "Pitt" Meinzer, you are a character and you know it. To the "Mädels" Jean, Steffi and Bea for always getting our back on the administrative stuff. To my room, which was a cool place to work not only because of the AC, but also due to Fabian α Isensee, Paul F. Jäger, Simon Kohl, Tobias Roß and Jens Petersen. I hope your portfolios will rise as will your drones.

Last but not least a thanks to family, Edda, Sepp, Leni, Noah and my girlfriend Nasrin. Your support from beginning to the amazing finishing party was just that. Amazing.

Abstract

Modern visceral surgery is often performed through small incisions. Compared to open surgery, these minimally invasive interventions result in smaller scars, fewer complications and a quicker recovery. While to the patients benefit, it has the drawback of limiting the physician's perception largely to that of visual feedback through a camera mounted on a rod lens: the laparoscope. Conventional laparoscopes are limited by "imitating" the human eye. Multispectral cameras remove this arbitrary restriction of recording only red, green and blue colors. Instead, they capture many specific bands of light. Although these could help characterize important indications such as ischemia and early stage adenoma, the lack of powerful digital image processing prevents realizing the technique's full potential.

The primary objective of this thesis was to pioneer fluent functional multispectral imaging (MSI) in laparoscopy. The main technical obstacles were: (1) The lack of image analysis concepts that provide both high accuracy and speed. (2) Multispectral image recording is slow, typically ranging from seconds to minutes. (3) Obtaining a quantitative ground truth for the measurements is hard or even impossible.

To overcome these hurdles and enable functional laparoscopy, for the first time in this field physical models are combined with powerful machine learning techniques. The physical model is employed to create highly accurate simulations, which in turn teach the algorithm to rapidly relate multispectral pixels to underlying functional changes. To reduce the domain shift introduced by learning from simulations, a novel transfer learning approach automatically adapts generic simulations to match almost arbitrary recordings of visceral tissue. In combination with the only available video-rate capable multispectral sensor, the method pioneers fluent perfusion monitoring with MSI. This system was carefully tested in a multistage process, involving *in silico* quantitative evaluations, tissue phantoms and a porcine study. Clinical applicability was ensured through in-patient recordings in the context of partial nephrectomy; in these, the novel system characterized ischemia live during the intervention. Verified against a fluorescence reference, the results indicate that fluent, non-invasive ischemia detection and monitoring is now possible.

In conclusion, this thesis presents the first multispectral laparoscope capable of video-rate functional analysis. The system was successfully evaluated in in-patient trials, and future work should be directed towards evaluation of the system in a larger study. Due to the broad applicability and the large potential clinical benefit of the presented functional estimation approach, I am confident the descendants of this system are an integral part of the next generation OR.

KURZFASSUNG

MOTIVATION

Die Überwachung funktionaler Parameter wie Gewebeoxygenierung und Blutvolumenanteil ist wichtig für viele minimal-invasive chirurgische Eingriffe. Vor Nierenteilresektionen sollte beispielsweise sichergestellt werden, dass der zu resezierende Teil nicht mehr durchblutet ist. Die dadurch bedingten Veränderungen der optischen Absorption und damit auch im Spektrum des reflektierten Lichtes können mittels das Auge "imitierender", konventioneller Kameras kaum wahrgenommen werden. Multispektrale Kameras erfassen subtilere Unterschiede im Spektrum und erlauben somit potentiell genauere Schätzung funktionaler Veränderungen.

Bislang scheiterte der Einsatz multispektraler Bildgebung in der (minimal-invasiven) Chirurgie an langen Bildaufnahmezeiten sowie der verzögerten oder nur ungenauen Schätzung der Oxygenierung.

Die grundlegende Zielsetzung dieser Arbeit war diese Limitationen im Rahmen der multispektralen Laparoskopie zu überwinden. Dadurch sollte erstmalig das flüssige Schätzen funktionaler Parameter für das komplette Blickfeld realisiert werden. Insbesondere wurden folgende Hypothesen untersucht:

- Wenige, ausgewählte Bänder erfassen die relevante Information.
- Maschinelle Lernverfahren im Zusammenspiel mit hochgenauen physikalischen Modellen ermöglichen schnelle und genaue Schätzung der funktionalen Parameter.

Eine zusätzliche Randbedingung war dabei, dass es keine andere praktikable Möglichkeit zur Bestimmung der Oxygenierung in Gewebe gibt. Daher bildeten realistische, simulierte Daten sowie unüberwachte Lernverfahren eine essentielle Grundlage zur Entwicklung eines geeigneten Schätzers.

BESCHLEUNIGUNG DER BILDAUFNAHMEZEITEN

Die Aufzeichnung multispektraler Bilder dauert gewöhnlich mehrere Sekunden. Bei der sequentiellen Erfassung ist die Aufnahmezeit abhängig von der Anzahl der aufgenommenen spektralen Bänder. Durch intelligente Auswahl der informativsten Bänder konnte diese Anzahl von 20 auf 8 gesenkt. Der dadurch ermöglichte Technologiewechsel reduzierte die Bildaufnahmezeit von 10,5 s auf 0,4 s.

Der entwickelte Algorithmus baut auf Arbeiten in der Fernerkundung auf und wählt das Subset von Bändern mit der maximalen differentiellen Entropie. Diese ist unter An-

nahme einer multivariaten Normalverteilung der Band-Bildintensitäten direkt proportional zur Determinante der Kovarianzmatrix des Subsets. In Experimenten mit Gewebephantomen und einer fünf Schweine umfassenden Studie konnte gezeigt werden, dass der vorgeschlagene Ansatz deutlich mehr Information erhält als Vergleichsverfahren.

Des Weiteren wurde erstmalig eine Kamera in der Laparoskopie eingesetzt, die multispektrale Bilder mit über 25 Hertz aufnehmen kann. Die Analyse dieser Bilder wurde maßgeblich durch den entwickelten Schätzalgorithmus ermöglicht.

Schätzung funktionaler Parameter durch Lernen der Monte Carlo Inversion

Als Ausganspunkt für die Schätzung dienen physikalische Modelle, welche die Interaktion von Licht mit Gewebe beschreiben. Bis jetzt waren diese entweder zu vereinfachend (Lambert-Beersches Gesetz) oder zu langsam (Monte Carlo Simulationen), um in der Praxis robuste Ergebnisse zu liefern.

Der entwickelte Ansatz kombiniert erstmalig hochgenaue physikalische Modelle mit maschinellen Lernverfahren für die multispektrale Laparoskopie. Durch das physikalische Modell erstellte Simulationen wurden dabei mittels genauer Kenntnisse der optischen Eigenschaften des Laparoskops und der Kamera in den Bandreflektanzraum transformiert. Der Domänenunterschied von Simulationen und realer Messung wurde explizit adressiert: Eine große Anzahl generischer Simulationen wurde mittels realer Messungen automatisch auf die gewünschte klinische Applikation angepasst. Das dabei zum Einsatz kommende Transferlernverfahren vermeidet die durch solche Ansätze oft verursachte hohe Varianz. Schließlich lernt der Algorithmus anhand der adaptierten Beispiele, die den Bildern zugrundeliegenden funktionalen Eigenschaften zu bestimmen. Dies ermöglicht dem Chirurgen die direkte Darstellung operationsrelevanter Parameter wie Gewebeoxygenierung und Blutvolumenanteil für das gesamte Blickfeld.

Die Evaluation des Algorithmus wurde in einem mehrstufigen Prozess mittels *in-silico* Simulationen, Gewebephantomen und *in-vivo* Aufnahmen von Schweinen durchgeführt. Abschließend wurde das System im Patienten evaluiert und mit Fluoreszenbildgebung als Referenz verglichen. Das entwickelte System ermöglichte dabei erstmalig die flüssige Überwachung der Durchblutung während einer Nierenteilresektion.

Schlusserfolg

In dieser Arbeit wurde das erste multispektrale System entwickelt, das die instantane Darstellung funktionaler Gewebeeigenschaften ermöglicht. Die funktionale Gewebeanalyse ist relevant für viele Eingriffe und wurde während einer Nierenteilresektion erfolgreich am Patienten evaluiert. Zukünftige Arbeiten sollten die Methode in einer größeren Patientenstudie weiter prüfen. Wegen der breiten Einsatzmöglichkeiten und dem potentiell großen Nutzen funktionaler Bildgebung bin ich zuversichtlich, dass die multispektrale Bildgebung Teil des OPs der Zukunft wird.

CONTENTS

Acknowledgements iii

1 Introduction 1

2 Foundation 9
2.1 Biophysical foundations . . . 10
2.1.1 Absorption in biological tissue . . . 11
2.1.2 Scattering in biological tissue . . . 14
2.1.3 Further reading . . . 15
2.2 Laparoscopic surgery . . . 15

3 State of the Art 17
3.1 Applications of multispectral imaging in endoscopy . . . 18
3.1.1 Cancer detection . . . 18
3.1.2 Perfusion monitoring . . . 19
3.1.3 Conclusion . . . 20
3.2 Recording of multispectral images in endoscopy . . . 21
3.2.1 Recording techniques . . . 21
3.2.2 Multispectral band selection . . . 24
3.2.3 Conclusion . . . 27
3.3 Analysis of multispectral images . . . 28
3.3.1 Calibration . . . 28
3.3.2 Functional parameter estimation . . . 29
3.3.3 Evaluating functional approaches . . . 32
3.3.4 Other techniques . . . 34
3.3.5 Conclusion . . . 35
3.4 The bigger picture . . . 35
3.4.1 Blood gas analysis . . . 35
3.4.2 Pulse oximetry . . . 36
3.4.3 Diffuse reflectance spectroscopy . . . 36
3.4.4 Fluorescence . . . 36
3.4.5 Other techniques . . . 37
3.4.6 Conclusion . . . 38
3.5 Conclusion . . . 39

4 Machine learning based Monte Carlo inversion 41
4.1 The forward model . . . 43
4.1.1 Layered tissue model . . . 43
4.1.2 Calculation of spectral reflectance . . . 44
4.1.3 Adaptation to imaging system . . . 45

4.2 Machine learning based inversion . . . 46
4.3 Experiments and results . . . 47
4.3.1 General setting . . . 48
4.3.2 Spectral reflectance . . . 49
4.3.3 Adaptation to imaging system . . . 51
4.3.4 Inverse model . . . 55
4.4 Discussion . . . 60
4.5 Conclusion . . . 63

5 Domain adapted parameter estimation 65
5.1 Methods . . . 66
5.1.1 Data set generation using a generic tissue model . . . 67
5.1.2 Domain Adaptation . . . 67
5.2 Experiments and results . . . 69
5.2.1 Experimental setup . . . 69
5.2.2 Validity of tissue model . . . 69
5.2.3 Performance of domain adaptation . . . 71
5.3 Discussion . . . 74
5.4 Conclusion . . . 75

6 Band selection 77
6.1 Methods . . . 78
6.1.1 Multispectral Imaging System . . . 78
6.1.2 Formal endoscopic image acquisition model . . . 79
6.1.3 Endoscopic Sheffield Index . . . 80
6.2 Experiments and results . . . 81
6.2.1 Phantom experiments . . . 81
6.2.2 *In vivo* experiments . . . 82
6.3 Discussion . . . 83
6.4 Conclusion . . . 86

7 First fluent in-patient analysis and recording 87
7.1 Video-rate laparoscope: image recording . . . 89
7.2 Video-rate laparoscope: image analysis . . . 90
7.3 Video-rate laparoscope: application . . . 94
7.4 Experiments and results . . . 94
7.4.1 General setting . . . 95
7.4.2 *In patient* application . . . 96
7.4.3 Evaluation of analysis method . . . 102
7.4.4 Evaluation of image recording . . . 107
7.5 Discussion . . . 109
7.6 Conclusion . . . 114

8 Concluding Thoughts 117

A Imec noise, linearity and imaging system transformation 121
A.1 Sensor noise and linearity . . . 121
A.2 Imaging system . . . 121

Bibliography **125**

List of Publications **143**

Attachments **144**

Special thanks to the Delft University of Technology for kindly providing their LaTeX–Template and allowing for slight modifications concerning colors and layout as used within this document.

1

INTRODUCTION

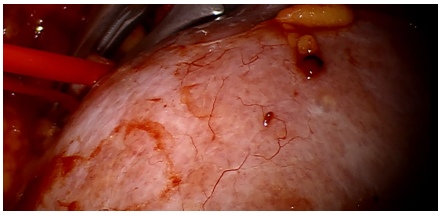
What is essential is invisible to the eye.

The fox from Antoine de Saint Exupéry's "The Little Prince"

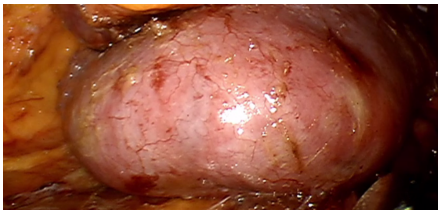
1

MOTIVATION

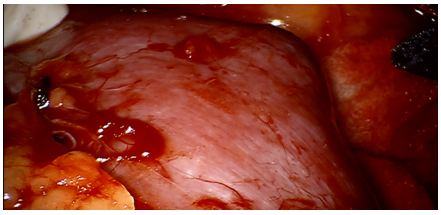
IMAGINE you're a surgeon. If you looked at the laparoscopic images in Fig. 1.1, could you make out which of the organs is sufficiently perfused with blood?



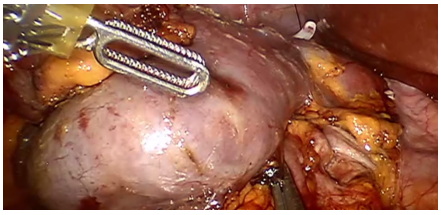
(a) Kidney perfused?



(b) Kidney perfused?



(c) Kidney perfused?



(d) Kidney perfused?

Figure 1.1: Could you tell through which organ blood is flowing? Resolution in footnote.

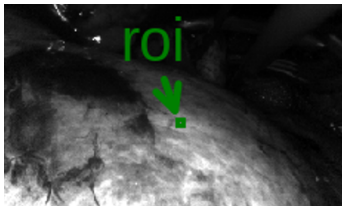
Even for expert physicians this is extremely challenging - a major problem, since detecting ischemia (lack of perfusion) often is a game of life and death: after anastomosis, the (re)connection of two tubular structures, leakage in the blood stream can be extremely critical. In a recent study, 14% of patients showed anastomotic leakage, but even experienced surgeons could only predict this with 44% sensitivity [106]. Partial nephrectomies are an example of inducing ischemia on purpose: to prevent bleeding, the blood flow is usually stopped before the kidney tumor is surgically removed. Selective clamping avoids harmful global ischemia by stopping the blood supply only to the tumorous area [67]. While this approach could substantially change future partial nephrectomies, an imaging technique that is suited to characterization of ischemia has yet to be found [10]. Colonoscopies are another example that illustrates the limitations of current imaging techniques. As a standard procedure recommended in most countries for everyone over the age of 50, gastroenterologists screen for possibly cancerous polyps. Although the widespread implementation of such screening programs has reduced the mortality considerably [9], many lesions are still overlooked by the physicians. In the case of flat adenoma, up to 42% are missed [90]. These flat adenoma might be identifiable by subtle changes in perfusion due to angiogenesis in early stages or hypoxia in later stages.

A large body of research aims at developing technical innovations to support the physician in surgical perception. Many traditional approaches focus on displaying preoper-

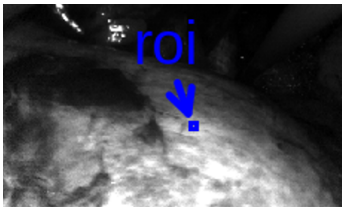
Upper left is not perfused (ischemic). Upper right is perfused, as is lower left. The kidney on the lower right is partly perfused: roughly below the lower instrument the kidney is ischemic, above it is not.

atively acquired images during the intervention. These “offline” methods cannot react to the dynamical state of organs or pathologies, leaving the aforementioned changes in perfusion undetected. This “blind spot” has spurred the novel branch of live functional imaging. Its goal is the detection and monitoring of changes in metabolism and blood flow that are mostly hidden to the human eye. One of the most successful examples for functional imaging is indocyanine green (ICG) fluorescence, which can reveal information about tissue perfusion after injection of a near-infrared (NIR) fluorescing dye. A laser excites the fluorophore, which binds to blood plasma and bile. This then emits a slightly shifted NIR signal that allows visualizing perfusion up to a depth of about 1cm. However, fluorescence is not quantitative and complicates the intervention by requiring injection of the dye, which leads to follow up questions such as: how much dye - and where is it to be injected?

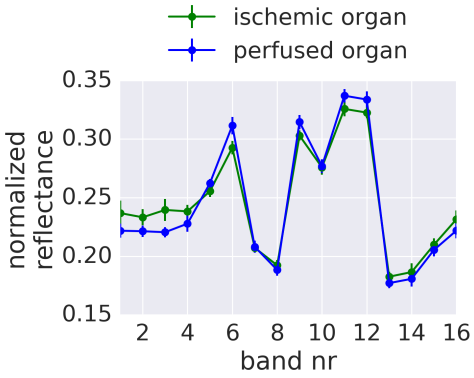
In this thesis I aim to pave the way for a new success story, a technique that is less invasive and more straightforward to use: **multispectral imaging** (MSI). MSI aims at detecting subtle changes in amount and oxygenation of blood, both of which influence the wavelength dependent absorption of light. Multispectral cameras record these changes by ceasing to “imitate” the eye, which has only a limited capability to detect the resulting color changes. Instead, multispectral imagers parse the electromagnetic spectrum at many, typically narrow, bands of light, often selected to suit a given task. Each pixel in a multispectral image thus encodes information about the functional state of a tissue, exemplified in Fig. 1.2. However, the raw measurements are of no use to the surgeon. To provide a real advantage, the underlying functional causes have to be revealed.



(a) ischemic organ band 1



(b) perfused organ band 1



(c) ROI comparison

Figure 1.2: Comparison of the spectra of perfused and ischemic tissue. (a) one band of an ischemic kidney recording and a 3x3 region of interest (ROI). (b) the same kidney, but this time perfused. (c) compares the normalized reflectances in these ROIs (mean with standard deviation as errorbars). The task of functional imaging is deducing the molecular changes leading to these differences.

1

OBJECTIVES

The primary objective of this thesis was to **pioneer fluent (video-rate) functional imaging by means of multispectral imaging for laparoscopy**; imagine a laparoscope, which will, at the press of a button, display information about perfusion. This technique should be quantitative, allow continuous monitoring and easy workflow integration.

TECHNICAL CHALLENGES

To enable fluent functional imaging, three major challenges had to be addressed:

C1 The lack of image analysis concepts that provide both high accuracy and speed. Image analysis involves inverting the multispectral camera measurements to determine the molecular tissue composition (see Figure 1.3). This inversion is complex and cannot be done analytically without restrictive assumptions. Further, the inversion cannot be learned from tissue measurements, as the necessary ground truth is often not available (see also challenge 2).

C2 A quantitative ground truth for the recordings is hard or even impossible to obtain. During an intervention, one has to resort to point measurements for reference. However, functional information for the whole field of view is not detectable with other methods.

C3 Multispectral image recording is often slow, typically ranging from seconds to minutes. This is caused by the need to record many narrow spectral bands. Depending on recording technique, this inevitably leads to either longer imaging times or images with a lower spatial resolution.

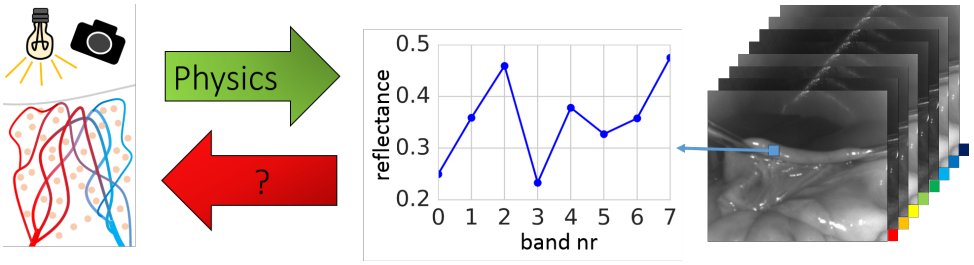


Figure 1.3: For known tissue compositions and light/camera arrangements (a), well worked out physics can be used to calculate tissue reflectance (b), that is the fraction of reflected light to incident light, which would be measured for a pixel in a multispectral image (c). However, solving the inverse problem, from the measured reflectance to the underlying tissue properties such as oxygenation or perfusion, is not obvious.

CONTRIBUTIONS

The central guiding hypotheses investigated in this thesis were:

H1 Exact simulations working in conjunction with powerful machine learning techniques can enable accurate, high-speed functional imaging.

H2 The relevant information can be captured by a small, selected subset of bands to enable faster image recording.

The investigation of these hypotheses spawned the following technical (T) and medical (M) contributions:

T1: Machine learning-based spectral decoding [195]. For the first time exact physical models are combined with powerful machine learning techniques to enable functional laparoscopy. The physical model creates highly accurate simulations of colonic tissue, which are transformed to arbitrary laparoscope and multispectral camera settings through a sophisticated model of the optical system. The resulting reflectance simulations teach the algorithm to rapidly relate multispectral pixel measurements to functional changes. Because learning from such simulations has to be performed with great care, a multistage evaluation process was carried out in a V-model like manner. Within this framework, a range of *in silico* evaluations, tissue mimicking phantoms, lab measurements and data from six pigs ensured that the taken assumptions are valid.

T2: Domain adaptation for MSI [196]. This contribution extends T1 to be applicable not only to colonic tissue, but almost arbitrary tissues encountered in visceral surgery (see Figure 1.4). It relies on a highly generic tissue model that aims to capture a large range of optical parameters that can be observed *in vivo*. Adaptation of the model to a specific clinical application based on unlabelled *in vivo* data is achieved by incorporating a new concept of domain adaptation that explicitly addresses the high variance often introduced by such methods. According to comprehensive *in silico* and *in vivo* experiments this approach enables estimation of functional properties for various tissue types without the need for incorporating specific prior knowledge on optical properties.

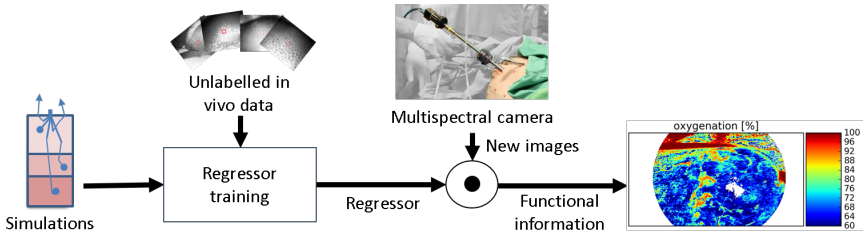


Figure 1.4: Technical contributions 1+2: Approach to functional estimation. Highly accurate simulations generate tissue, reflectance pairs. These pairs are fed to a machine learning regressor, which learns the inverse function so that it can derive functional information about the tissue from the measured reflectances. Additionally real recordings can be used to remove bias introduced by learning from simulations.

T3: Information-theoretic band selection [194]. Faster recording of multispectral images is enabled by capturing only a small, but discriminative subset of bands (see Figure 1.5). The developed band selection method contrasts with competing approaches by neither relying on labelled training data nor on a sophisticated light-tissue interaction model. Instead, it extends a concept from the related field of remote sensing to work in the endoscopic environment. The resulting algorithm selects bands which maximize the retained differential entropy in a completely unsupervised manner. A phantom and *in vivo* porcine study, involving data from five surgeries, support hypothesis H2.

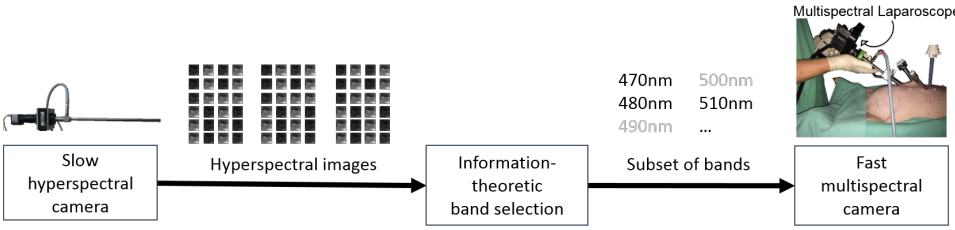


Figure 1.5: Technical contribution 3: The band selection method from chapter 6 can aid in designing new cameras by selecting the most useful bands. For this purpose, some hyperspectral images are taken. The band selection algorithm proposed in this thesis can be employed to choose the best multispectral bands for future recording.

T4: Video-rate, functional imaging capable laparoscope. The final technical contribution is the first video-rate multispectral laparoscope for functional analysis. While existing systems are bulky and/or slow, this system retains the look and feel of conventional laparoscopes (see Figure 1.6). This important step towards clinical use was realized by first-time application of a novel multispectral camera in laparoscopic environments and further advancement of the powerful analysis methods. Clinical applicability was ensured through in-patient recordings, during which the system characterized ischemia live during the intervention. Verified with a fluorescence baseline, the results indicate that fluent, non-invasive ischemia detection and monitoring is now possible.

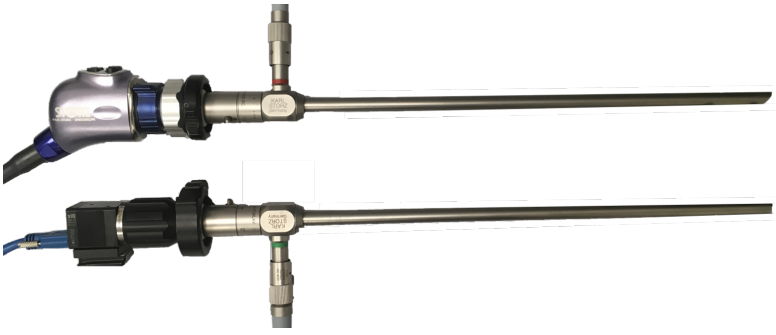
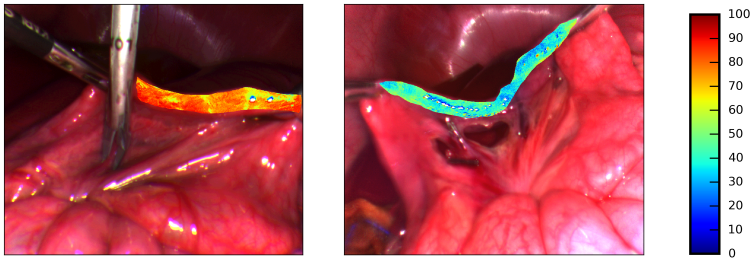


Figure 1.6: Technical contribution 4: A standard RGB laparoscope (top) compared to the multispectral laparoscope developed in this thesis (bottom).

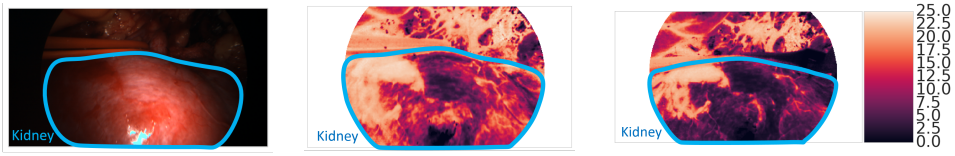
M1: Oxygenation estimation for tissue and cancer characterization. Pig experiments were conducted in collaboration with the Department for Visceral Surgery of the Heidelberg University Hospital. The developed systems were able to detect changes in oxygenation of several organs, including kidney, liver and bowel (see Figure 1.7). This result could have important implications, e.g. for predicting organ function after anastomosis, transplantations and cancer resections. Furthermore, oxygenation estimation might be valuable during cancer screening, as hypoxia is an important tumor hallmark.



(a) Bowel before clipping (b) Bowel after clipping

Figure 1.7: Medical contribution 1: oxygenation monitoring, here in a porcine small bowel. (a) RGB and oxygenation overlay on the well oxygenated small bowel segment, while a clip was applied. (b) Hypoxia after clipping the blood supply to this small bowel segment. The colors code oxygenation in percent.

M2: Ischemia monitoring in partial nephrectomy. The system developed in T4 was evaluated in collaboration with the Department for Urology of the Heidelberg University Hospital. The in-patient results suggest that ischemia induced before partial nephrectomies can be detected by a sharp drop in the estimated amount of blood in tissue (see Figure 1.8). Reperfusion lead to a restoration of the pre-ischmeic values within two seconds. This result could lead to a paradigm shift towards selective clamping in partial nephrectomy, because it allows previously not possible, highly spatially resolved, immediate and repeated monitoring of kidney perfusion.



(a) RGB before ischemia (b) Functional before ischemia (c) Functional during ischemia (30s)

Figure 1.8: Medical contribution 2: ischemia monitoring in a partial nephrectomy patient. (a) Perfused kidney and (b) the corresponding functional estimate. (c) Functional estimate 30s after ischemia was induced by clamping. The colors code the blood volume fraction in tissue [%].

1

OUTLINE

The next chapter establishes the foundations necessary to put this thesis in context: the biophysical properties of light-tissue interaction and a short summary of laparoscopic surgery. After establishing this basis, Chapter 3 thoroughly reviews the state of the art of MSI with a focus on applications, recording and analysis. Chapters 4 to 7 are the core methodological chapters, table 1.1 summarizes the relation of these chapters to the challenges, hypotheses and contributions. Each chapter presents the method, experiments and results in combination with a chapter discussion and conclusion. The thesis concludes with the author's reflections on the presented work and on the road ahead in Chapter 8.

Table 1.1: Overview on methodological chapters.

	Challenges	Hypothesis	Contribution
Chapter 4	C1, C2	H1	T1, M1
Chapter 5	C2	H1	T2
Chapter 6	C2, C3	H2	T3
Chapter 7	C1, C3	H1	T4, M1, M2

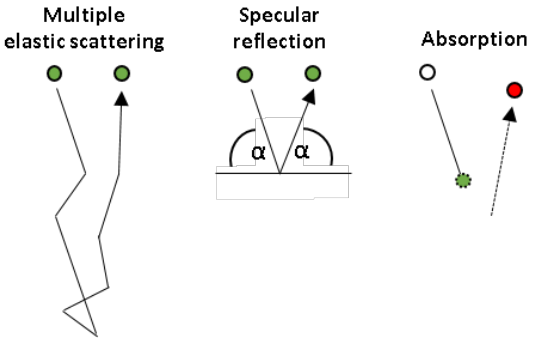
2

FOUNDATION

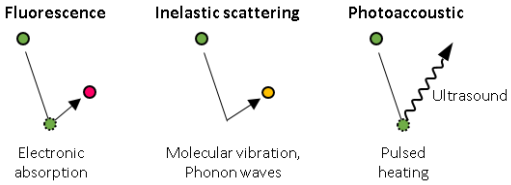
MULTIPECTRAL imaging is a highly multidisciplinary topic, at the intersection of Biology, Physics, Medicine and Computer Science. Since the thesis was written from a Computer Scientist's perspective, this chapter introduces the background knowledge from the other fields that is necessary to put the work in context. First, biophysical foundations are established in Section 2.1, then minimally invasive surgery is introduced in Section 2.2.

2.1. BIOPHYSICAL FOUNDATIONS

On each pixel of the multispectral image, a molecular signature of the tissue is recorded. To understand how this signal is formed, first a firm grasp on light and tissue interaction needs to be established. Biophotonics is the field which studies these interactions, some of which are summarized in Figure 2.1. This chapter first focuses on the optical phenomena necessary to understand multispectral imaging (MSI): absorption (Section 2.1.1) and scattering (Section 2.1.2).



(a) Interactions relevant for multispectral imaging.



(b) Interactions relevant for other Biophotonics techniques.

Figure 2.1: Examples for light-tissue interactions. Photons are denoted by circles, with colors representing the wavelength. Absorption events are indicated by dotted lines. (a) Multispectral cameras measure light reappearing at the surface due to multiple scattering events. Specular reflections at the tissue surface are generally a source of nuisance, because they do not encode any tissue specific information but just "mirror" the light-source. Because light at different wavelengths is absorbed at different magnitudes, white light, which consists of every color of light, will reappear as colored light. (b) The shift in color caused by fluorescence and inelastic scattering can give clues on present fluorophores and the chemical tissue composition. Thermoelastic expansion caused by the heating with pulsed laser light leads to ultrasound waves, which are measured in photoacoustic imaging. Figure in parts reproduced from [203].

For MSI, the dominant interactions are **absorption** and elastic **scattering** (or just "scattering" for convenience). Figure 2.2 gives a schematic overview on how these physical properties influence the signal. To briefly summarize MSI: absorption and scattering change dependent on the molecular composition of tissue. Since both quantities not only vary spatially but also with wavelength, measurements at multiple wavelengths gives a system of equations which can be solved with respect to the molecular composition.

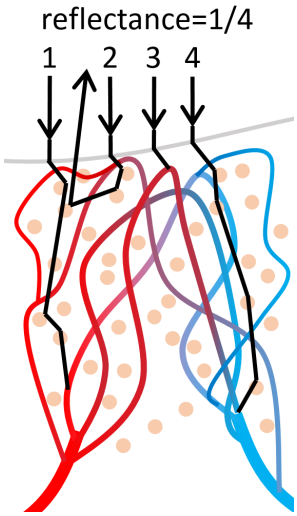


Figure 2.2: As the exemplary four photons enter the tissue they get scattered and absorbed. Some will reappear at the surface after multiple scattering events. The ratio between diffusely backscattered photons and total number of photons is called diffuse reflectance. This ratio is dependent on absorption and scattering, but also on the wavelength of light. Absorption mainly takes place at the oxygenated (red) and deoxygenated (blue) blood vessels. Scattering happens at small particles like cells, sub-cellular organelles and collagen fibers, visualized as yellow circles.

2.1.1. ABSORPTION IN BIOLOGICAL TISSUE

Hemoglobin, the molecule that transports oxygen in blood, often is the only notable optical absorber in internal organs [30, 191]. Hemoglobin is responsible for the red color of blood and internal organs and is present in two main forms: oxygenated and deoxygenated [191]. Deoxygenated hemoglobin binds with oxygen when passing through the lungs. In organs haemoglobin travels through arteries to smaller arterioles and finally the capillaries, thin vessels through which the oxygen is released to the surrounding cells due to the difference in oxygen pressure (and the drop in pH which is a consequence of higher concentration of CO_2 [113]). From the capillaries hemoglobin travels to venules and veins, through the heart and is resaturated with oxygen in the lungs [163]. Figure 2.3 gives an example for capillaries, arterioles and venules in human mucosal tissue.

Optical absorption in biological tissue is also due to several other chromophores, like melanin, bilirubin, β-carotine, fat and water [98]. In this thesis I mainly investigate the

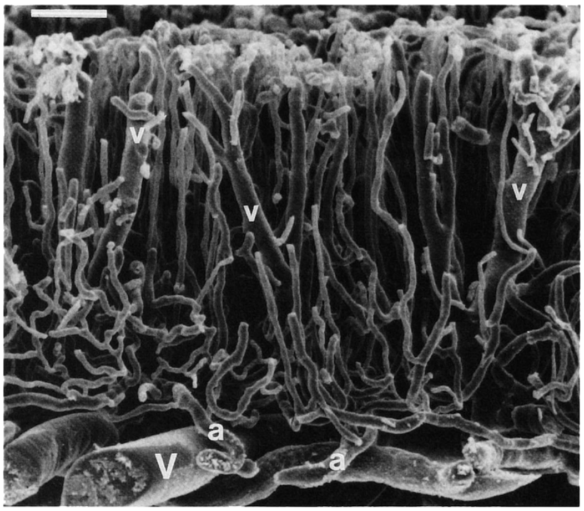


Figure 2.3: Microvessels in a humon colon crossection (bar=100μm). Blood flows through the arterioles (a) which divide into capillaries. From the capillaries blood drains into venules (v) and finally into veins (V). Reprinted with permission from [168].

visible wavelength range from roughly 450nm to 700nm. In these ranges water and fat absorption play minor roles. Note however that water, which is transparent in our everyday world, is opaque in the infrared. Melanin is mainly present in skin tissue and the eyes, and doesn't play a major role in the visceral applications of interest in this thesis. Bilirubin and β-carotine absorption is usually insignificant compared to hemoglobin. Note, however, that the validity of this statements depend on a case to case basis. A counterexample is the gallbladder, which has a distinct color due to the bile containing bilirubin.

Absorption is usually quantified by the **absorption coefficient** $\mu_a(\lambda)$ [cm^{-1}], defined as the probability of photon absorption in a medium per unit infinitesimal path length. Its reciprocal gives the mean free path length before an absorption event. Blood contains a mixture of both oxygenated hemoglobin HbO_2 and deoxygenated hemoglobin Hb. This relationship is called oxygenation and can be defined as [191]:

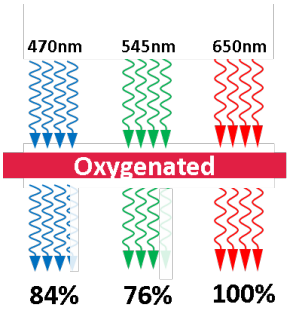
$$s = \frac{C_{\mathrm{HbO_2}}}{C_{\mathrm{HbT}}} \tag{2.1}$$

where C_{HbT} refers to total hemoglobin concentration $C_{\mathrm{HbT}} = C_{\mathrm{HbO_2}} + C_{\mathrm{Hb}}$ and $C_{\mathrm{HbO_2}}$, C_{Hb} are molecular concentrations [mol L^{-1}] of oxygenated and deoxygenated hemoglobin[1]. The absorption coefficient μ_a in tissue can be calculated as [91]:

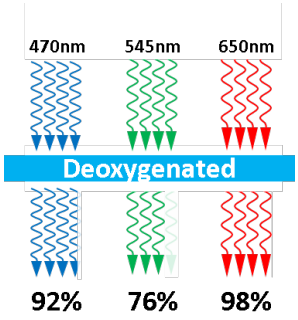
$$\mu_a(\lambda) = \log(10) \cdot \nu_{\mathrm{Hb}} \cdot C_{\mathrm{HbT}} \cdot (\epsilon_{\mathrm{HbO_2}}(\lambda) \cdot s + \epsilon_{\mathrm{Hb}}(\lambda) \cdot (1-s)) \tag{2.2}$$

[1] Hemoglobin concentration can alternatively be given in [gL^{-1}]. In this case, it has to be divided by the gram molecular weight of hemoglobin, which is 64,500g mol^{-1} to convert to molar concentration. A typical value for hemoglobin concentration is 150g L^{-1} [152].

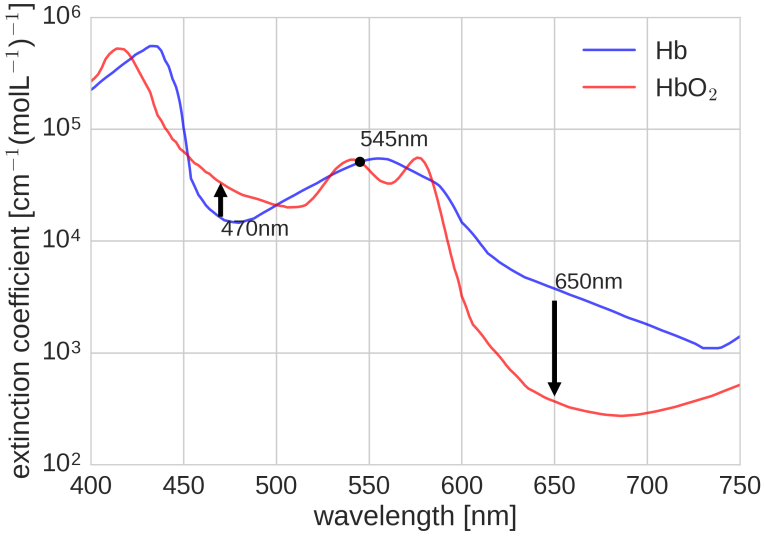
where υ_{Hb} indicates the volume fraction in tissue occupied by blood and $\epsilon(\lambda)$ the wavelength dependent extinction coefficient [$\mathrm{cm}^{-1}\frac{1}{\mathrm{molL}^{-1}}$]. This wavelength dependence of the extinction coefficients is the key to functional imaging and shown in Figure 2.4. It makes measuring s and the product $\upsilon_{\mathrm{Hb}}C_{\mathrm{HbT}}$ possible by determining the absorption coefficient at a minimum of two wavelengths [191].



(a) Transmission oxygenated capillary



(b) Transmission deoxygenated capillary

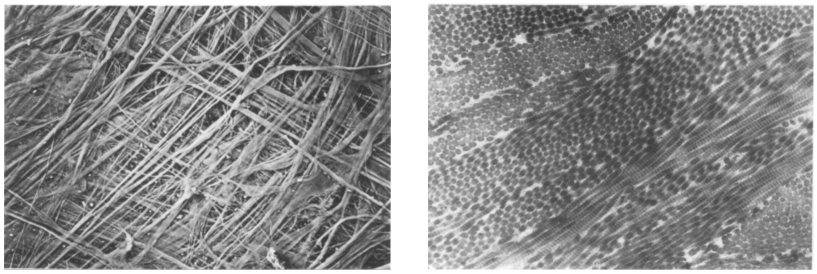


(c) Hemoglobin extinction coefficient

Figure 2.4: Oxygenated (HbO_2) and deoxygenated (Hb) blood absorb light differently depending on wavelength. This is shown for capillaries of 9μm diameter in (a) and (b). This wavelength dependent absorption is the key to functional imaging. It can further be seen why blood appears red: most light is absorbed in the blue and green wavelength range. The absorption of hemoglobin is determined by the extinction coefficients [152] shown in (c). Light at 545nm is absorbed the same for both variants, a so called isosbestic point. Although the extinction coefficient technically includes both scattering and absorption, absorption dominates scattering in hemoglobin [191].

2.1.2. SCATTERING IN BIOLOGICAL TISSUE

If only absorption were present, one would not be able to measure a signal when both the detector (camera/eyes) and the light source are on the tissue surface - as it is the case in surgery. Due to (possibly multiple) scattering events photons reappear on the surface, and can thus be measured. These measurements encode information about hemoglobin concentration and oxygenation because of the probabilistic absorption on the path of the photon. Light which is directly reflected at the surface (specular reflection) does not encode function information but the spectrum of the light-source and thus is not of special interest [62].



(a) Human colon collagen fibre network x74

(b) Human colon collagen fibre network x10,600

Figure 2.5: Collagen is the main structural protein in the body. (a) shows the layer-wise organization in the colon. (b) shows a higher zoom to the fibers, which cause light to scatter. Disorganization of the collagen fiber network can be associated with cancer [91]. Reprinted with permission from [180].

Scattering occurs when a photon is absorbed and re-emitted in a possibly different direction [143]. It is most probable if the scattering structure matches the optical wavelength and has a refractive index different to the surrounding medium [191]. In tissue this may be caused by collagen fibrils (Fig. 2.5), cells or subcellular organelles such as lysosomes, mitochondira and Golgi apparati [64, 98, 191]. It can be described by two factors: (1) the frequency of interaction and (2) a factor describing the directionality of scattering [191]. The frequency of scattering is usually quantified by the **scattering coefficient** μ_s [cm^{-1}], defined as the probability of photon scattering in a medium per unit infinitesimal path length. As in the case of absorption coefficient, the reciprocal gives the mean free path length before a scattering event occurs [191].

Assuming scattering is symmetric relative to the direction of the incoming photon, the directional component can be described by a probability distribution of the scattering angle Θ. The expectation of the cosine of Θ is referred to as **anisotropy**, quantified by the anisotropy factor $g = E(\cos(\Theta))$ [64]. Thus, an anisotropy of 1 describes virtually no scattering, 0 isotropic scattering and -1 completely backward scattering. When inserted in the approximate Henyey Greenstein [89] phase function

$$p(\Theta) = \frac{1}{4\pi} \frac{1-g^2}{(1+g^2-2g\cos(\Theta))^{\frac{3}{2}}} \tag{2.3}$$

the anisotropy allows sampling from the probability density function $p(\Theta)$ [64].

Both μ_s and g are wavelength dependent and can be calculated exactly using Mie's scattering theory [132] or fitted to measurements with heuristic, exponential curves [98]. Mie theory is a direct consequence of the Maxwell equations, applicable in the case of homogeneously distributed spherical particles, when some properties like the particles radii and refractive index are known [91]. Although particles in tissue are not spherical, some studies suggest that modelling tissue as homogeneous spheres is a plausible approximation [29]. Experimental determination of both μ_s and g is challenging [98]. In practice, one often measures the reduced scattering coefficient $\mu_s' = \mu_s(1-g)$ [cm^{-1}] and one of g or μ_s. Experimental measurements of g suggest tissue scatters highly in the forward direction [98] and anisotropy increases with wavelength. This is surprising, as Mie theory would suggest that scattering is more isotropic when wavelengths get larger than particle size. One of several possible explanations following [98] is that the scatterers in tissue are not homogeneous and of the same size, but there are mesoscopic ($\gg 10\mu$m) structures which cause this effect.

2.1.3. FURTHER READING

"Biomedical Optics - Principles and Imaging" [191] from Lihong V. Wang and Hsin-I Wu is a book tailored specifically to biomedical optics. A comprehensive overview on tissue optics from a modelling and experimental point of view can be found in "Optical properties of biological tissues: a review" from Steven L. Jacques [98]. Džena Hidowić-Rowe and Ela Claridge show how a detailed model of colonic tissue can be derived from experimentally measured literature values in "Modelling and validation of spectral reflectance for the colon" [91]. If one wants to expand upon the physical foundation of absorption and scattering, "Absorption and scattering of light by small particles" by Bohren et al. [20] starts from the Maxwell equations. This book is not focused on biomedical optics, but might be recommendable for people with a strong physics background and interest.

2.2. LAPAROSCOPIC SURGERY

Laparoscopic (also called minimally invasive) surgery (MIS), aims to minimize damage inflicted to patients during visceral interventions. Metallic or plastic tubes of usually 5-10mm diameter called trocars are placed into small incisions made in the abdomen. Insufflating the abdomen with CO_2 through the first trocar creates the pneumoperitoneum, which ensures enough working space following procedure. The first trocar is used to host the laparoscope (Figure 2.6). The lens system projects the image from the abdomen onto a camera attached at the end. Additional trocars (usually 1-4) are placed with the help of the laparoscope to host instruments during the intervention.

Although modern MIS techniques can be traced back over a hundred years [121], rapid, widespread acceptance only took place in the late 80s and early 90s [108, 187]. Today, MIS is commonplace in procedures on inner organs [203], with for example 96% of cholecystectomies in the United States of America being performed laparoscopically

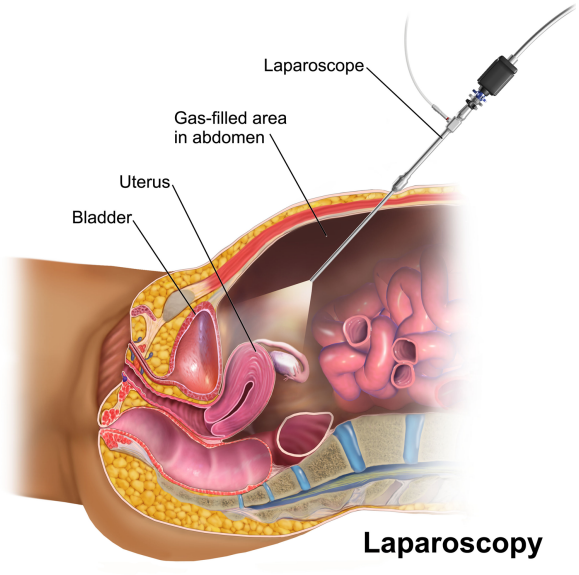


Figure 2.6: Schematic depiction of a laparoscope inserted in the pneumoperitneum. Figure reprinted with permission from [172].

[183]. MIS have been found to result in smaller scars, fewer complications and quicker recovery compared to open surgery [203]. Together with the related flexible endoscopy, the field is growing rapidly, with the market projected to reach 38 billion US$ by 2018 [203].

The lack of haptic feedback in minimally invasive interventions require the physician to rely on visual cues alone. This makes tasks like tissue discrimination during cancer resections more difficult [203]. The same challenge arises in flexible endoscopy, where a flexible tubular instrument is used e.g. to screen for cancer in the gastrointestinal tract. In this thesis, I refer to endoscopy for both laparoscopy and flexible endoscopy. The following state-of-the art chapter will explore possible uses for MSI in endoscopy, followed by an in-depth analysis of multispectral recording and analysis.

3

State of the Art

MULTISPECTRAL imaging (MSI) was first proposed for use in remote sensing in the mid 80s [71]. Since then, the technique has been adopted in many fields, including pedestrian detection [96], art conservation [55, 119], food quality control [52] and detecting details in crime scenes that would otherwise be invisible [114]. In the biomedical field, MSI techniques have been used literally from head [49, 56, 92] to toe [202] in the human body.

However, in minimal-invasive surgery MSI is still in a research stage with no commercial laparoscopic devices available. To enable clinical use of MSI, the state of the art should be advanced in three aspects:

1. Multispectral imaging must provide clear clinical benefits over existing techniques - either through better patient outcomes or simpler workflow integration.
2. Multispectral image recording should be at a speed of 25Hz or faster and at HD resolution to provide immediate and exact feedback.
3. Analysis methods have to be in real-time and at video rates, while not compromising on performance.

This chapter will analyze advances in these topics in the above order.

3.1. APPLICATIONS OF MULTISPECTRAL IMAGING IN ENDOSCOPY

Multispectral imaging devices record the spectral signatures of tissue for the whole field of view of the camera. This high resolution information characterizes the tissue surface, as the reappearing light does not penetrate tissue deeper than about 1 to 3mm [17]. Because MSI needs accurate knowledge about the spectral distribution of the light, the inherently controlled lighting conditions of endoscopic interventions make these the special interest and focus of this review and thesis. Possible interventional uses include organ classification [3, 133, 204] and identification of important structures such as blood vessels, nerves or the ureter [4, 58, 141, 142, 158]. In my opinion, most promising for endoscopic applications are early cancer detection for screening purposes and perfusion monitoring. Both will be analyzed in depth in the following.

3.1.1. CANCER DETECTION

The hope that multispectral imaging could aid in early cancer detection is mainly attributed to two effects: angiogenesis and hypoxia. Angiogenesis, the formation of new blood vessels, is an important cancer hallmark . This neovascularisation is induced by the tumor due to its need for nutrients such as oxygen [83]. Angiogenesis happens already in premalignant stages, and therefore could serve as an early indicator during cancer screenings. Hypoxia is defined as low oxygen tension and can be observed in tumor areas with malfunctioning vasculature, which are often significant in size [45]. As multispectral imaging can estimate total hemoglobin (linked to angiogenesis) and oxygenation (linked to hypoxia), there is reasonable hope it can help in cancer detection, e.g. to increase adenoma detection rates during screenings.

Kaneko et al. [105] reasoned that real-time oxygenation estimation in flexible endoscopy can help in cancer screening. Though not truly multispectral (only three, carefully selected bands), this work is interesting because of its innovative approach and convincing study design. Oxygenation estimation was done in real-time (not further specified) and evaluated on 40 patients. Significant differences were detected between neoplastic and non-neoplastic areas in esophagus and colorectum. Similarly, Han et al. [82] aimed for early detection of colorectal cancer with a true multispectral system. They directly classified colorectal cancer, meaning classification on the camera measurements without estimating functional parameters as oxygenation. Images were acquired from 12 patients and it was not mentioned how the tumor regions were identified besides being manually marked. Martinez-Herrera et al. developed a light modulating, multispectral gastroscope [129]. They collected data from 17 patients and direct classification was performed, comparing various feature sets and classifiers. Hohmann et al.[93] evaluated a multispectral system on 14 gastric cancer patients. The tumor margins were determined by medical experts. It was noted that this ground truth might not be completely accurate, as medical experts can not determine these with certainty. Several classifiers were tried to directly classify cancerous tissue. The results showed quite low sensitivity and specificity. Fazwy et al. characterized another endoscopic system [51] for measuring blood volume fraction and oxygenation in the lung with possible application in lung cancer detection. They performed measurements on 10 healthy patients, so no functional analysis of malignancies was done. Gu et al. [77] presented preliminary results on 12 gastroscopy patients, showing that multispectral images provide higher contrast for cancerous tissue than RGB imaging.

While endoscopic patient studies such as the above are rare, multispectral cancer detection in *ex vivo* [5, 36, 57, 74, 117, 201], animal models [2, 124, 125, 170] and easily accessible cancers like head and neck [14, 81, 122, 173] and skin [42, 54, 84] are widespread. Comprehensive overviews on the role of multispectral imaging in cancer detection can be found in [123] and [126].

3.1.2. PERFUSION MONITORING

Perfusion can be defined as the volume of blood that flows through the capillaries of a certain tissue or organ in an interval of time [193]. Ischemia is the stopping or reduction in perfusion and can lead to irreversible damage. Depending on the organ, this damage will occur within minutes to hours. The damage is caused by the imbalance between supply and demand, since cells still consume oxygen and other nutrients, but no new oxygenated blood is arriving [47]. Because of reduced oxygenation and the lower amount of blood in ischemia, multispectral imaging has potential to characterize perfusion.

Ischemia can be prompted by several reasons. It can for example be caused by a physician voluntarily. An example for this is partial nephrectomy, where a kidney tumor is surgically removed. To reduce bleeding during the resection, ischemia can be induced by clamping the supplying hilar vessel. Studies suggest that (warm) ischemia can be sustained about 20 minutes before damages occur [179]. Olweny et al [145] used a multispectral system to monitor kidney oxygenation during robot-assisted partial nephrec-

tomy in 18 patients, aiming to predict postoperative kidney function in the future. Several other similar studies on the kidney in open settings were conducted by this group [94, 182, 209].

Ischemia can also be a unwanted side effect after surgery. Two prominent examples are organ transplants and anasotomsis, the reconnection of two tubular sections. Clancy et al. [32, 33] performed uterine transplants in rabbit and sheep models with a multispectral laparoscope. They found that post transplantation oxygenation has a lower baseline levels compared to pre-transplantation readings. These measurements were in line with data from pulse oximetry acquired in parallel. In future studies they aim to investigate if the measured oxygenation values can serve as a predictor for long time survivability. In another variant of the system [30] they investigated measurement of bowel oxygen saturation. The main focus of this work was the comparison of oxygenation determined by multispectral imaging and BGA during clamping of bowel segments in a porcine model. The determined pearson correlation coefficient between BGA and multispectral imaging was 0.76. Clancy argued comparing oxygenation of both systems needs some consideration: Blood for the BGA has to be drawn from a larger vessel, while multispectral imaging will measure an average of venous and arterial blood in the tissue imaged at a specific pixel. Furthermore, BGA readings are highly temperature dependent, a 2° Celsius difference in temperature can account for up to 9% difference in oxygenation measurement. They argued that systems like these could help predict success of bowel anastomosis, e.g. after cancer resection.

3.1.3. CONCLUSION

To date, no one breakthrough application for endoscopic MSI has been identified. Two of the most promising areas for minimally invasive multispectral imaging are cancer detection and perfusion monitoring. In colonic cancer screenings, the goal can be to increase adenoma detection rate compared to white-light colonoscopy. Perfusion monitoring could help in assuring that the correct vessel has been clamped before partial nephrectomy. Currently this can be tested with ICG fluorescence, which disturbs the surgical workflow and can potentially be replaced by completely passive MSI. Other applications like anastomosis and organ transplant success verification are relevant applications as well. However, here the benefit of MSI will be harder to test, as for example anastomotic leakage has indications besides postoperative ischemia.

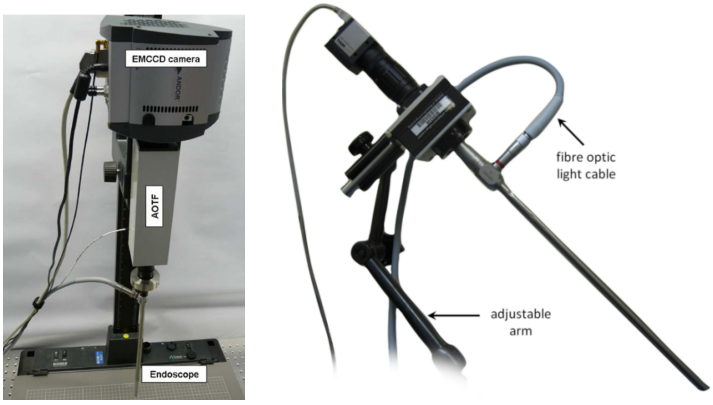
While there are strong indications that MSI could help in the aforementioned application, reliable proof of clinical benefit is hard to establish and currently missing. Therefore, an application with a clear reference method should be identified. The application proposed in this thesis is perfusion monitoring in partial nephrectomy. The system was compared to the more invasive ICG fluorescence baseline in Chapter 7, indicating that MSI can detect ischemia by monitoring blood volume fraction. Two other main challenges for MSI are not-standardized recording and analysis, which are discussed in the following.

3.2. RECORDING OF MULTISPECTRAL IMAGES IN ENDOSCOPY

Use in laparoscopy places high requirement on multispectral systems. The ideal systems should

1. be lightweight and compact, so it can be attached to a laparoscope and held during the intervention
2. record images at video rates (25Hz and above), to give immediate feedback and avoid motion artifacts or blurring
3. provide high resolution images (ideally HD) to capture intricate details
4. be able to display conventional RGB images

Current multispectral endoscopes are quite far from this ideal piece of hardware. Two techniques are commonly employed for multispectral image recording in endoscopy: staring systems and snapshot systems. Figure 3.1 shows typical systems as found in literature. The next section reviews these techniques and their endoscopic implementations. Section 3.2.2 discusses band selection techniques, which can benefit both staring and snapshot systems.



(a) Leitner et al. [117] (b) Clancy et al. [30]

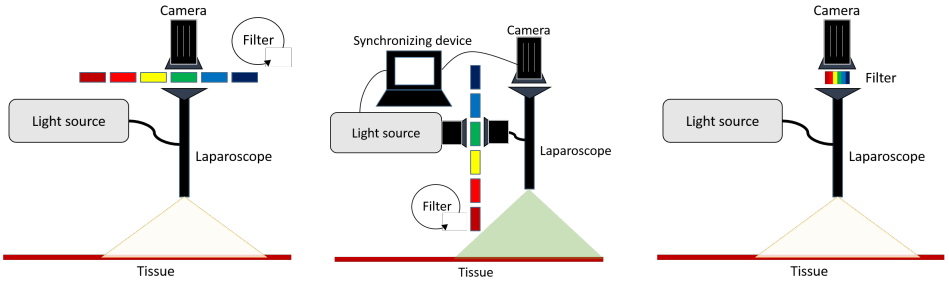
Figure 3.1: Two staring systems that filter light at the camera level. (a) The highly sensitive, but bulky EMCCD camera and the fast switching AOTF allow relatively fast recording of eight bands in 5Hz. (b) The slow switching LCTF and standard monochrome camera record an image stack in about 3s, but feature a smaller form factor. Pictures adapted with permission from [117] and [30].

3.2.1. RECORDING TECHNIQUES

Light back-scattered to the surface encodes information about the present molecules, varying at wavelengths. Since measuring at one exact wavelength would not yield enough photons, multispectral recording techniques integrate them from within a band of light.

Recording several bands of light leads to a 3D image cube with two spatial and one spectral dimension. The broadness of the bands is often characterized by the full width at half maximum (FWHM) [118], the distance in nm of the two wavelengths where transmission is at 50% of the maximum.

Multispectral images can be recorded using point-scanning (whiskbroom), line-scanning (pushbroom), staring or snapshot approaches [62, 118, 123]. Whiskbroom and pushbroom techniques perform spatial scanning and are not suitable for interventional MSI [123]. This leaves the spectral scanning staring systems and snapshot systems which record both spatial and spectral information at once, summarized in Fig. 3.2.



(a) Staring filtering at camera (b) Staring filtering at light source (c) Snapshot filtering at pixel level

Figure 3.2: Three ways to record multispectral images. Staring system perform time sequential wavelength scanning either (a) in front of the camera or (b) after the light source. This filtering can be done by LCTFs, AOTFs or fast filter wheels. (c) Snapshot systems acquire the full spectral-spatial data-cube at once, e.g. by filtering directly on pixel level.

STARING SYSTEMS

Staring systems acquire images with full spatial information at the cost of temporal resolution by sequentially switching the recorded band. This switching takes place at light source level or before the detector. Most commonly, filter wheels [31, 51, 77, 82, 195, 196], liquid crystal tunable filters (LCTF) [30, 32–34, 210] and acousto-optical tunable filters (AOTF) [117] are used for switching in endoscopic snapshot systems. LCTF and AOTF systems are flexible in wavelength selection, an often important requirement for research purposes. AOTFs enable fast switching of wavelengths in about 0.1ms, LCFTs switch considerably slower, taking around 150ms [61], albeit offering better blocking efficiency and acceptance angles than AOTFs [117]. Both AOTF and LCTF suffer from low throughput of about 30% [61]. The transmission efficiency of competing fast filter wheel systems is near 100% [31]. However, these systems suffer from bulky setups and cannot switch wavelengths without manipulating the device. Other disadvantages of fast filter wheels mentioned in literature are slow wavelength switching, misregistration due to filter movement, mechanical vibration and narrow spectral range [118]. However, besides the slow wavelength switching, which is faster than conventional LCTF but slower than for AOTFs, these disadvantages could not be observed in this thesis.

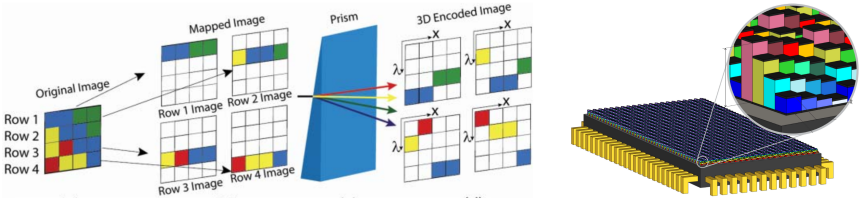
A simple ways to realize such a multispectral staring system is to place an AOTF, LCTF or fast filter wheel between endoscope and a monochromatic camera [30–34, 117, 210] (Fig. 3.2a). The LCTF-based systems acquire image stacks in 3 to 41 seconds in the reviewed literature. The differences are caused mainly by number of recorded bands, light source and camera quantum efficiency. The AOTF technique has been explored in an ex vivo setting by Leitner et al. [117]. They were able to record eight band multispectral images at 5Hz or 51 band images at 1.25s. The system is quite bulky, using a highly sensitive EM-CCD camera as detector (see Fig. 3.1a). Due to the better light throughput and relatively fast switching, Clancy et al. [34] were able to record eight-band multispectral images at 4Hz with a fast filter wheel.

A somewhat more involved, second option is switching bands at the light source as in [51, 94, 105, 145, 182, 209] (Fig. 3.2b). This requires a control unit to synchronize light source and camera [51]. Fawzy et al. [51] presented a system which filters light by means of a fast filter wheel. Each of the slots was occupied by one of six customized triple-band filters. These triple-band filters had three narrow peaks in the red, green and blue. A conventional RGB camera was used to collect the reflected light. The system was able to record 3x6=18 band multispectral images at 15Hz. Similar systems, but with classical one band filters, were presented in [82] and [77]. Hohmann et al. used a modified light source to illuminate tissue in 6 different bands in gastroscopy. Their system acquired multispectral images in 2Hz. Kaneko et al. [105] recorded images using a conventional RGB probe. They illuminated tissue sequentially with broad band light and laser light which is sensitive to oxygenation. In [209] Zuzak et al. presented the digital light processing DLP® hyperspectral system. In DLP® light is first dispersed by a grating. The spectrally unfolded light hits a micro-mirror array with 768 by 1024 fast switching mirrors. Each of the mirrors either directs light to a heat sink or propagates it forward towards the target tissue. The system can therefore produce almost arbitrary shapes of light. Recording with 3Hz is possible in three-shot mode, where three selected bands of light are used for illumination. The DLP® was patented by Zuzak et al. [208] and used in various kidney related studies [94, 145, 182].

SNAPSHOT SYSTEMS

Snapshot imagers acquire the complete spectral and spatial information in a single image, at the cost of resolution [118]. Various techniques have been proposed [80], but only few have found application in biomedical imaging [62]. Even less have found endoscopic usage, likely due to device compactness and, in the case of tomographic and Fourier based solutions, computational complexity [62]. One of few published endoscopic systems is the image mapping spectroscopy (IMS) device presented in [109]. The IMS concept is layed out in Fig. 3.3a. It is based on a custom designed image mapper, which contains micro mirrors to map spatially close zones to isolated zones on a CCD camera. A prism is used to spectrally resolve the pixels in these isolated zones. The mapping back to a 3D multispectral datacube is then a simple reshuffling operation. IMS systems need a large-format camera, which normally have low frame rates [62] (e.g., 5.2 frames per seconds in [109]). Spectrally resolving detector arrays (SRDA) are another technique for snapshot multispectral imaging. The imec (Leuven, Belgium) SRDA sen-

sor uses Fabry-Pérot filters at pixel-level, arranged in a mosaic pattern (see Fig. 3.3b). It can record images with up to 170Hz and records 16 bands in the visible or 25 bands in the near-infrared. This and the Pixeltech (Largo, FL, USA) PixelCam™ which only records four relatively broad bands, are to my knowledge the only commercially available snapshot sensors.



(a) IMS snapshot concept

(b) imec snapshot concept

Figure 3.3: Two principles for snapshot recording. The IMS concept (3.3a) maps spatially close regions to separate zones on the CCD using micro mirrors. Before the light reaches the CCD it is spectrally distributed by a prism. The imec concept (3.3b) performs filtering using Fabry-Pérot elements on top of each pixel. The pattern of these filters is repeated in a mosaic pattern. Pictures adapted with permission from [109] and [63].

Currently, staring systems outnumber snapshot devices. No clear standard technology has been established for snapshot imaging, although where are some competing technologies, like computed tomographic imaging spectrometer (CTIS), IMS and SRDA. The first snapshot multispectral endoscope was presented by Kester et al. [109], who invented the IMS concept. The presented IMS endoscope operates in visible wavelength ranges and is able to record 200x200x48 images with 5Hz. Another recently proposed snapshot system featured a flexible 10x10 fiber optic [120]. The distal end of this fiber optic was rearranged to a 100x1 line. This line is resolved spectrally by a spectrograph and imaged with a CCD, similar to pushbroom multispectral imagers. Rearranging the image pixels restores the spatial information. With this technology, Lim et al. manage to record images with 10x10x756 resolution in 6Hz. The imaged wavelengths are in the range of 400-1000nm.

3.2.2. MULTISPECTRAL BAND SELECTION

Band selection can help mitigate the aforementioned problems of slow image recording and low resolution: sequentially recording staring techniques suffer from long image recording times, often in the range of seconds [16]. Aside from the surgeon not patient enough to wait for seconds for one image, tissue and instrument movement misalign the recorded stack, thus rendering the measured spectra invalid. Some sequential techniques, such as the fast filter wheel, have only limited slots available, so reduction to this number of slots is a must. With proper band selection imaging time usually reduces linearly with the number of bands for staring techniques. The one-shot recording snapshot techniques have differing issues. While they can record an image stack all at once, the image resolution decreases with the increasing number of bands. Modern surgery however is performed at full HD resolutions (1920 x 1080 pixel) and some vendors already

Table 3.1: Overview on relevant band selection methods.

Group	Modality	label-free	Method	Data	Application
Wood et al. [197]	Fluorescence	no	wrapper	ex vivo	cancer detection
Han et al. [82]	MSI	no	filter	in vivo	colorectal cancer
Lu et al. [125]	MSI	no	filter	in vivo	cancer detection
Nouri et al. [141, 142]	MSI	yes	filter	in vivo	surgical visualization
Gu et al. [77]	MSI	yes	filter	in vivo	gastric abnormalities
Preece et al. [153]	MSI	yes	wrapper	in silico	physiological parameters
Wirkert et al. [194]	MSI	yes	filter	in vivo	physiological parameters

are offering 4K (4096 x 2160 pixel) solutions. This dwarfs the spatial resolution of current multispectral snapshot sensor such as the Ximea (Muenster, Germany) MQ022HG-IM-SM4X4-VIS (512 × 272 pixel), with drawbacks that subtle textural changes in tissue might remain hidden. Band selection can mitigate this drastic difference, as a smaller number of bands recorded leads to higher resolutions.

Note that band selection is different from dimensionality reduction methods as principal component analysis (PCA), which reduce the dimensionality by finding a subset composed of combinations of bands/features. This does not reduce imaging time, because usually all bands have to be recorded to compute the axes found by PCA. Band selection is however closely linked to variable/feature selection, in which algorithms can be grouped roughly in filter, wrapper and embedded methods [79]. Filter methods determine the best features as an independent preprocessing step. Wrapper methods use the performance metric for a given task (e.g. accuracy of tumor detection) to rank feature sets. Finally, embedded methods are directly part of the training process of a machine learning algorithm trying to find a solution to a specific task. While filter methods can rely on unsupervised correlation analysis, wrapper and embedded methods usually need labelled training data.

Due to its early adoption in the field, a large body of work in band selection is available in the field of remote sensing (see e.g. [13, 27, 78, 161, 167]). Moreover, band selection is employed in fields like food safety [43] and histopathology [148]. This state of the art review will be restricted to algorithms with relations to interventional imaging. See table 3.1 for an overview of this approaches, described briefly in the following.

Wood et al. [197] explored band selection in context of distinguishing fluorophores. To this end they selected the bands which maximize the Area Under the Curve (AUC) of a Naïve Bayesian classifier's Receiver Operating Charactersitic (ROC). More specifically

they used a greedy algorithm to remove bands which contribute the least to the AUC while having a low AUC if taken on their own. The algorithm was evaluated on phantoms and lung cancer biopsies, which were stained with an exogenous fluorophore. They concluded that three wavelengths provide essentially as much information as using all sixteen wavelengths.

Han et al. [82] selected the most discriminative bands for detecting gastric tumors. These were selected by maximizing the symmetric class-conditional Kullback-Leibler (KL) divergence, a sub-optimal, greedy algorithm [43] search algorithm, chosen because the number of subsets grows exponentially with the number of bands. This algorithm iteratively adds bands to the selection set based on the largest incremental increase on the symmetric KL measure. After selecting the bands in this manner, a SVM was trained to detect cancerous tissue on a pixel level. The algorithm was evaluated on 12 patients and a total of 21 colorectal tumors and bands recorded in the visible wavelength range. They concluded that reasonable tumor detection performance can be achieved with 5 out of 28 bands.

Lu et al. [125] performed band- and feature selection for hyperspectral images recorded in a mouse head and neck tumor model. They employed the maximal relevance and minimal redundancy (mRMR) framework [150]. Maximal relevance is characterized by high mutual information between the bands and the classes. Minimal redundancy is characterized by minimal mutual information between bands. As in [82], the band selection results were generated by iterative adding of the band which maximizes a combination of these two criteria. The authors recorded hyperspectral data with a total of 226 bands in the visible and near infrared wavelength. Tumor v non-tumor reference was established using *in vivo* GFP fluorescence. The authors resorted to pseudo GFP ground truth due to difficulties matching the images to real, but deformation-prone histopathological ground truth. Due to the difficulty to establish reliable reference data, I think band selection methods should ideally not require labelled in vivo data (the label-free column in table 3.1). The following approaches aim to find the best bands without the need for ground truth.

Nouri et al. [141, 142] inspected a number of unsupervised, thus label-free, band selection algorithms. The algorithms, originating from the remote sensing community, were evaluated within the context of hyperspectral ureter surgery. Compared to other state-of-the art works, the band selection was performed in a large spectral range (400-1700nm), enabled by using a system consisting of two cameras, one for the VIS-NIR and one for the NIR-SWIR spectral range. Instead of finding bands which separate a certain class, the authors aimed to find the best three bands to visualize instead of RGB to the surgeon. They evaluated the differing methods by several contrast and entropy measures and measured how the bands can differentiate structures as the ureter and fat. Two findings here are of special interest in our context: a variant of the Sheffield index, which is also employed in this thesis (Chapter 6.1.3) shows best results for ureter discrimination

and the fact that inclusion of middle infrared information provides substantially better discrimination results. Unfortunately incorporating middle infrared bands is not possible for silicon based standard cameras, as the material is not sensitive to light in this range. The work of Nouri et al. [141] is similar to our work [194] and was published basically at the same time. Our work further differs by the endoscopic setting and the oxygenation estimation as target criterion.

Gu et al. [77] selected three bands to discriminate gastric abnormalities from benign tissue. As [141, 142] they aimed to replace the RGB image with the selected bands. Their algorithm first selected the band with the highest variance from a set of 27 bands in the visible. Subsequent bands were added iteratively with the criterion of minimizing mutual information compared to the previous selection. The bands were determined using 29 images from 12 patients with gastric abnormalities. The authors claim that the selected bands increase contrast, however this statements seems to lack quantifiable evidence.

Preece et al. [153] were selecting bands with a different idea. First Kubelka-Munk light transport theory based simulations were created for assessing pigmentation of human skin. After adapting these simulations to a set of virtual filters, a genetic algorithm was employed to find the best subset for recovering papillary dermis thickness and blood/melanin content. They ensured the bands can invert the parameters uniquely by differential-geometric reasoning. Incorporation of ground truth coming from simulations allowed circumventing problems related to real ground truth determination mentioned in [125].

3.2.3. CONCLUSION

No off the shelf recording solution is available for endoscopic multispectral imaging. The four requirements mentioned in the beginning of this section are met by no single system. I predict the successful system will only record a very limited, but highly expressive, number of bands, making band selection, as presented in Chapter 6, a prerequisite for future system design.

Devices for recording split into staring and snapshot systems, each with unique properties. While staring systems mainly suffer from slow image recording, snapshot devices acquire images at higher rates, at the cost of resolution. Staring systems which manipulate the light source have some advantages. They do not require bulky imagers, but can use the conventional camera, as e.g. done in [51, 105]. Moreover, they can modulate light intensity to the needed output for narrow or broad bands. This allows design of systems which mix broad RGB bands with selected, task specific narrow bands as in [105]. Since the light source is modified, these devices are harder to certify for in-patient use [93] and need synchronization of camera and light source. Acquiring a high number of spectral bands with systems like these has so far proven unsuccessful. Snapshot systems on the other hand do not suffer from image misalignment and leave the light source untouched.

The completely passive recording technique might make using these systems easier in clinical practice. On the downside, current snapshot devices lack high-resolution image acquisition.

Chapter 7 presents the to my knowledge first laparoscope which acquires multispectral images in 25Hz and above. This compact system is the first to fulfill both the first and second initially mentioned requirement, which was made possible by the first endoscopic use of the imec snapshot sensor. Using methods presented in Chapter 7.2, the system can further display RGB information. The research question of high resolution MSI is remaining. However, the system offers higher spatial resolution (512x272x16) than existing snapshot devices. I strongly believe that multispectral band selection as presented in Chapter 6 is one of the key element to removing this last obstacle.

3.3. ANALYSIS OF MULTISPECTRAL IMAGES

Multispectral images can be analyzed by unraveling the functional tissue properties (Section 3.3.2) or by directly trying to estimate the end target, such as the presence or absence of a disease (Section 3.3.4). A challenge when employing functional approaches is the evaluation, which is discussed in Section 3.3.3. Both approaches to analysis commonly employ the calibration steps described in the next section.

3.3.1. CALIBRATION

In biomedical applications, multispectral images are of interest because they can capture information about the molecular properties of the underlying tissue. However, raw multispectral image measurements are dependent on the light source, device optics and camera characteristics. Consider two images recorded of the same tissue but under different light sources. The intensity of the second light source being twice that of the first in the red wavelengths. This would cause the recordings in the second image to be brighter in the corresponding wavelengths. A calibration step is typically performed by acquiring a dark and a white image. The dark image, $D(x, y, \lambda)$, is recorded in complete darkness and the white image, $W(x, y, \lambda)$, is taken from a standard white target[1]. The multispectral image $I(x, y, \lambda)$ is transformed to band reflectance R by applying the transformation [123]

$$R = \frac{I - D}{W - D}, \tag{3.1}$$

where D accounts for the dark current of the camera and W for the lighting conditions. All mathematical operations are performed pointwise. The images are all of the same size $I, W, D, R \in R^{hxwx|\mathbf{b}|}$, with the first two dimension being the image height h and width w and the last one being the number of spectral bands $|\mathbf{b}|$.

For a static scene, with the camera and light source being at same pose for tissue and white target, the above transformation is sufficient. This can for example be the case

[1] Commonly recorded from a completely diffuse reflecting surface, using National Institute of Standards and Technology certified Spectralon targets, which diffusely reflects 99% of light in the visible [48, 123]

when haemodynamics are monitored on the open brain's surface or in microscopy. This cannot be guaranteed in the dynamic scenes encountered in multispectral endoscopy . Consider again two images of the same tissue, but the light source is being closer to the tissue when recording the second image. The measured image will be brighter in all bands, because of more incident light per surface area. If the surface is assumed to be Lambertian[2] , measurements change multiplicatively in spectral dimension, dependent on pose of light source and camera [36]. Dividing by the mean in spectral direction [196] removes the multiplicative factor. By applying this transformation information can be lost, since the molecular changes in tissue can also lead to multiplicative changes in spectral reflectance.

Depending on camera setup, additional calibrations can be necessary. The dark current can be time and temperature dependent and the camera response can be non-linear. Calibration methods in such situations have been discussed in [76, 86, 107, 127].

3.3.2. FUNCTIONAL PARAMETER ESTIMATION

One way to interpret the multispectral images is to estimate the underlying functional information of the tissue for each image pixel. After linking the multispectral image pixel measurements to causing functional effects, like blood volume fraction and oxygenation, so called parametric maps can be displayed to the surgeon, see Fig. 3.4 for an example. These functional images can be used to guide surgical decision. To be able to calculate the functional information, usually a physical model is needed to establish the function-measurement relationship. The most prominent models are based on the Beer-Lambert law or Monte Carlo reasoning, briefly explained in the following.

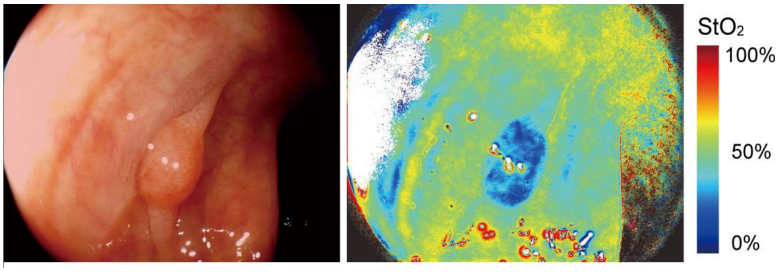


Figure 3.4: Oxygenation map (right), revealing a hypoxic adenoma in the center. Left image shows the same scene as RGB. Reprinted with permission from [105].

BEER-LAMBERT

Most systems for interventional *in vivo* multispectral imaging as [30, 138, 209] use linear estimation approaches based on the modified Beer-Lambert law [92, 162]. It states, that

[2] A surface which is diffusely reflecting isotropically in the spectral domain and the intensity follows the Lambert cosine law [115]

the following relationship holds for spectral reflectances $r(\lambda)$ and absorption $a(\lambda)$:

$$a(\lambda) = -\log(r(\lambda)) = \mu_a(\lambda)L + G \tag{3.2}$$

with L being the path length travelled by the photons and G is a geometry dependent term e.g. accounting for scattering losses. If hemoglobin is the only notable absorber, this leads to:

$$a(\lambda) = \epsilon_{\mathrm{HbO_2}}(\lambda)C_{\mathrm{HbO_2}}L + \epsilon_{\mathrm{Hb}}(\lambda)C_{\mathrm{Hb}}L + G. \tag{3.3}$$

3

If more than one measurements are available, indicated by vector notation $\mathbf{a}$, we write for convenience:

$$\mathbf{a} = -\log\mathbf{r} = H\mathbf{x}. \tag{3.4}$$

Here, H is the matrix composed of $(\boldsymbol{\epsilon}_{\mathrm{HbO_2}}, \boldsymbol{\epsilon}_{\mathrm{Hb}}, \mathbf{1})$ and $\boldsymbol{x} = (C_{\mathrm{HbO_2}}L, C_{\mathrm{Hb}}L, G)^T$. Using linear least squares regression, the optimal solution $\hat{\boldsymbol{x}}$ is determined by:

$$\hat{\boldsymbol{x}} = (H^T H)^{-1} H^T \mathbf{a}. \tag{3.5}$$

This enables calculation of oxygenation by $s = \frac{C_{\mathrm{HbO_2}}L}{C_{\mathrm{HbO_2}}L + C_{\mathrm{Hb}}L}$. Note, that total hemoglobin concentration can only be determined up to a constant factor L, because $LC_{\mathrm{HbT}} = L(C_{\mathrm{Hb}} + C_{\mathrm{HbO_2}})$.

While being fast, the method is based on the assumptions that light travels an equal path length L in tissue regardless of wavelength, that scattering is constant and that absorption is low compared to scattering [92]. These assumptions do not hold up in real tissue. Red light penetrates the tissue deeper than blue light, because blood absorbs photons less in the red wavelength range. The scattering is also wavelength dependent, e.g. dropping by about 20% from 500nm to 620nm within the bowel [30, 98]. Sometimes, the linear regression may calculate physically impossible negative values. In these cases one can resort to much slower non negative least squares regression.

From a theoretical standpoint Beer-Lambert's law is defined for spectral reflectances $r(\lambda)$, so reflectances at specific wavelengths, while the camera measures reflectance r_{camera} in a certain band of light $\mathbf{b}_i$ (after the transformations in the previous subsection). In many practical application, like [30, 33, 34, 209], the Beer-Lambert approach is nonetheless used for the band reflectance measurements. Regression is then based on:

$$\mathbf{a}_{\mathrm{camera}} = B\,H\mathbf{x}. \tag{3.6}$$

With B (for bands) being the linear transformation matrix to transform spectral reflectances to band reflectances. However, in reality the transformation happens on image acquisition level, $\mathbf{r}_{\mathrm{camera}} = B\,\mathbf{r}$. This means, that Equation 3.6 cannot be used for regression with $\mathbf{r}_{\mathrm{camera}}$, as

$$\mathbf{a}_{\mathrm{camera}} = BH\mathbf{x} = -B\log\mathbf{r} \neq -\log B\mathbf{r} = -\log\mathbf{r}_{\mathrm{camera}}. \tag{3.7}$$

When narrow bands are recorded, this might not play a big role, which is probably why most endoscopic approaches [30, 33, 34, 94, 145, 182, 209] used the Beer-Lambert based

approach and Equation 3.6 for oxygenation determination. In case of broadband cameras like RGB or the camera used in Chapter 7 this effect can however not be neglected anymore. One possible solution is to transform band reflectances to spectral reflectances as a preprocessing step. For this purpose, several approaches have been explored. Nishidate et al. [139, 140, 200] used Wiener estimation to restore spectral reflectance. They determined the Wiener matrix, a linear transformation of the band reflectances which minimizes squared reflectance reconstruction error [200]. This transformation of RGB data to multispectral data is highly underconstrained and can be made more stable by adding a regularization term [103]. Non-linear spectral reflectance reconstruction has also been done in the community [57, 100], using a neural network based approach [128]. An approach outside of medicine has recently been proposed at CVPR [198]. It estimates both illumination and spectral reflectance in a bi-linear system by leveraging linearity assumptions and knowledge about typical illuminations and reflectances.

MONTE CARLO

Alternative to the simple Beer-Lambert approach are more accurate, non-linear, forward models based on diffusion approximation [18, 50, 157], δ-P_1 [51, 164], Kubelka-Munk [35, 46, 188, 189], Monte Carlo methods [36] and others [57, 202]. Most of these come in closed form, so at least the forward model can be evaluated rapidly. This is not the case for the widely accepted gold standard for describing photon transport in biological tissue, Monte Carlo (MC) simulations [206].

The MC simulations can produce arbitrary exact results at the cost of run time by simulating photon transportation [206]. They estimate quantities which can be characterized by an expectation value of many independently propagated photons. Examples for these quantities are spectral reflectance and fluence [191]. MC simulations can be applied if the medium is loosely packed so that scattering events can be considered as independent [191]. MC programs offer varying degrees of freedom, allowing varying complexity in modelling e.g. geometry and physical phenomenon.

The multi-layered Monte Carlo (MCML) is probably the most commonly employed inhomogeneous Monte Carlo model [206]. In MCML [190] tissue is modeled as infinite, homogeneous slabs. Each of these slabs is characterized by its depth d, g, μ_a, μ_s and refractive index n. Photons are fired perpendicular in an infinitely narrow beam onto the tissue, in which they are probabilistically propagated and absorbed. E.g., the Henyey Greenstein phase function from Equation 2.3 is used to sample the scattering angle from g. The back-reflected photons (more correct: photon packets) are recorded and their ratio to totally fired photons gives spectral reflectance. Note that simulations are not defined by the wavelength of light in MCML. The wavelength is indirectly coded by setting the wavelength dependent parameters μ_a, μ_s and g. That means one run of the simulation framework will give spectral reflectance at only one exact wavelength. MCML is not capable of simulating complex geometry or effects as fluorescence. For this, more complex frameworks like Geant4 [8] need to be used.

The drawback of all MC methods is that they are time and resource intensive, since the propagation of many photons has to be simulated individually. To mitigate this, modern

approaches employ the graphics processing unit (GPU) to simulate photon migration in parallel [7]. Additional speed-up can be achieved by using non-optimal acceleration techniques [206]. Like Beer-Lambert, MC present a *forward model* for photon propagation. However, because the problem is non linear, inversion is not as easy as in the Beer-Lambert case, where simple least squares fitting is a reasonable choice. On the contrary, inverting the Monte Carlo forward model is an active area of research, summarized in the following.

Claridge et al. investigated Monte Carlo inversion [36] on excised colon cancer tissue. The forward simulations were created with a sophisticated layered model for colonic tissue from a previous work [91]. They varied three parameter related to blood content, mucosa thickness and scattering in five steps to create a total of $3^5 = 125$ simulated spectra. The other parameters were set to standard values. For inversion the spectrum closest to the measurements w.r.t. euclidean norm was determined. They found that inverting spectra like this is non-unique within the noise level. Consequently, subsequent spatial regularization was employed under the assumption that parameter values change smoothly.

This thesis aims at inversion using machine learning. In other literature, this was explored in context of skin [188, 189] and the ocular fundus [175]. The ocular fundus work [175] is to my knowledge the only work combining Monte Carlo simulation and machine learning. It used a neural network to derive parameters like macular pigment and retinal hemoglobins. Main drawback of their work was that the modelled spectra do not fit the real recordings. Before inversion, they therefore applied a heuristic "shift and scale" operation on the simulations. They noted this mismatch might be caused by insufficient variability in their model and that comprehensive in vivo measurements of absorption and scattering properties are not available. Machine learning was used for melanoma screening in [188, 189]. In both works Kubelka-Munk theory was used for forward modelling. The earlier work [189] regression was done with support vector machine and evaluated in silico. In the later work [188] nearest neighbor classification to find the nearest spectrum replaced support vector regression. Advantage of the nearest neighbor approach is, that the determined spectrum can be compared to the measured one as a sanity check. The method was used to determine physiological parameters for benign caucasian, asian and african american skin. While the values determined was as expected, for example showing higher melanin concentration in african skin, an evaluation with melanoma was not performed.

Many approaches can be found in the related fields of Diffuse Reflectance Spectrography (DRS) which acquires multispectral data at single points. In this field, forward simulations are inverted using lookup tables [88, 137, 166], Levenberg–Marquardt [207], Gauss-Newton [146], phantoms [11] and machine learning [28]. While the challenges are not quite the same, methods in DRS are of interest for translation to multispectral imaging.

3.3.3. EVALUATING FUNCTIONAL APPROACHES

One of the detriments of the functional estimation methods is that they are notoriously hard to evaluate. The main difficulty is that no established reference technique exists,

which can provide functional values such as oxygenation in the whole camera field of view.

For reference values established during the intervention, one can resort to point measurements. These can be taken via blood gas analysis (BGA) by drawing blood from a vessel within the field of view [30]. The main disadvantage is that this directly measures arterial or venous blood from within the vessel, while the MSI measures the average oxygenation in tissue. Furthermore, BGA has to be carefully calibrated, as it is highly temperature sensitive. Another approach for acquiring point measurements are pulse oximeters. These can provide point transmission measurements of arterial blood at a peripheral part of the organ and also suffer from the fact they record arterial blood only [32]. Ground truth oxygenation measurements can be circumvented by correlation analysis to other parameters. This was e.g. done by Best et al. [16], who tried to correlate high or low oxygenation values in partial nephrectomy patients with glomerular filtration rate (GFR) to predict the outcome of the surgery. Kaneko et al. [105] compared local changes in oxygenation between neoplastic and non-neoplastic regions. In this thesis a fluorescence baseline was chosen to evaluate if MSI can detect kidney ischemia after clamping the organ.

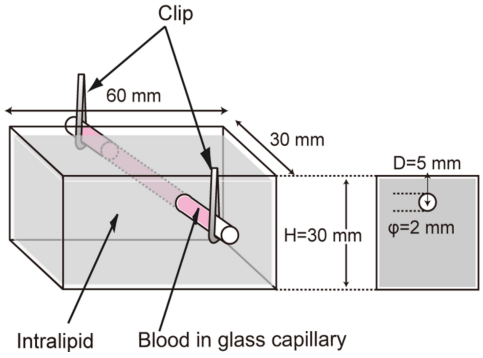


Figure 3.5: Schematic depiction of the oxygenation phantoms from [105]. A capillary was simulated by a glass tube filled with blood, surrounded by intralipid for scattering. Oxygenation was changed by adding sodium hydrosulphite and controlled by transmission measurements. Reprinted with permission from [105].

Reference can also be established by means of tissue mimicking phantoms. These typically combine a scattering medium with an absorber, both are bound by a matrix material. Popular choices for scattering media include milk, intralipid or Polymer microspheres. Matrix materials range from water and agar to RTV silicone, dependent on requirements such as solidity and time stability [151]. Absorption for determining oxygenation should be provided by whole blood, in other cases materials like ink can be sufficient. Sodium hydrosulphite can be used to reduce oxygenation. Phantoms can mimic capillaries by putting blood in a glass tube surrounded by intralipid [105] (see Figure 3.5) or imitate the layered structure of tissue [166]. A phantom study was also conducted in this thesis, Chapter 4.3.2. To date, phantoms are not standardized and measuring the ground truth values for the phantoms is not easy. For an excellent review on phantom creation techniques, please refer to [151].

Yet another option is to utilize highly accurate simulations for validation, as done in this thesis and other works such as [36, 102, 153, 175]. The main advantage of this approach is that ground truth values for the simulations are inherently available. The main disadvantage is that errors go unnoticed if the simulations are not corresponding to reality. Validity of the simulations was ensured by tissue mimicking phantoms and spectrometer measurement in this work.

3.3.4. OTHER TECHNIQUES

Kaneko et al. proposed a more heuristic approach to functional imaging. They recorded three carefully selected bands. One band in the red should capture variations in illumination, caused e.g. by distance and angle to tissue. Red was chosen, because absorption in this region is low thus this band is sensitive mostly to changes in lighting, e.g. caused by distance and angle to tissue. A second broad band in the green spans several isosbestic points (points where oxygenated and deoxygenated blood absorption is the same). They reasoned this band is robust to variations in oxygenation. The third band is a small band centered at 473nm, where oxygenated and deoxygenated hemoglobin absorption is differing and thus sensitive to oxygen. Using the red and the green band, they normalized the 473nm against changes in lighting and total hemoglobin absorption. This approach is capable of real-time oxygenation estimation. Unfortunately it is not clear from the publication how the authors arrive at absolute values for oxygenation, as the system as described should only be able to measure relative differences.

Functional parameter estimation is one way to interpret multispectral images. The main drawback with the approach is the need for a physical model. Options for analyzing to circumvent this need might thus be appealing. One widely adopted approach is taking the spectral measurements directly and train machine learning algorithms like SVM [2, 82, 93, 117, 125, 129] to detect structures of interest. This approach can also be extended to jointly process spectral and spatial information as texture [124, 133, 204]. One downside is that acquiring labeled training data is often not straightforward. In the important example of cancer detection it is challenging to map the histopathological ground truth back to the *in vivo* measurement, because the histopathological analysis involves cutting the tissue and sever deformations [125]. Another disadvantage is that changes to the imaging system, like the camera, necessitate re-acquisition of the training data.

To circumvent the physical modeling *and* the necessity of ground truth label acquisition, unsupervised clustering methods like principal component analysis (PCA) have been proposed [58, 157, 158, 210]. PCA finds a subset of orthogonal axis in data space which minimizes data projection error under Euclidean distance. The resulting images in PCA space are similar to the parametric maps of the functional approach. Downside is that these maps are difficult to interpret [123]. Furthermore, the PCA axes are data dependent, so they might change depending on application, imaging system and data set.

3.3.5. CONCLUSION

Interpretation of multispectral images is an active area of research. Functional parameter estimation has the advantage of displaying information familiar to the physician and with clear physical meaning, such as oxygenation. Identifying the best physical model for this task is not trivial - it must not be too simplistic to capture important relationships, but also shall not offer too many hard to tune degrees of freedom. Direct or unsupervised classification methods can be an alternative to functional imaging, but need recorded data for training the algorithms. In my opinion the choice of algorithm should be task dependent. If data acquisition and labelling is easy I recommend skipping the physical modelling and instead directly classify the image. This option was explored by students supervised by me in the context of automatic organ labelling [133, 204]. If the physician is on the search for indications with a clear physiological interpretation, I recommend functional estimation as done in this thesis.

As of today, approaches for functional estimation in minimally invasive environments were based on linear assumptions [30, 33, 34, 94, 145, 182, 209], with limitations recognized in literature [30, 92, 138, 202]. More sophisticated non-linear inference has been explored in other domains, but has not been translated to endoscopic environments. Most likely reason are difficulties to perform non-linear inversion on high resolution images in video rates. Furthermore, more complicated models rely on correct ranges for several free parameters. Measuring typical ranges for these is complex and thus not always available [18, 98, 175]. In this thesis I present a method to enable video-rate Monte Carlo inversion based on machine learning techniques (chapters 4 and 7). To specifically address the often unknown parameter ranges, the first method to domain adaptation is presented in Chapter 5. This method automatically adapts a general data set to the desired tissue by incorporating real measurements. Validation of functional techniques is challenging and was performed with a multi-stage V-model like fashion, involving *in silico* simulations, phantom measurements and an *in vivo* porcine study (see Chapter 4.3).

3.4. THE BIGGER PICTURE

While this thesis focuses on functional imaging from multispectral images, competing techniques for functional imaging should not go unnoticed. A special focus will be given to techniques which are capable of monitoring perfusion or close to multispectral imaging and clinical translation.

3.4.1. BLOOD GAS ANALYSIS

Modern blood gas analysis (BGA) was pioneered by John W. Severinghaus in the 1950 [165]. Among other things, BGA usually measures partial oxygen pressure by a platinum cathode and a silver/silver chloride anode [1]. At constant temperature, the current resulting from the consumption of oxygen at the cathode is linearly linked to the partial oxygen pressure [1]. Partial oxygen pressure is linked to oxygenation by the oxygen–hemoglobin dissociation curve [37]. The shape of this curve is dependent on factors

like temperature and thus is a possible source of error [30, 37]. BGA devices are largely used by clinical pathologists today [165]. In anesthesiology and critical care, the technique of choice rather is pulse oximetry because it is non-invasive and provides immediate and continuous feedback [37].

3.4.2. PULSE OXIMETRY

The principle of oximetry, first described in [130], performs transmission measurements of at least two wavelengths of light in the red and infrared. Beer-lamberts law is used to calculate oxygenation as explained in Chapter 2.1.1. However, measurements can be disturbed by scattering losses and additional absorbers as melanin [199]. The central idea in pulse oximetry is that while the other absorbers stay constant, arterial blood arrives newly with each heartbeat [199]. By dividing the dynamic part by the static component, one calibrates to only measuring the dynamic part of the signal, and thus only measuring arterial blood [199]. The probes are placed on the finger, earlobe or toe [37]. Today, pulse oximetry is one of the most common monitoring modalities in the OR [104].

3.4.3. DIFFUSE REFLECTANCE SPECTROSCOPY

Pulse oximeters perform transmission measurements. These can be easily taken at extremities like fingers and toes, where light can shine through. However, if one wanted to measure oxygenation directly on site at large organs, a different technique is needed. Diffuse refletance spectroscopy (DRS) measures light diffusely reflected to the surface [166]. Commonly, one fiber delivers light, while collecting fibers at varying distances measure the reflected light. While the instrumental setup is relatively straightforward for DRS [166], analysis involving the diffuse approximation [207] or Monte Carlo methods [146, 166] is still ongoing research. DRS can at least theoretically enable non-invasive point measurements of oxygenation, hemoglobin concentration and thickness of the first epithelial layer [166]. The DRS principle powers commercially available oxygenation measurement probes such as the moorVMS-OXY (Moor Instruments, Cologne, Germany). Concepts of the contact based DRS are similar to multispectral imaging, which becomes apparent by an alternative name for multispectral imaging: imaging spectroscopy [118].

3.4.4. FLUORESCENCE

Both DRS and pulse oximetry can enable continuous point based monitoring of tissue oxygenation. Like MSI, fluorescence is a so called wide-field technique, providing non-contact measurements in a large field of view. In case of laparoscopy this field of view is equal to the normal field of view of the laparoscope. This is a significant advantage, as it allows the physician to monitor e.g. perfusion in a large area at once.

Fluorophores can be excited with light of a suitable wavelength to raise them to a higher energy level [191]. As they relax to the ground state, they may emit light of a slightly different, longer wavelength [191]. This process is called fluorescence and the occurring

wavelength shift is called the Stokes shift. Because of the Stokes shift, it is possible to separate incident and fluorescing light. One can e.g. excite the ICG fluorophore with a 780nm diode and collect the fluorescing light using a bandpass filter at 820nm.

Fluorophores can be endogeneous or exogeneously injected. In autofluorescence endogenous molecules like collagen are excited and might give clues w.r.t. precancerous lesions as some studies suggest [156]. The only exogeneous fluorophore currently approved for clinical use is ICG. It binds to blood and bile fluid and can be identified in up to 1-2cm depth, because it operates in the so called optical near infrared "window" with low absorption from hemoglobin and water. ICG is used in urology and surgery for ischemia and sentinel lymph node detection [6]. Due to its success across many disciplines, laparoscopic vendors offer ICG capable devices[3]. The next frontier for florescence imaging will be moving away from the relatively unspecific ICG and towards tumor binding dyes [203]. These could let tumors "glow" during surgery, potentially solving many practical issues like tumor margin detection and metastasis identification [203]. In 2011, the first in-patient fluorescence-guided surgery with a tumor specific dye was conducted [185].

3.4.5. OTHER TECHNIQUES

Fluorescence introduced a way of light-tissue interaction different from the explained absorption and scattering. Yet another possible, but rare interaction is elastic scattering, also referred to as Raman scattering. *Raman spectroscopy* investigates the shift in wavelength due to elastic scattering [191]. Because molecules have different peaks w.r.t. this shift, Raman spectroscopy encodes specific molecular information, in fact more specific than DRS or fluorescence [40]. However, since Raman scattering is very rare compared to inelastic scattering events, the technique is often referred to as insensitive [24]. Today, there are already endoscopic research probes for point based Raman spectroscopy measurement [144], however, clinical translation has yet to jump several hurdles [26].

Polarization is another fundamental physical property of electromagnetic waves which can be leveraged in imaging. Structural proteins such as collagen change the polarization state of light. As disorganization of the collagen fiber matrix is an important cancer hallmark [91], potential measurement of collagen fiber orientation could give important cues in early cancer detection [66]. First polarization endoscopes have been realized in research settings [154]. Clinical adoption of polarization imaging has to overcome several challenges both from an analytic and experimental point of view [66].

Structural frequency domain imaging (SFDI) is a recent technique first proposed in 2009 [39]. By illuminating the tissue surface with at least two sinusoidal illumination patterns with differing frequencies, one can determine absorption coefficient and the reduced scattering coefficient at the wavelength of the illumination source. Thus this technique can be employed for oxygenation measurement and is based on optical absorption and scattering as multispectral imaging. First in-patient trials showed promising results [68]

[3] Examples: IMAGE1 S, Karl Storz (Tuttlingen, Germany). 1588 AIM Platform, Stryker (Kalamazoo, Michigan, US). da Vinci Si, Intuitive Surgical (Sunnyvale, California, US). PINPOINT, Novadaq (Mississauga, Kanada).

and recent progress was made towards real-time acquisition and processing of SFDI data [186].

When rough surfaces are illuminated with coherent light they show random interference patterns, referred to as *laser speckles* [12]. When small objects like blood cells move, this speckle pattern changes as well [12]. By analyzing the statistics of the speckle movement, in time or in space using a small ROI, one can draw conclusion about blood flow velocities [12]. By measuring blood flow, laser speckle provides additional information compared to solely oxygenation measurements [92]. Laser speckle was first proposed in the 1980 and is most prominently used to analyze brain haemodynamics [19, 44]. Recently, the technique has been translated to the endoscopic domain and tested in an ex vivo setting [160].

Photoacoustic tomography *PAT* (sometimes also referred to as optoacoustic tomography) promises to determine information about hemoglobin absorption deep inside the tissue. PAT is a contact based techniques, which illuminates tissue with laser light and recollects a sound wave with a (conventional) ultrasound device (light in - sound out) [184, 203]. These sound waves are emitted due to thermoelastic expansion happening when by laser pulses are absorbed [15, 192]. If a tunable laser is used, PAT can determine oxygenation similarly to multispectral imaging by measuring absorption at multiple wavelengths, as e.g. done in [112]. This and the high penetration depth compared to other optical techniques are the main advantages of PAT [203]. The actually measured signal is a product of absorption and light fluence. Accounting for the fluence term which disturbs the absorption measurement is an open research question [184]. PAT is a highly active field of research with first commercial systems on the market [184] and recent advantages in transabdominal characterization of e.g. Chron's disease [112].

Perfusion cannot only be assessed with optical techniques. *Doppler ultrasound* e.g. is a contact based technique which can asses perfusion in large blood vessels by examining the doppler shift of the ultrasound signal [38]. Blood-oxygen-level dependent contrast function magnetic resonance imaging (*Bold fMRI*) non-invasively measures oxygen saturation [70]. The magnetic field surrounding blood cells is dependent on the oxygen state of hemoglobin: deoxygenated hemoglobin is paramagnetic due to unpaired electrons while oxygenated hemoglobin is diamagnetic. This change in the magnetic field causes spin relaxation times to change, making oxygenation detectable in the T2* image.

3.4.6. CONCLUSION

Techniques for functional imaging differ considerably. Some notable axes of variation are leveraged physical phenomenon, energy needed, penetration depth, contact requirement and the measured parameter.

Already used in clinical practice, the most advanced method is in the authors opinion fluorescence. The main advantage of MSI is the easier workflow integration, as it does not necessitate injection of a dye. Especially in the common use case of ischemia monitoring MSI can potentially replace fluorescence. Although both MSI and fluorescence

can be used for ischemia detection, the techniques provide complementary information; oxygenation can only be determined by MSI, but fluorescence penetrates tissue deeper and has the unique potential to bind directly to tumors .

Techniques like DRS and BGA are not competing to multispectral imaging, as they offer point-based measurements. However DRS and BGA have been used to evaluate multispectral techniques [30, 68]. PAT, doppler ultrasound and fMRI have radically different workflow integration and thus will be used in differing clinical scenarios. Emerging techniques such as laser speckle, polarization imaging and SFDI could become interesting additions or alternatives to MSI. Which techniques are fit for widespread clinical use remains to be seen.

3.5. CONCLUSION

This chapter reviewed the state of the art with respect to medical application, recording hardware, analysis methods and other techniques for functional imaging.

In the authors opinion, one of the most important needs for MSI is demonstrating the patient benefit after incorporation of the technique into the clinical workflow. The state of the art so far has focused on preliminary, preclinical studies. Some smaller patient trials have been conducted, indicating that MSI can help spotting tumors and characterize perfusion. However, in none of these the real clinical benefit of the technique has been assessed by comparison with a reference technique. The multispectral system developed for this thesis was evaluated by means of a fluorescence reference in Chapter 7. These in-patient results indicate MSI is suited for perfusion monitoring and should be further evaluated in larger scale studies.

Recording of MSI is either done through spectral or spatial scanning. Both methods have disadvantages, either long imaging times or a lower resolution image. Both disadvantages might be mitigated by intelligent selection of the most informative bands, as done in Chapter 6. The authors opinion is that the fast recording snapshot techniques are preferable to slower staring techniques in endoscopic applications. Slow recording leads to misaligned or blurred multispectral images, which makes analysis very challenging. This is more detrimental than missing intricate details due to the low resolution of snapshot techniques. However, the choice of snapshot devices is very limited. Chapter 7 describes the first use of the imec snapshot sensor in laparoscopy, premiering video-rate MSI recording in this field.

The key advantage over traditional non-invasive diagnostic medical imaging devices is that functional imaging can provide live feedback to the surgeon. To realize this advantage, the analysis of the images must be immediate and expressive. So far, MSI analysis in laparoscopy relied on the restrictive assumptions of the Beer-Lambert law. These restriction to not apply for the method proposed in the next two chapters, which combines powerful Monte Carlo methods with state of the art machine learning techniques.

4

MACHINE LEARNING BASED MONTE CARLO INVERSION

Parts of this chapter have been published in the International Journal of Computer Assisted Radiology and Surgery (CARS), Special issue: IPCAI **11**, 6 (2016) [195].

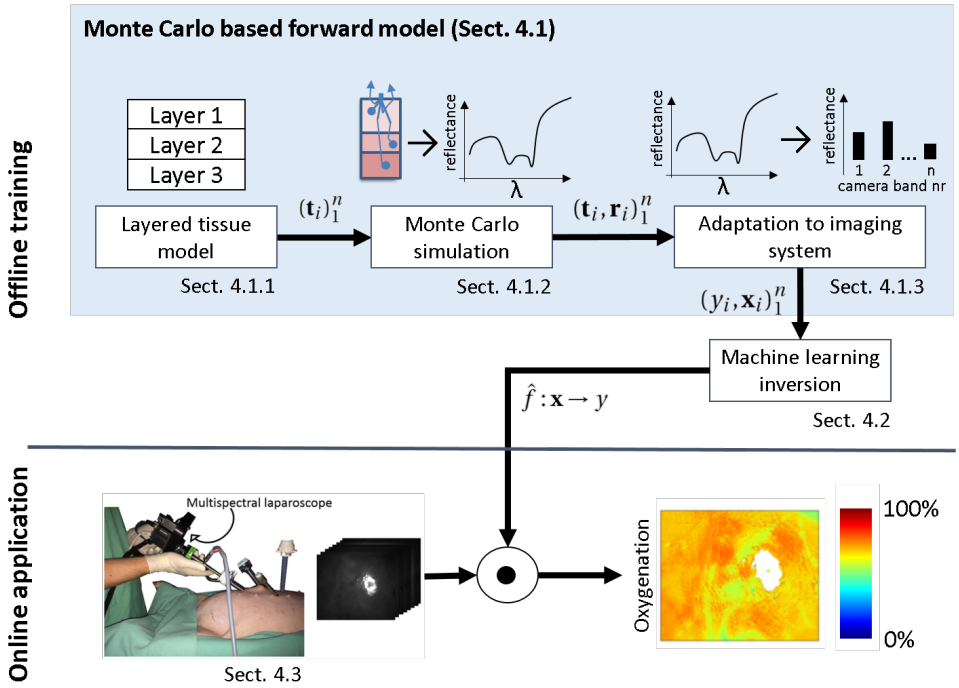


Figure 4.1: Overview on the approach to functional laparoscopy. In the offline phase, n examples $(\mathbf{t}_i)_1^n$ are drawn from a point-wise model of artificial tissue. For each of these samples, highly accurate simulations calculate the spectral reflectance $\mathbf{r}_i$. After transforming this spectral reflectance to band reflectance $\mathbf{x}_i$ space, it is ready to serve as input for a machine learning algorithm. The algorithm learns how to relate the reflectance measurement of a pixel to a functional tissue parameter $y_i \in \mathbf{t}_i$. This functional parameter can e.g. be tissue oxygenation or blood volume fraction. In the online phase, the trained machine learning based regressor $\hat{f}$ is applied to real multispectral images and within milliseconds generates an estimate for each (multispectral) pixel in the image.

PERFUSION visualization is highly relevant for surgical treatments as it can e.g. point out ischemic regions. Changes related to perfusion can also be closely linked to cancer, because important hallmarks of cancer such as hypoxia (deprivation of oxygenated blood) and angiogenesis (formation of new blood vessels) are characterized by changes in oxygenation and blood volume fraction.

Multispectral imaging (MSI) can potentially monitor intricate changes in tissue perfusion, which are quantified by blood volume fraction and oxygenation and cannot easily be seen by the human eye. Deciphering this information, i.e., estimating the molecular composition of tissue on the basis of multispectral images, remains challenging. Most systems for interventional *in vivo* multispectral imaging [30, 105, 138] use linear estimation approaches based on the modified Beer-Lambert law [162]. While fast, these methods are restricted to estimating oxygenation, require small multispectral bandwidths and rely on a number of unrealistic assumptions regarding tissue composition (See Chapter 3.3.2). More sophisticated *ex vivo* tissue analysis methods leverage highly accurate Monte Carlo methods for more accurate assessment of physiological parameters. The

drawback of these methods is that they are too slow for real-time estimation of tissue parameters when dealing with high-resolution multispectral images; they are thus not suited for application during surgery or interventional procedures. For an in-depth discussion of these topics, please refer to Chapter 3.3.

To overcome these issues, the approach presented in the following is the first capable of video-rate estimation of physiological parameters during minimally invasive interventions by combining the accuracy of Monte Carlo based approaches with the speed of state-of-the-art machine learning (ML) algorithms. The central idea is to simulate a reflectance measurement caused by an artificial tissue (Section 4.1) and use these tissue/reflectance pairs as input for ML regression (Section 4.2). The resulting regressor can be applied to multispectral images acquired during an intervention to estimate the tissue parameters for each pixel. Figure 4.1 shows an overview of the presented method. This method was evaluated in a multi-stage process, involving simulations, tissue phantoms, and an in vivo study on six pigs (Section 4.3). The chapter ends with a discussion of the method and experiments in Section 4.4 and the conclusion in Section 4.5.

4.1. THE FORWARD MODEL

The forward model consists of three parts: (1) a layered tissue model formally describes the composition of possible tissues (Section 4.1.1). (2) Highly realistic simulations of spectral reflectances are computed using tissues drawn from the model (Section 4.1.2). (3) Transformation of the simulation to the camera space in Section 4.1.3. This process of sample creation and transformation is represented by the blue box in the offline training part of Figure 4.1.

4.1.1. LAYERED TISSUE MODEL

Inspired by the typical layer structure of the colon and other epithelial tissues, a tissue $\mathbf{t}$ inspected during minimally invasive surgery is modeled as a layered structure. Each layer $\mathbf{l}$ is characterized by a set of optically relevant tissue properties: $\mathbf{l} = \{v_{\text{hb}}, s, a_{\text{mie}}, b, g, n, d\}$. Where the parameters describe the following:

v_{hb} blood volume fraction [%], the amount of blood occupying a unit volume of tissue

s the ratio [%] of oxygen-bound hemoglobin to total hemoglobin, also referred to as oxygenation

a_{mie} a parameter quantifying the amount of scattering [cm^{-1}] by the reduced scattering coefficient at 500nm.

b the scattering power, a term which characterizes the exponential wavelength dependence of the scattering (see Eq. 4.2)

g anisotropy factor, characterizes the directionality of scattering

n the refractive index

d the layer thickness [μm]

the optical and physiological parameters a_{mie}, b, g, v_{hb} and s influence the optical absorption and scattering coefficients. As in [91], the absorption coefficient μ_a at wavelength λ is calculated by

$$\mu_a(v_{\text{hb}}, s, \lambda) = v_{\text{hb}}\, c_{\text{hb}}\, (s \cdot \epsilon_{\text{HbO2}}(\lambda) + (1-s) \cdot \epsilon_{\text{Hb}}(\lambda))\, \ln(10)\, (64{,}500\,\text{g}\,\text{mol}^{-1})^{-1}. \tag{4.1}$$

ϵ_{HbO2} and ϵ_{Hb} are the molar extinction coefficients of oxygenated and de-oxygenated hemoglobin[1] and c_{hb} is the molar concentration of hemoglobin in human blood. As discussed in Chapter 2.1.1, hemoglobin, the oxygen transporter in human blood, can be assumed to be the only notable absorber. As in [36] oxygenation is assumed to be constant across layers. This is a reasonable assumption if the layers share a common blood supply as e.g. in the colon. The reduced scattering coefficient μ_s' is calculated by an empirical power law

$$\mu_s'(a_{\text{mie}}, b, \lambda) = a_{\text{mie}}(\frac{\lambda}{500\text{nm}})^{-b}. \tag{4.2}$$

This simple heuristic approximation to measurements of reduced scattering was proposed in [98] for the visible wavelength range. The anisotropy factor g is assumed constant over the wavelength range. With g and μ_s' , the scattering coefficient μ_s can be calculated by $\mu_s(a_{\text{mie}}, b, \lambda) = \frac{\mu_s'(a_{\text{mie}}, b, \lambda)}{1-g}$ [98]. Refer to Chapter 2.1 for more details on the above parameters. Pixel independence is implicitly assumed, by modelling tissue as homogeneous, infinite slabs. This leads to less assumptions on the 3D composition of tissue as e.g. vessels, but prevents modelling cross-talk between pixels.

A TISSUE MODEL FOR COLONIC TISSUE

The tissue instances $\mathbf{t}$ have to be drawn from tissues T expected during intervention. If e.g. colon anastomosis success verification is the target application, a model for colonic tissue is needed. An exemplary model is shown in table 4.1. The values were chosen to mimic colonic tissue literature when available and drawn randomly from these parameter ranges to form $\mathbf{t}$. The molar hemoglobin concentration c_{hb} was set to $120\text{g}\text{L}^{-1}$, a typical value in the colon as opposed to $150\text{g}\text{L}^{-1}$ in general human tissue [91].

4.1.2. CALCULATION OF SPECTRAL REFLECTANCE

The spectral reflectance denotes the portion of light returning to the surface for each wavelength. This quantity is a property of the tissue and independent on the measurement system. To generate such a reflectance spectrum $\mathbf{r}$ from a tissue drawn from the layered model, a function r_{sim} is evaluated at wavelengths λ

$$r(\lambda, \mathbf{t}) = r_{\text{sim}}(\lambda, \mathbf{t}) = r_{\text{sim}}(\lambda, \mathbf{l}_1, ..., \mathbf{l}_k). \tag{4.3}$$

[1] Values taken from [152]

Table 4.1: Parameter ranges for the colon tissue model and their usage in the simulation set-up as described in Section 4.1.2.

	v_{hb}[%] [36]	s[%]	$a_{\text{mie}}[\frac{1}{\text{cm}}]$ [98]	b_{mie} [98]	g	n [36]	d[mm] [91]
layer 1 :	0-10	0-100	18.9±10.2	1.289	.8-.95	1.36	.6-1.1
layer 2 :	0-10	0-100	18.9±10.2	1.289	.8-.95	1.36	.4-.8
layer 3 :	0-10	0-100	18.9±10.2	1.289	.8-.95	1.38	.4-.6

$\mu_a(v_{\text{hb}}, s, \lambda) = v_{\text{hb}}\,(s \cdot \epsilon_{\text{HbO2}}(\lambda) + (1-s)\cdot\epsilon_{\text{Hb}}(\lambda))\,\ln(10)\,120\text{gL}^{-1}\,(64{,}500\text{g mol}^{-1})^{-1}$
$\mu_s(a_{\text{mie}}, b, \lambda) = \frac{a_{\text{mie}}}{1-g}(\frac{\lambda}{500\text{nm}})^{-b_{\text{mie}}}$

simulation framework: GPU-MCML[7], 10^6 photons per simulation
simulated samples: 10K
sample wavelength range: 450-700nm, stepsize 2nm

In this thesis, a multi-layered MC (MCML) approach was chosen for evaluation of r_{sim}, because MC models are widely considered to be the gold standard for calculating how light travels through biological tissue (see Chapter 3.3.2). The MC tissue optics simulation irradiates multi-layered tissue with photon packets [190]. Depending on the properties of the layers, the photons will be probabilistically reflected, scattered and absorbed. Among other attributes, the photons reflected at the tissue surface due to (possibly multiple) scattering events can then be measured.

The MC simulation was evaluated in the [450nm, 720nm] wavelength interval in 2nm steps. The open-source GPU-MCML implementation from [7] was used as simulation framework. The number of photon packets fired was set to 10^6 in all simulations and the diffuse reflectance was taken as the reflectance value.

4.1.3. ADAPTATION TO IMAGING SYSTEM

The simulations evaluated in the last section yield spectral reflectances at specific wavelengths. The spectral reflectance is an inherent property of the tissue and independent on the measurement system and lighting. In this section the camera independent spectral reflectances are converted to values proportional to image intensities measured by the camera.

The multispectral sensor integrates this reflectance over the bands $b_j \in \mathbf{b}$. The camera intensity measurement i_j is dependent on the following factors:

$r(\lambda, \mathbf{t})$ the spectral reflectance for the tissue.

$l_j(\lambda, \mathbf{p})$ linear factors subsuming the response of b_j, the quantum efficiency of the camera, the irradiance of the illuminant (e.g., Xenon or Halogen) and other components in the imaging system like the transmittance of the laparoscope optic [175]. Can be dependent on the position in the image $\mathbf{p}$.

$\tau(\cdot)$ the camera response function, often intentionally non-linear to cover a higher

range of irradiances within the possible image intensities [76]

w_j camera noise for band j, mainly coming from (1) shot noise due to the particle nature of light and (2) dark noise, which is dependent on sensor temperature and linear in integration time [86]. The constant offset caused by the dark noise is assumed to be zero in the following, as it can be removed by simple subtraction of a dark image in practice.

For large image intensities, the often Poisson distributed noise sources can be approximated by multiplicative Gaussian noise [86, 107, 127]. Because all reasoning is done on pixel level, geometric distortions caused by lens effects are not relevant. Putting all of this together, an image intensity can be simulated by

$$i_j(\mathbf{t}) = \tau\left(w_j \cdot \int_{\lambda_{\min}}^{\lambda_{\max}} l_j(\lambda, \mathbf{p}) \cdot r(\lambda, \mathbf{t})\,\mathrm{d}\lambda\right). \tag{4.4}$$

Assuming $\tau(\cdot)$ is linear and that the spatial inhomogeneities of the light source are independent of wavelength $l_j(\lambda, \mathbf{p}) = \alpha(\mathbf{p}) l_j(\lambda)$ the model simplifies to

$$i_j(\mathbf{t}) = \alpha(\mathbf{p}) \cdot w_j \cdot \int_{\lambda_{\min}}^{\lambda_{\max}} l_j(\lambda) \cdot r(\lambda, \mathbf{t})\,\mathrm{d}\lambda \tag{4.5}$$

The factor $\alpha(\mathbf{p})$ accounts for constant multiplicative changes in intensity. The reason for these changes are differences in distance or angle of the camera to the tissue and the internal scaling of electrical current to values measured by a camera [36].

4.2. MACHINE LEARNING BASED INVERSION

With the method described so far, simulated image intensities depend on the image brightness $\alpha(\mathbf{p})$ and the spectral distribution of light reaching the sensor $l_j(\lambda)$.

To make the data more robust to changes in $l_j(\lambda)$, each band intensity is divided by the its volume $1_j = \int_{\lambda_{\min}}^{\lambda_{\max}} l_j(\lambda)\,\mathrm{d}\lambda$. The results is referred to as band reflectance in this thesis:

$$r_j(\mathbf{t}) = \frac{i_j(\mathbf{t})}{1_j} = \alpha(\mathbf{p}) \cdot w_j \cdot \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{l_j(\lambda)}{\int_{\lambda_{\min}}^{\lambda_{\max}} l_j(\lambda)\,\mathrm{d}\lambda} \cdot r(\lambda, \mathbf{t})\,\mathrm{d}\lambda. \tag{4.6}$$

This corresponds to the calibration with a white target for real measurements described in Chapter 3.3.1.

The multiplicative $\alpha(\mathbf{p})$ is accounted for by normalizing with the sum of all bands, transforming the data to what is further referred to as normalized reflectance:

$$r_{j,n}(\mathbf{t}) = \frac{\alpha(\mathbf{p}) r_j(\mathbf{t})}{\alpha(\mathbf{p}) \sum_k r_k(\mathbf{t})} = \frac{r_j(\mathbf{t})}{\sum_k r_k(\mathbf{t})}. \tag{4.7}$$

thus removing dependence on $\alpha(\mathbf{p})$. As a final transformation – log is taken from the normalized reflectances and the resulting values are divided by their ℓ^2 norm, transforming

the data to what is called normalized absorbance x_j in the following:

$$x_j(\mathbf{t}) = \frac{-\log r_{j,n}(\mathbf{t})}{\sqrt{\sum_k (-\log(r_{k,n}(\mathbf{t})))^2}}. \tag{4.8}$$

This last transformation is strictly speaking not necessary, but leads to better experimental results in regression. Given a noise model W and a tissue T, the training input data for the regressor $\mathbf{X} = (\mathbf{x}(\mathbf{t}_1), ..., \mathbf{x}(\mathbf{t}_m))^T$ can be simulated.

Any machine learning regressor can be trained for specific tissue parameters $y \in \mathbf{t}$, such as oxygenation and blood volume fraction. A machine learning regressor is a function f_θ, parametrized by parameters θ. It maps input features, in this case the described $\mathbf{x}$, to labels y, in this case functional parameters. Machine learning methods aim to find the parameters which minimize the expected risk R:

$$\hat{f} = \arg\min_{f_\theta} R(f_\theta, X_{\text{train}}, \mathbf{y}_{\text{train}}, l(\cdot,\cdot)) = \frac{1}{|X_{\text{train}}|} \sum l(f_\theta(\mathbf{x}_{\text{train}}), y_{\text{train}}) \tag{4.9}$$

It thus tries to determine the function which maps the training features X_{train} closest to the training labels $\mathbf{y}_{\text{train}}$. The notion of closeness is provided by the real valued loss l, which is often amended by a regularizing term to prevent overfitting to the training data. The functions, often also called hypothesis, and minimization technique differ substantially depending on method.

In the example of random regression forests, the functions are defined by an ensemble of decision trees. The path a feature traverses in a tree depends on comparisons taking place in the nodes. The output of the function is calculated from the leafs the feature landed in. Because the determination of the optimal trees is NP-complete, learning the trees is usually done by a greedy algorithm, which iteratively adds nodes [23]. Due to the multidisciplinary focus of this thesis, a detailed description of machine learning and the used methods is beyond the scope of this thesis. Please refer to good books on the subject, such as [134].

APPLICATION TO *in vivo* RECORDINGS

The trained regressor can be applied to each pixel of an *in vivo* acquired multispectral image. This image needs to undergo the same transformations as the simulations. First the images $\mathbf{I}$ are transformed to normalized reflectance by correcting by a dark $\mathbf{D}$ and white image $\mathbf{W}$: $\mathbf{X}_I = \frac{\mathbf{I}-\mathbf{D}}{\mathbf{W}-\mathbf{D}}$ and dividing by the mean in the spectral dimension (also see Chapter 3.3.1). As the simulations in Equation 4.8, the data is further transformed to absorbance and normalized by an additional ℓ^2 norm. This ensures real recordings and simulations are in the same space, which is independent of multiplicative changes in illumination.

4.3. EXPERIMENTS AND RESULTS

The experiments assess the approach in a V-model-like fashion, depicted in Fig. 4.2. This section starts with the evaluation of the spectral reflectance simulations, on which all

other reasoning is based on in Section 4.3.2. Following the transformation from spectral reflectance to normalized reflectance space is investigated in Section 4.3.3. Ultimately, Section 4.3.4 investigates the inversion method, both *in silico* and *in vivo* in a porcine study.

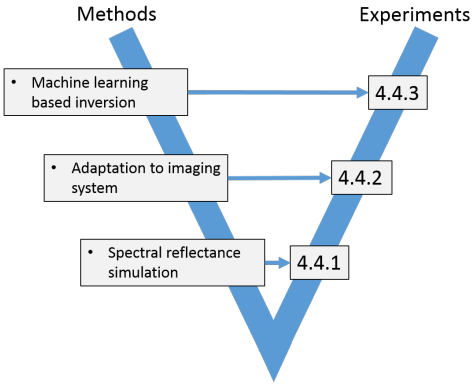


Figure 4.2: Experiments have been performed in a v-model-like fashion. The machine learning builds on band reflectances, which build on the spectral reflectances; thus experiments have been performed in the reverse order to ensure the basis for the methods on top is sound.

4.3.1. GENERAL SETTING

Closeness of spectra is measured with the ℓ^2 norm. Variations with standard deviation σ around a mean μ (e.g. around a mean camera measurement) are given in signal-to-noise ratio SNR:= $\frac{\mu}{\sigma}$.

SIMULATIONS

If not mentioned otherwise, all experiments were conducted with 15,000 training samples generated from the model specified in Section 4.1.1 and tested with a separate set of 5,000 samples. A sample consists of a pixel-wise band reflectance measurement $\mathbf{x}$ and the functional parameters $\mathbf{y} = (v_{hb}, s)^T$.

IN VIVO RECORDINGS

A custom-built multispectral laparoscope was constructed during Justin Iszatt's Bachelor thesis, supervised by me. The system combines a Richard Wolf (Knittlingen, Germany) laparoscope and light source with a Pixelteq (Largo, FL, USA) 5 Mpix VIS-NIR Spectrocam (see Fig. 4.3). To limit the effects of chromatic aberration, achromatic lenses were used. The Spectrocam is fast filter wheel based system with eight filter slots, synchronized with a monochromatic, integrated camera which records 12bit images. The filters can be chosen from a pool of narrow or broad filters of available filters or customized. The eight selected filters for this work had central wavelengths of 470nm, 480nm,

511nm, 560nm, 580nm, 600nm, 660nm and 700nm, determined with the method from Chapter 6. The full width at half maximum of the bands is 20nm, except for the 480nm band where it is 25nm. The sequential acquisition of one multispectral image stack takes 400ms and images are downsampled images to one fourth of the original size. This leads to a resolution of 1228x1029 pixels, which is similar to modern laparoscopic HD optics. Contrary to [30], no further post-processing in the form of image registration or Gaussian smoothing was performed. The calibration steps described in Section 4.2 were performed, for which the dark image was recorded by covering the laparoscopic optic and the white image was taken using a Labsphere (New Hampshire, US) Spectralon® white reference target.

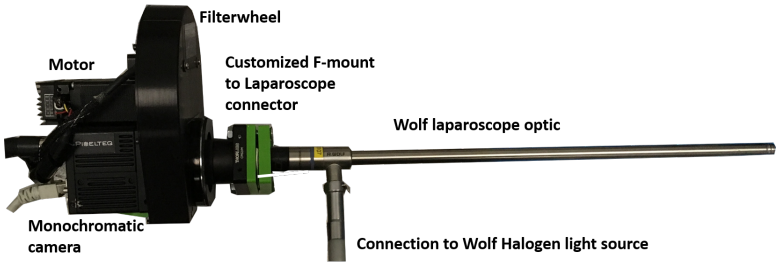


Figure 4.3: The multispectral laparoscope, constructed during the Bachelor thesis of Justin Iszatt. It combines the fast filter wheel based Spectrocam with a conventional laparoscope. The connection from laparoscope to F-mount is customized and houses an achromatic lens. The eight filters of the Spectrocam were chosen with the method from Chapter 6.

Camera noise estimation. A rough estimate of the camera's SNR was determined by calculating the mean intensities for each band, using all downsampled images acquired during one experiment. The mean intensities were subtracted by the dark current and divided this result by a camera noise estimate for the determined mean intensities. This leads to SNRs ranging from 29 for the 470nm band to 47 for the 660nm band. The differences in SNR are mainly caused by the hemoglobin absorption, light source spectrum and camera quantum efficiency. To train the random forest, the SNR was set to ten to account for variations caused by camera-tissue movements. This is also motivated by the *in silico* results, which indicated that the method is relatively robust to conservatively estimated noise levels.

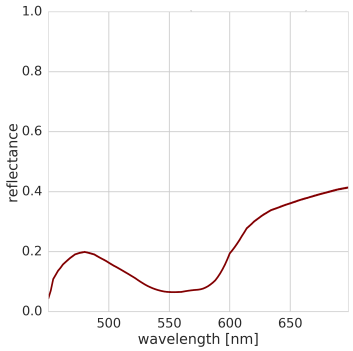
4.3.2. SPECTRAL REFLECTANCE

The spectral reflectances were created from a layered tissue model (Section 4.1.1) with Monte Carlo simulations (Section 4.1.2). The experiments in this section investigate the variations in the simulations and how they compare to spectrometer measurements from artificial tissue.

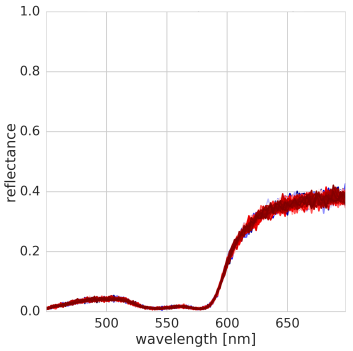
MONTE CARLO SIMULATION VARIATION

The purpose of this evaluation is determining the noise level which has to be expected due to the random nature of MC.

Ten samples were drawn from the colon tissue model specified in table 4.1 and each simulation was performed 100 times. The minimum measured SNR for one wavelength over all simulated spectra was 185. Figure 4.4a shows the corresponding spectra. For comparison, the same evaluation was done with a factor of 1000 less photons. This leads to a drop in minimum SNR to 5, shown in Fig. 4.4b.



(a) The spectrum with the worst SNR (10^6 photons)



(b) The spectrum with the worst SNR (10^3 photons)

Figure 4.4: Investigation of noise in the MC simulation process. The left image shows the worst result for the simulated number of photons. Since the 100 spectra almost perfectly overlap, almost no difference can be seen. As comparison, the right image shows an example with a factor of 1,000 less photons.

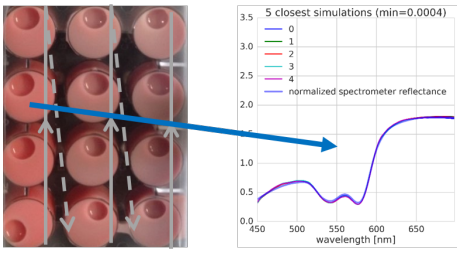
PHANTOM EXPERIMENTS

To make sure the spectral reflectance simulation are corresponding to real measurements, they were compared to tissue phantoms.

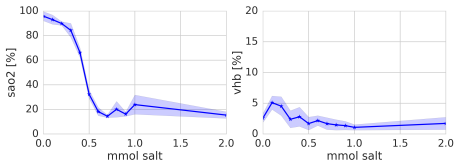
In a joint work with Fraunhofer PAMB, liquid phantoms were created in Jasmin Mangei's Bachelor Thesis. Fresh porcine blood served as absorber, scattering was simulated using intralipid. After flushing with pure oxygen, the saturation was controlled by adding sodium hydrosulphite salt in different concentration. Three well plates were created, featuring different amounts of blood and intralipid. In each plate oxygenation was varied in twelve wells, starting from full oxygenation. The amount of blood and intralipid chosen for plate one mimics typical colonic tissue. Plate two and three vary these parameters by increasing blood volume (plate two) and scattering (plate three).

A spectrometer recorded the reflectance created by each of the wells. The 15,000 simulations created by the colonic tissue model specified in table 4.1 were compared to these reflectance measurements to test if the created simulations correspond to the phantom

reflectances. The worst mean squared difference was 0.0048 (median: 0.00076). A summary of the results for the three plates can be found in Figures 4.5, 4.6 and 4.7.



(a) Plate 1 (2% blood, 1% intralipid) and the best overall match



(b) Plate 1 parameter estimates

Figure 4.5: The first well plate with normal blood volume fraction and scattering values. (a) The best fit of simulations to spectrometer measurements. The direction of increasing amounts of deoxygenating salt is shown by the grey lines. (b) shows the oxygenation (sao2) and blood volume fraction (vhb) mean and standard deviation estimated from the five best fits. As expected oxygenation drops from full oxygenation. Blood volume fraction should remain constant but varies around the (correct) values of 2%.

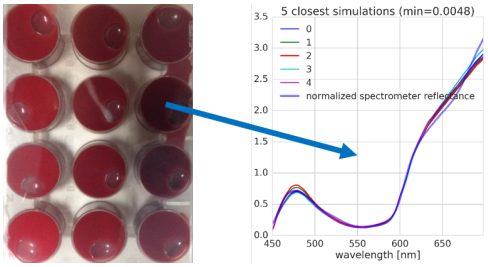
4.3.3. ADAPTATION TO IMAGING SYSTEM

Realistic spectral reflectances are an important first step. However, the transformations happening in the image recording process need to be also taken into account when learning from simulations. This section investigates the assumptions made in Section 4.1.3.

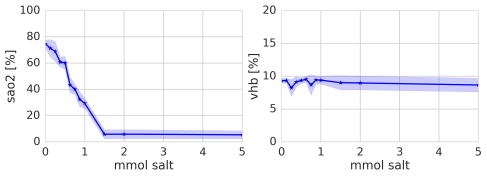
LIGHT SOURCE

Moving from equation 4.4 to equation 4.5 assumes spatial and spectral independence of the illumination: $l_j(\lambda, \mathbf{p}) = l_j(\lambda)a(\mathbf{p})$. This is true if the normalized light irradiance is independent on the exiting angle from the laparoscope tip. To test for this the laparoscope was attached about 7cm above a white tile. The spectral response of the white tile was measured by spatially sweeping the spectrometer over the tile. If the independence condition is met, the normalized spectrometer counts are equal independent on position on the homogeneous diffuse surface, and thus independent on the exit angle of light.

Figure 4.8 shows both the normalized and the not normalized measurements. The median SNR of the normalized measurement was 39 (minimum: 11) in the 460-710nm range.



(a) Plate 2 (15% blood, 1% intralipid) and the worst match of all well plates



(b) Plate 2 parameter estimates

Figure 4.6: The second well plate with increased amount of blood (vhb). The 15% of blood in the well plates exceeds the highest simulated amount of 10% (a) The overall worst fit of measurements to simulations, probably caused by insufficient range of simulations (b) A drop in oxygenation (sao2) can be detected, but does not start from the expected full sao2. Blood volume fraction peaks at 10%, which is the highest value in the simulation database (see Table 4.1).

CAMERA NOISE AND LINEARITY

A linear camera response is assumed in equation 4.5. To check, the laparoscope was placed about 7cm above a white tile and the integration time was varied to record typical intensity values encountered during in vivo scenarios. One hundred recordings of a central 10x10 ROI per integration time were taken to get a stable estimate for the mean intensity and SNR.

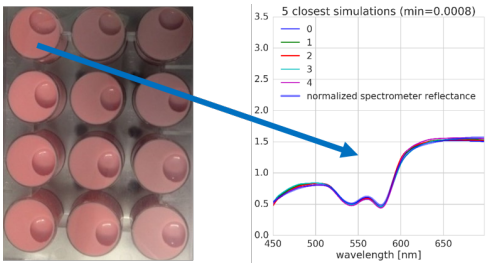
Figure 4.9a shows the line fit to the mean recordings and the SNR for different exposure times. Figure 4.9b shows the worst fit of these recordings to a normal distribution, determined by a D'Agostino and Pearson's based test[2].

IMAGING SYSTEM

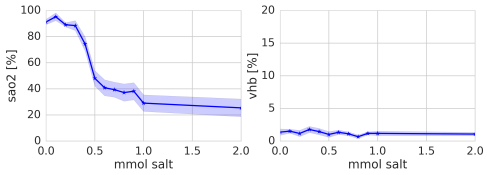
Light exiting from the light source is reflected on the tissue surface, passes through the laparoscope, the filter and finally reaches the camera, where it is transformed into electric current. This experiment evaluates these transformations are correctly represented in the model.

The relative irradiance of the light source and the transmissions of the laparoscope and of the filters were measured. Quantum efficiency of the camera was taken from its ref-

[2]Tested with the scipy.stats.mstats.normaltest function of scipy v0.19 [101]



(a) Plate 3 (2% blood, 3% intralipid) and the median match over all well plates



(b) Plate 3 parameter estimates

Figure 4.7: The third well plate with increased amount of scattering. (a) The median match of measurements to simulations over all well plates shows how the simulations and the measurement coincide. (b) The drop in oxygenation (sao2) and the values close to the correct amount of blood (vhb) were estimated from the five best fits.

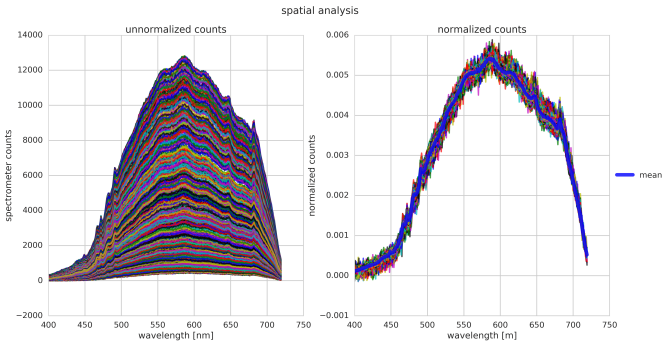
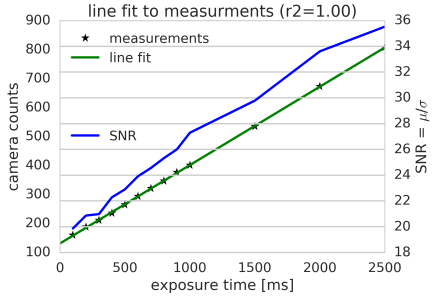


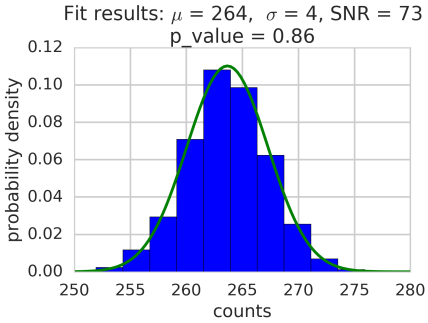
Figure 4.8: Raw and normalized spectrometer measurements of light source under varying light exit angles. The median SNR of the normalized spectra to the mean spectrum is 39.

erence sheet. To test if these components are correctly taken into account, color tiles[3] with known reflectance were measured with the camera. For each color tile, the transformation described in equation 4.7 was applied without adding noise and compared to the normalized mean of ten camera measurements. Figure 4.10 shows an overview on these matchings for all color tiles. The mean error of all color tile measurements is -0.7%

[3] Measured from a X-Rite (Grand Rapids, MI, USA) ColorChecker classic



(a) Mean measurements, line fit, and the SNR



(b) Worst normal fit. The measurement histogram is depicted in blue, the Gaussian induced by mean and standard deviation is overlayed in green.

Figure 4.9: (a) Measuring linearity for typical ranges of camera brightness commonly encountered during interventions. The line fits to the measurements almost perfectly, indicating the camera is linear. SNRs, determined from 100 repeat measurements of a central 10x10 ROI each, grow approximately linear with camera brightness. (b) shows the measurements which fit worst to a normal distribution (500ms exposure time).

and the standard deviation is 5.4%.

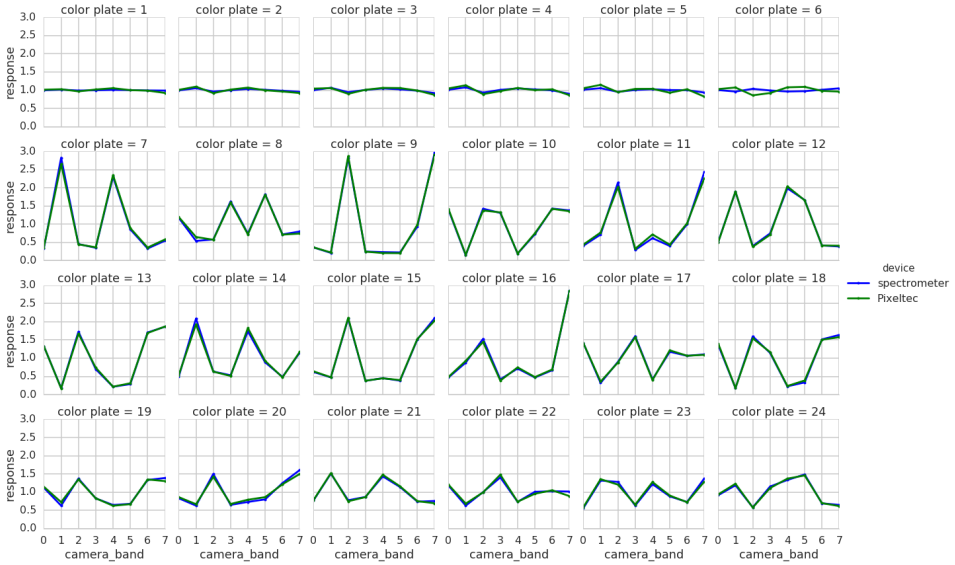


Figure 4.10: Color tiles measured with spectrometer and camera. The artificial optical system was applied to the spectrometer measurements to transform them into camera space. The first row are color tiles of constant reflection, starting with 99% reflectance in the upper left.

SIMULATIONS MATCHING TO IN VIVO

In a final test it is investigated if simulations and *in vivo* measurements occupy the same subspace.

For this purpose, bowel images from six pigs were collected and measurements of a 100x100 center ROI were extracted and dark and white corrected (Eq. 3.1). Simulations created as in table 4.1 were transformed following equation 4.8 and Gaussian noise w with a mean of 1 and standard deviation of 0.1 was added to simulate a SNR of 10.

The first three principal components explain 97% (=79%+12%+6%) of the simulated variance. For *in vivo* data, 95% (=89%+3%+3%) of the variance lies on first three principal components of the simulated data. For qualitative assessment the *in vivo* measurements were projected on the first two principal components of the simulated data in Fig. 4.11. The real bowel measurements lie close to the mode of the simulations in this subspace.

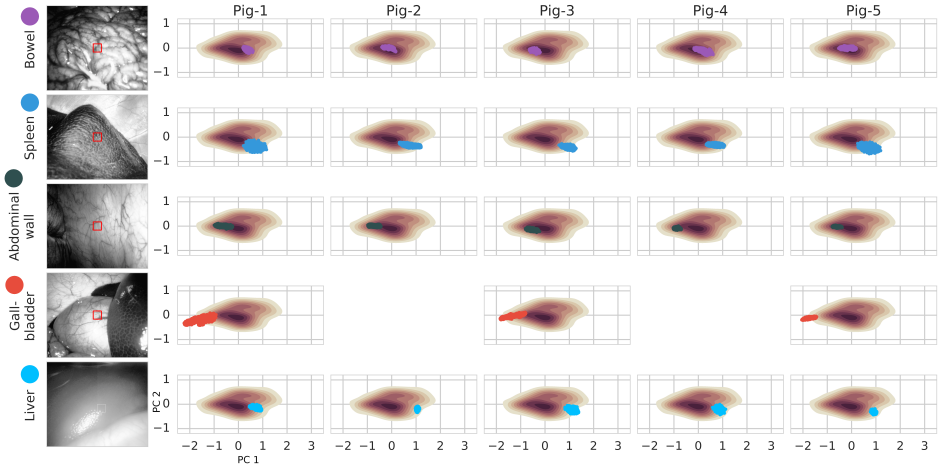


Figure 4.11: Bowel, spleen, abdominal wall, gallbladder and liver measurements from five pigs projected onto the first two principal components of the simulated reflectance data plotted in brown. The images on the left show the 560nm band recorded for the first pig. The depicted measurements are taken from the red ROI. For some pigs, no Gallbladder measurements were performed. The bowel measurements lie close to the mode of the simulated data. The other organs lie more to the boundaries of the simulated density estimates. This is the expected result, as the simulations specified in Table 4.1 were created to mimic colonic tissue.

4.3.4. INVERSE MODEL

The previous experiments were designed to ensure the simulations resemble real tissue and that the measurement process is understood. The experiments in this section examine the inversion performance. A set of *in silico* (Section 4.3.4) and *in vivo* (Section 4.3.4) experiments assesses the accuracy, robustness and run-time of the new method compared to the widely applied linear Beer-Lambert regression model described in 3.3.2.

INVERSION METHOD AND TRAINING

To validate our approach to functional estimation, a prototype was implemented within the Python scikit-learn framework [149]. Random forest (RF) regression [23] was chosen for the inversion (see Chapter 4.2), due to its rapid evaluation and training speed. The RF parameters were determined by five-fold cross-validation and grid search on the training data, varying the maximum forest depth from three to ten in increments of one. The minimum samples per leaf were evaluated for one, five, ten, 20 and 100. The number of trees was set to ten to keep the computational effort manageable. Experiments with larger and deeper forests showed no change in performance. The best forest parameter setting had a depth of nine and a minimum of ten samples per leaf. These parameters were used in the subsequent experiments.

The noise for each band was varied by modifying the signal-to-noise ratio SNR: $\sigma_{w_j} = \frac{x_j}{\mathrm{SNR}}$. If not mentioned otherwise, the SNR was set to ten. To simulate a typical multi-spectral camera, the spectrum was parsed in 10nm increments from 470 – 680nm. For each of these central wavelengths a 10nm sliding average simulated the filter bandwidth $b_k(\lambda)$.

IN SILICO QUANTITATIVE VALIDATION

In the *in silico* experiments it was investigated how factors like the number of samples, noise and domain switch influence the regression result.

Dependency on noise. In these experiments, Gaussian noise of equation 4.8 was varied. First the performance of the classifier while adding noise of the same distribution for training and testing data (Fig 4.12) was investigated. For SNRs above ten our approach outperforms the linear Beer-Lambert approach. Since the determination of noise is not trivial the effect of different magnitudes of training and testing noise (Fig 4.13) was also investigated. Even under these conditions our approach showed lower errors for SNRs above 20.

As v_{hb} cannot be measured by the baseline method, no comparison has been made here. The median absolute error stays below $1.6 \pm 1.2\%$ regardless of noise and below $1 \pm 1.1\%$ for SNRs larger than 10.

Performance under domain switch. Knowledge about the model may not always be available. Furthermore, malignancies can change the parameter ranges. An example of such a malignancy are carcinomas which, due to angiogenesis, may show abnormally high values for v_{hb}. To test the effect of different parameter ranges, samples were created from a second model with the values specified in table 4.2. As in the colon model oxygenation was assumed constant across layers. As an additional constraint all layers were normalized so they totaled a maximum of 2mm in depth.

When training with this model, and testing on the colon model the median absolute errors of the proposed method and the baseline are the same. The proposed method's

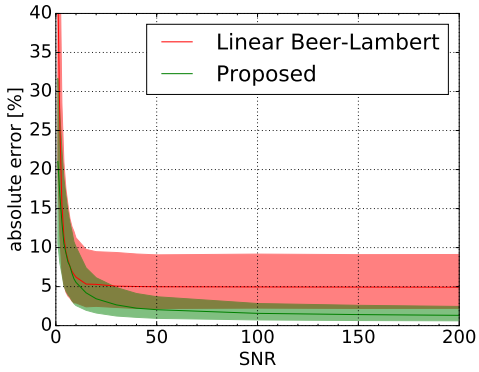


Figure 4.12: Absolute oxygenation estimation error if training and testing data are based on the same noise distribution: zero-mean Gaussian noise with standard deviation corresponding to the SNR. Depicted are the median and interquartile ranges as a function of the amount of noise for baseline and the proposed random forest based method with 15,000 training samples. Reprinted with permission from Wirkert et al. [195].

Table 4.2: Parameter ranges for the generic tissue model.

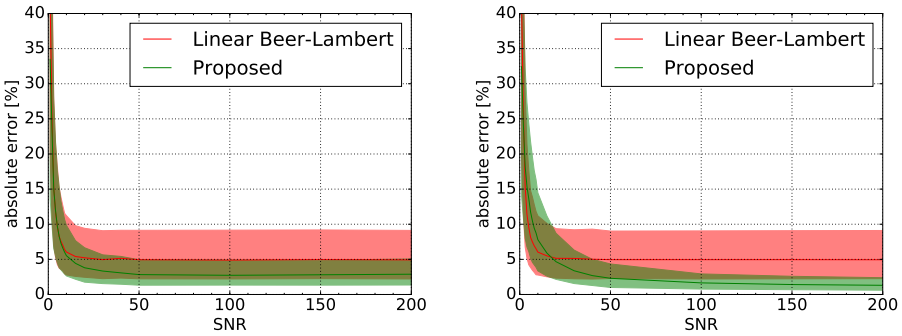
	v_{hb}	s	a_{mie}	g	n	d
$\mathbf{l}_1$	0-100%	0-100%	$18.9 \pm 10.2\text{cm}^{-1}$ [98]	0.8-0.95	1.33-1.54 [98]	0-2000μm
$\mathbf{l}_2$	0-100%	0-100%	$18.9 \pm 10.2\text{cm}^{-1}$ [98]	0.8-0.95	1.33-1.54 [98]	0-2000μm
$\mathbf{l}_3$	0-100%	0-100%	$18.9 \pm 10.2\text{cm}^{-1}$ [98]	0.8-0.95	1.33-1.54 [98]	0-2000μm

75% quartile error is 3% higher and the 25% quartile error is 0.7% lower than the baseline.

Accuracy and run-time compared to state-of-the-art regression methods. The Python implementations (scikit-learn v0.18.1 [149]) of the proposed random forest regression were compared with with Support Vector Regression (svr) and k-nearest neighbor (knn). The parameters for svr were selected with grid search to be the radial basis function kernel (rbf) with $C = 100$ and $\gamma = 10$. Five neighbors were used for knn as in [188] and the algorithm parameter of the scikit implementation was set to auto. The experiment was conducted on a Intel Core™i7 CPU@3.20GHzx6 machine.

As can be seen in Fig. 4.14, the random forest regression is at least two orders of magnitudes faster than either of the baseline approaches. It took 0.18±0.01 seconds to evaluate s and v_{hb} for a megapixel image. The median absolute error was 5.4% for rf, 4.8% for svr and 5.5% for knn. The linear Beer-Lambert method evaluates one megapixel in 0.03 seconds.

Number of samples. As the generation of training data is time consuming even with the graphics processing unit (GPU) accelerated MC simulation used in this publication, the data necessary for training the regressor was analyzed. Figure 4.15 shows stabilization after training with about 10^4 samples.



(a) varying testing noise, fixed training to a SNR of ten

(b) varying testing noise, fixed training to a SNR of 200

Figure 4.13: Effect of unequal test and training noise. Depicted are the median absolute error in oxygenation estimation and corresponding interquartile ranges as a function of the amount of noise for baseline and the proposed random forest based method with 15,000 training samples. Reprinted with permission from Wirkert et al. [195].

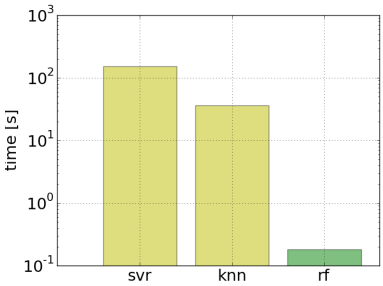


Figure 4.14: Mean computation time for estimation oxygenation for a multispectral imaging stack of dimension 1,000x1,000x8, where eight is the number of spectral bands used. Compared are the proposed random forest approach (rf), Support Vector Regression (svr) and k-nearest neighbors (knn). Reprinted with permission from Wirkert et al. [195].

QUALITATIVE IN VIVO EXPERIMENTS

When blood supply to tissue is stopped, both oxygenation and blood volume fraction should go down. The qualitative *in vivo* experiments, conducted in collaboration with the Department for Visceral Surgery of the Heidelberg University Hospital, investigated if the expected drop could be measured.

Small bowel experiment. In a first experiment, blood supply to a small bowel segment was stopped in a porcine model by clipping three vessels connected to the segment. To simulate real operating conditions the multispectral video of the breathing swine was recorded while moving the camera and the instruments. Mean oxygenation of the clipped bowel segment was determined by manually segmenting the bowel seg-

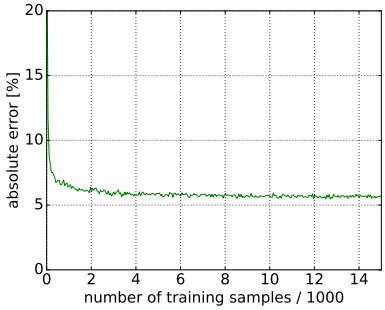
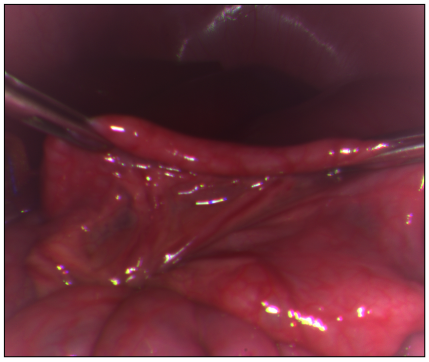
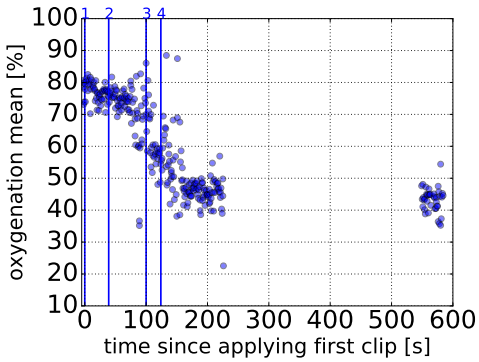


Figure 4.15: Median absolute error in oxygenation estimation as a function of the number of training samples. The results stabilize after training with about 10^4 samples. Reprinted with permission from Wirkert et al. [195].

ment and excluding specular regions by thresholding. A sharp drop in oxygenation was detected after clipping. Figure 4.16b visualizes the estimation result derived from all 315 multispectral image stacks of the recorded video. Example frames during clipping are shown in Figure 4.17.



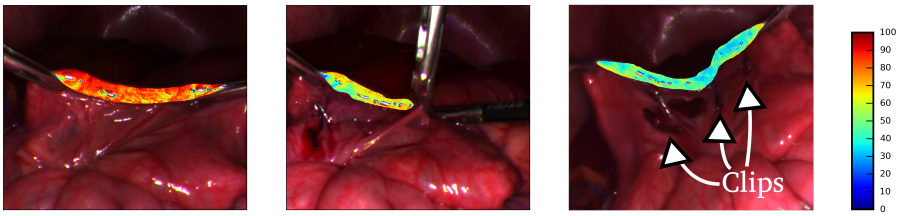
(a) bowel before clipping



(b) oxygenation

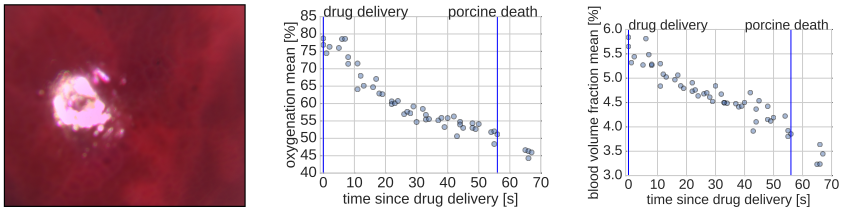
Figure 4.16: Qualitative validation of the proposed method in a porcine small bowel (a). The plot (b) shows the mean oxygenation in the small bowel segment as a function of time. The vertical lines show the time points at which the three vessels were clipped. The first two clips were applied to the same vessel. The high variations before setting the third clip and after setting the forth clip are caused by fast camera movements. After setting the last clip, the camera was removed and reinserted about nine minutes after setting the first clip. Example frames during clipping are shown in Figure 4.17. Reprinted with permission from Wirkert et al. [195].

Liver experiment. In a second experiment, images of a porcine liver were recorded directly after lethal drug delivery. Mean oxygenation was determined over all image pixels except for specular regions, which were excluded by thresholding. A steady drop of oxygenation and blood volume fraction was measured as shown in Fig. 4.18, which shows the estimation results for all 51 recorded image stacks.



(a) before clipping (b) setting of the fourth clip (c) nine minutes after applying first clip

Figure 4.17: Visualization of bowel oxygenation estimation. The clipped colon segment was segmented to estimate mean oxygenation. The color bar shows oxygenation in percent. Reprinted with permission from Wirkert et al. [195].



(a) liver before drug delivery (b) mean oxygenation over time (c) mean blood volume fraction over time

Figure 4.18: (a) Qualitative validation of the proposed method in a porcine liver. The plots show the (b) mean oxygenation and (c) blood volume fraction after lethal drug delivery as a function of time. Specular pixels were excluded by simple thresholding. Reprinted with permission from Wirkert et al. [195].

4.4. DISCUSSION

In this chapter the first machine learning based method to estimate physiological parameters from laparoscopic multispectral images was presented. It is based on generating labeled training data using a physical model and highly accurate MC simulations. Random forests are used to invert the MC spectra and derive oxygenation and blood volume fraction for every pixel in the recorded image.

The estimation of these parameters is valuable for monitoring and detecting ischemia. This can be of relevance in operations like partial nephrectomies, colectomies or organ transplants. The method could also be applied in minimally-invasive cancer-screenings such as colonoscopy, where it could help in finding polyps and flat adenoma. To guarantee practical applicability in these scenarios, special emphasis was given to developing a fast, and robust method without compromising on estimation performance.

Unlike the compared state-of-the-art Beer-Lambert method, the presented approach is capable of estimating both oxygenation and blood volume fraction. Due to the underlying MC framework, it is not restricted to assumptions such as constant light penetration depth and scattering losses. In the following the forward model, the proposed inversion

technique and the results are discussed.

Forward model. Oxygenated and de-oxygenated hemoglobin were chosen as sole absorbers in the model because they are the only notable absorbers of visible light in human tissue besides melanin, which, however, is mainly contained in the skin [98] and thus irrelevant for visceral applications. Note that the model allows for easy integration of further absorbers (if necessary) by modification of equation 4.1. Modification of the tissue composition requires investigation as to whether the generated spectra can still be inverted, e.g. by using the regression techniques presented in this paper.

The anisotropy is hard to measure and is not well understood for human tissues, as experiments and the results from theoretical analysis by Mie's scattering theory do not coincide [98]. Most experiments come to the conclusion that human tissue is strongly forward scattering with quite high levels of g in the visible range [98]. Therefore, the anisotropy factor was modelled to be in a range that covers most of the the experimental values depicted in [98].

Inversion by random forest regression. Random forest regressors were chosen because they are capable of near real-time regression of megapixel multispectral images. Additionally random forest regressors are inherently multi-variate and thus allow joint oxygenation and blood volume fraction estimation. However, as can be seen in chapter 7, neural networks are equally suited for the task.

From a machine-learning standpoint, the Gaussian noise w added to the reflectances in Eq. 4.8 necessary to prevent over-fitting of the regressor. The term models noise from the camera, tissue/camera movement during image acquisition and model inaccuracies. A simple noise model was chosen, which adds zero mean Gaussian noise dependent on reflectance and SNR: $w_i \sim \mathcal{N}(0, \frac{r_i}{\mathrm{SNR}})$. In future works more complex models could be explored. One possibility could e.g. be to add more noise to bands which are expected to record darker images. Model inaccuracies can e.g. be the presence of additional unknown absorbers, chromatic aberration, cross talk between pixels caused by inhomogeneous tissue or tissue structures not modelable by a multi-layer model. To account for the latter two, 3D MC simulations would be a viable option for future experiments. Such a set-up would need careful design to both ensure the modeling of realistic tissue and be general enough to cover a relevant tissue variability

To account for constant multiplicative illumination changes, the ℓ^1 norm was applied. Other normalizations proposed to account for these changes are the usage of image quotients [175] or division by the integral of the reflectance spectrum [36]. The $l1$ norm was chosen because it is more robust to noise than image quotients; the integral can be tricky to calculate if the spectral bands are unevenly spaced and sparse.

The additional normalizations of transformation to absorbance and further ℓ^2-normalization could be left out in principle because non-linear regressors were applied. It was found however, that doing these normalizations improves the mean absolute regression

errors by more than 4% in the presented method. The also analyzed support vector regression did not necessitate the ℓ^2 normalization and transformation to absorption. The k-nearest neighbor result dropped by 10% when omitting normalizations.

An additional use-case of the developed method is as an *in silico* testing stage for hardware set-ups. It could, for example, help in choosing the most relevant multispectral bands. Furthermore the framework can be used to compare different inversion techniques. The Python framework was made available on GitHub[4].

Experimental set-up. MC models are considered to be the gold standard for calculating how light travels through biological tissue (Section 3.3.2). For this thesis, the MCML implementation was chosen to create the reflectance spectra. Aside from being more accurate than e.g. the diffusion approximation or the modified Beer-Lambert law it is easy to configure and flexible.

Its main drawback is the time consumption. With the set-up described in Section 4.3, the evaluation of one reflectance spectrum took about 16s by an off-the-shelf desktop PC with an NVIDIA GeForce GTX 660 Ti graphics card. The simulation of 15,000 spectra used for training our regressor therefore took less than three days. Because it is only required for training, this one-time investment of three days' computation was not seen as critical.

Experimental results. Experiments in section 4.3.2 indicated that the simulated spectral reflectances show low variance with repeat simulation and fit well to spectrometer measurements of phantoms with varying oxygenation, blood and scattering. The blood volume fraction for plate one and three estimates a correct approximate mean of 2%. Plate two has a blood volume fraction of about 15% which is not contained in the simulations (which peak at a maximum of 10%), the fitted simulations consequently are close to the maximum of 10%. Somewhat disconcerting, the v_{hb} estimates for plate one vary around the constant, correct value, overestimating v_{hb} for high oxygenation values and underestimating for low oxygenation values. This behavior cannot be observed in plate three and will thus be investigated in further phantom studies. The estimated oxygenation values start from high baselines and reduced when more salt was added, which is in line with expectations. Further phantom studies are currently ongoing in cooperation with Fraunhofer PAMB and will additionally provide quantitative oxygenation values.

The experiments in Section 4.3.3 support the modeling assumptions made in Section 4.1.3. Spatial/spectral independence of the light source is fulfilled. Further the camera response is linear, and camera noise is sufficiently close to normal distribution. Additional experiments investigated the understanding of the optical system and that the measurements by the camera and simulations vary in the same subspace. A finding here is that 95% of the measurements variance is taking place in the first three components of the simulated data, but a much higher portion lies in the first PC compared to the simulated data (89% versus 79%). Probable reason for this is a systematic difference of

[4] https://github.com/swirkert/ipcai2016

recordings and simulations: all oxygenation levels are present in the simulations, while the measurements only capture healthy, and thus well oxygenated, tissue.

The *in silico* study in Section 4.3.4 investigated the performance of the proposed machine learning MC inversion approach. The inversion method outperforms the baseline method for SNRs above ten and is as good as the baseline for lower SNRs. This is true even if the training noise is set to a fixed SNR of ten and the testing noise is varied. A high drop in performance occurred when applying the method to data from a different domain. The next chapter investigates reducing this performance gap by employing the first method to domain adaptation for multispectral functional imaging. Other advanced non-linear methods showed comparable performance but lack the rapid evaluation speed shown by the choice of random forests. The methods inverts 10^6 spectra with more than 5Hz. As the camera records images with 2.5Hz the method is fast enough for real-time processing of these images.

In both *in vivo* experiments, the drops in oxygenation and blood volume fraction have been observed, caused by clamping or lethal injection. The initial value of oxygenation of 70-85% in both experiments is in line with literature [68]. The drop in oxygenation of the bowel is in line with the experiments performed in [30]. The drop in oxygenation of the liver after euthanasia is also expected. The blood volume fraction probably decreases due to loss of blood pressure and subsequent drainage of blood from the liver surface due to gravity. The results on liver tissue are especially encouraging, as the regressor was trained on data tailored to colonic tissue.

The SNRs of the down-sampled camera images was determined to range from 29-50 in the experiments. The recorded images were quite dark with mean values within the lowest 10% of the camera's dynamic range. This can be ameliorated with a different light source. This and the next chapter rely on relatively dark Halogen light. Further, Halogen is brighter in red compared to blue. Thus the darker blue bands exhibit more noise, an effect which is currently not modeled. The system developed in chapter 7 uses a Xenon light source which does not suffer from these problems. Additionally, due to the 400ms long image acquisition, tissue/camera movement introduces noise due to misaligned, sequentially recorded multispectral image stacks. To this end, a snapshot camera is used in Chapter 7 to record images fluently with 25Hz.

4.5. CONCLUSION

The method described here features both the flexibility and realism of Monte-Carlo based approaches coupled with the speeds that are comparable to those of simple online methods. According to extensive *in silico* and *in vivo* experiments, this method is more flexible and accurate than the commonly used Beer-Lambert law-based regression and orders of magnitude faster than regression approaches developed for skin and *ex vivo* multispectral image analyses. Due to its robustness, non-linear estimation capability and rapid execution time the method has a high potential for future application in interventional multispectral imaging.

One remaining issue is the need to know the approximate tissue composition to be able to create forward simulations. The next chapter aims to generalize the approach so that the method can be to almost arbitrary tissues encountered during visceral interventions.

5

Domain adaptation for machine learning based Monte Carlo inversion

Parts of this chapter have been published in the 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) 2017 [196]

THE last chapter presented an approach to estimate oxygenation and blood volume fraction from multispectral images (MSI). Due to the lack of real annotated data, it relies on simulations of tissue reflectances to learn the transformation from reflectance to physiological parameter. The simulations utilize Monte Carlo methods, and are based on a layered model of tissue presented in Chapter 4.1.1. However, the bounds and distribution of the model parameters are hard to determine and highly dependent on experimental conditions [98].

Ideas in the field of domain adaptation (DA) can help overcome this bottleneck. DA is a branch of machine learning which addresses problems caused by mismatching marginal distributions of source and target domains' input data $p(\mathbf{x}_{\text{source}}) \neq p(\mathbf{x}_{\text{target}})$ [147]. This correspond to simulated source data and real target data in this thesis. The approaches to DA can be grouped into: (1) instance-transfer methods [75, 176, 177], which use the source instances for training, but assign a weight to each instance. These weights are high if the probability of drawing the sample in the source domain is small compared to the probability of drawing it in the target domain and vice versa. Instance-transfer methods are applicable if samples from the target domain have a non-zero probability of occurring within the source domain. If this is not the case (2) feature-representation transfer methods can be used. These either find a transformation to a feature representation invariant to the domain shift [53, 60, 73, 178]; or transform the features of the source domain to resemble features from the target domain [22].

This chapter describes the first use of DA methods to tackle the problems of unknown specifics for model parameters. This instance-transfer method is described in Section 5.1. The following study in Section 5.2 with seven pigs shows that (1) the model captures a large amount of the variation in real tissue and (2) the DA based approach enables highly accurate functional parameter estimation. The chapter closes with a discussion (Section 5.3) and a brief conclusion (Section 5.4).

5.1. METHODS

The functional parameter estimation method presented in Chapter 4 is based on accurate knowledge of the biophysical properties captured in the layered tissue model. However, depending on the experimental conditions, the values of these parameters can have significant variations. For example, blood volume fraction measurements in the colon from Zonios et al. [207] and Skinner and O'Brien [169] differ by 1-2 orders of magnitude [91]. This motivates the need for a more general framework that is less reliant on tissue specific parameter measurements.

The developed approach, illustrated in Fig. 5.1, aims to compensate for the lack of detailed prior knowledge related to optical tissue properties by applying DA. The hypotheses are that (1) a highly generic tissue model can be used to obtain spectral reflectances that cover a large range of multispectral measurements observable *in vivo* and that (2) samples of real (unlabelled) measurements can be used to adapt the data to a new target domain. Section 5.1.1 describes how the generic data set is generated while Section 5.1.2. introduces the approach to DA.

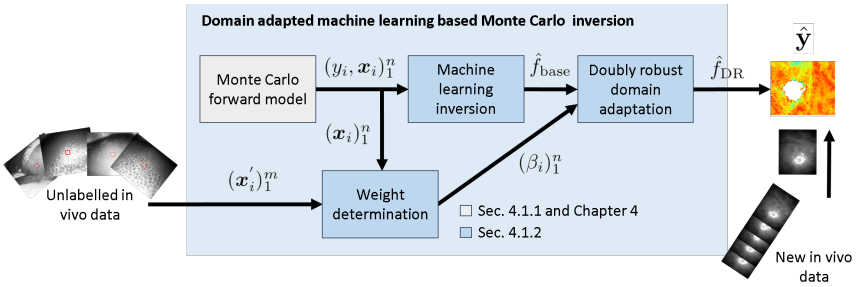


Figure 5.1: Overview on the approach. First a large number of generic, labelled, tissue-reflectance samples are simulated. The desired physiological parameter y_i and the reflectances adapted to the imaging system $\mathbf{x}_i$ are used for training a base regressor $\hat{f}_{\text{base}}$ with the methods described in the last chapter. Weights $(\beta_i)_1^n$ are calculated to fit the simulated data to measurements. These weights adapt the base regressor to the *in vivo* measurements. Applying the adapted regressor $\hat{f}_{\text{DR}}$ to new images yields their functional parameter estimates $\hat{\mathbf{y}}$.

5.1.1. DATA SET GENERATION USING A GENERIC TISSUE MODEL

The method relies on a single comprehensive data set, which can be used with various camera setups and is representative of a large variety of target structures in the human body. This data set captures a wider variety of tissue parameters compared to the data set tailored to colonic tissue from Chapter 4. By increasing the range, the density of samples naturally decreases. To account for this, the data set size was increased by a factor of 50.

The layered tissue model developed in Chapter 4.1.2 was used to create the data set consisting of camera independent tissue-reflectance pairs $(\mathbf{t}_i, \mathbf{r}_i(\lambda, \mathbf{t}_i)), i \in \{1...n\}$. The parameters were varied within the ranges shown in Table 5.1 for each of the three layers.

As before each layer is described by its value for blood volume fraction v_{hb}, scattering coefficient a_{mie}, scattering power b_{mie}, anisotropy g, refractive index n and layer thickness d. Oxygenation s is again kept constant across layers. Following the values in [98] and in contrast to Chapter 4, b_{mie} is varied, covering all soft, fatty and fibrous tissues. Ranges of v_{hb} were increased by a factor of three to potentially model pathologies. Together with values for hemoglobin extinction coefficients ϵ_{Hb} and ϵ_{HbO2} from the literature, absorption and scattering coefficients μ_a and μ_s are determined for usage in the Monte Carlo simulation framework. The simulated range of wavelengths λ is large enough for adapting the simulations to cameras operating in the visible and near infrared. To account for a specific camera setup, spectral reflectance $\mathbf{r}_i$ is transformed to normalized reflectances $x_{i,j}(\mathbf{t})$ at the jth spectral band, using Equations 4.6 and 4.7 from Chapter 4.1.3. In this chapter we skip the additional normalization to absorption (Equation 4.8) to limit the number of design choices.

5.1.2. DOMAIN ADAPTATION

The combination of normalized reflectances and corresponding physiological parameter $(\mathbf{X}, \mathbf{y})$ serve as training data for any machine learning regression method f to obtain

Table 5.1: The simulated ranges of physiological parameters, and their usage in the simulation set-up as described in Section 5.1.1. Important changes compared to the last chapter are marked in bold font.

	v_{hb}[%]	s[%]	$a_{\text{mie}}[\frac{1}{\text{cm}}]$	b_{mie}	g	n	d[mm]
layer 1-3:	**0-30**	0-100	5-50	**.3-3**	.8-.95	1.33-1.54	**.02-2**

$\mu_a(v_{\text{hb}}, s, \lambda) = v_{\text{hb}}(s\epsilon_{\text{HbO2}}(\lambda) + (1-s)\epsilon_{\text{Hb}}(\lambda))\ln(10)150\text{gL}^{-1}(64,500\text{gmol}^{-1})^{-1}$
$\mu_s(a_{\text{mie}}, b, \lambda) = \frac{a_{\text{mie}}}{1-g}(\frac{\lambda}{500\text{nm}})^{-b_{\text{mie}}}$

simulation framework: GPU-MCML[7], 10^6 photons per simulation
simulated samples: **500K**
sample wavelength range: **300-1000nm**, stepsize 2nm

the regressor $\hat{f}_{\text{base}}$, which corresponds to a regressor without DA. The functional parameter estimates during an intervention can be determined using this baseline regressor by evaluating $\hat{f}_{\text{base}}(\mathbf{x}') = y'$ for each recorded multispectral image pixel $\mathbf{x}'$.

5

Working with tissue samples from a simulated source domain $p_s(\mathbf{x}, y)$ will inevitably introduce a bias with respect to the target domain $p_t(\mathbf{x}, y)$. Covariate shift occurs if $p_s(y|\mathbf{x}) = p_t(y|\mathbf{x})$ and therefore $\frac{p_t(\mathbf{x},y)}{p_s(\mathbf{x},y)} = \frac{p_t(y|\mathbf{x})p_t(\mathbf{x})}{p_s(y|\mathbf{x})p_s(\mathbf{x})} = \frac{p_t(\mathbf{x})}{p_s(\mathbf{x})} = \beta(\mathbf{x})$. If p_t is contained in the support of p_s, weighing the sample loss by β in the empirical risk (Equation 4.9) can correct for the covariate shift [95].

The appeal of this method is, that only recordings $\mathbf{x}'_i$ and no labels y'_i are necessary for adaptation. While the concept of covariate shift has been applied with great success in a number of different medical imaging applications [72, 87], challenges related to transferring it to this problem are estimation of high dimensional $\boldsymbol{\beta}$, and high variance introduced by weighting, both addressed in the next two subsections.

FINDING β WITH KERNEL MEAN MATCHING AND RANDOM KITCHEN SINKS

Kernel mean matching (KMM) is a state-of-the-art method for determining $\boldsymbol{\beta}$ [75, 95]. KMM minimizes the mean distance of the samples of the two domains in a reproducing kernel hilbert space (RKHS) $\mathscr{H}$, using a possibly infinite dimensional lifting $\phi : \mathbb{R}^m \rightarrow \mathscr{H}$. In its original formulation [95], the kernel trick $k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle$ was used to pose KMM as a quadratic problem. In this problem domain it is not feasible because calculating the Gram matrix is quadratic in the (high) number of samples. To overcome this bottleneck, the KMM objective function [95]

$$\arg\min_{\boldsymbol{\beta}} \| \frac{1}{n}\sum_{i=1}^{n} \beta_i \phi(\mathbf{x}_i) - \frac{1}{n'}\sum_{i=1}^{n'} \phi(\mathbf{x}'_i) \|^2, \text{ s.t. } \beta_i \in [0, B] \text{ and } |\sum_{i=1}^{n} \beta_i - n| \leq n\frac{B}{\sqrt{n}} \tag{5.1}$$

is minimized directly. This objective function matches the empirical means of simulations and real data in $\mathscr{H}$. The first boundary condition limits the maximum influence of individual training samples, the second condition ensures $\beta_i\, p(\mathbf{x})$ is close to a probability distribution [75]. The variables n and n' denote the number of samples in the source

and target domain. Equation 5.1 cannot be minimized directly due to the possible infinite dimension of ϕ. The random kitchen sinks method [155] finds an approximate representation of $\phi(\mathbf{x}_i) \approx z(\mathbf{x}_i)$ by sampling from the Fourier transformation of a shift invariant kernel. This enables solving the convex KMM objective function in its non-kernelized form using a standard optimizer.

DOUBLY ROBUST COVARIATE SHIFT CORRECTION

Estimators trained using weighted samples can yield worse result than estimators not accounting for the covariate shift. The reason is that only few samples are effectively "active", leaving less samples to train the algorithm. However, using the unweighted training samples often leads to a reasonable, but biased, estimator [159]. Intuitively the unweighted base regressor $\hat{f}_{\text{base}}$ defined in Section 5.1.1 can be used to obtain an initial estimate. Subsequently, another estimator aims to refine the results with emphasis on the samples with high weight. This is the basic idea of doubly robust (DR) covariate shift correction [159]. Specifically, residuals $\delta_i = y_i - \hat{f}_{\text{base}}(\mathbf{x}_i)$ weighted by $\boldsymbol{\beta}$ from the Section 5.1.2 train an estimator $\hat{f}_{\text{res}}$ on $(\mathbf{X}, \boldsymbol{\delta}, \boldsymbol{\beta})$. The final estimate is $\hat{f}_{\text{dr}}(\mathbf{x}') = \hat{f}_{\text{base}}(\mathbf{x}') + \hat{f}_{\text{res}}(\mathbf{x}')$.

5.2. EXPERIMENTS AND RESULTS

The quality of the generic tissue model (Chapter 5.2.2) as well as the performance of the DA-based physiological parameter estimation approach (Chapter 5.2.3) was validated on a comprehensive *in silico* and *in vivo* MSI data set (Chapter 5.2.1).

5.2.1. EXPERIMENTAL SETUP

The multispectral images were recorded with the multispectral laparoscope described in Chapter 4.3.1. The images were taken from seven pigs and six organs (liver, spleen, gallbladder, bowel, diaphragm and abdominal wall) in a laparoscopic setting. For all experiments the data set described in Chapter 5.1.1 was used for training. The standard deviation σ_w of noise term w was set using $\sigma_w = \frac{x_j^o}{SNR}$, where x_j^o is the normalized reflectance without added noise. If not mentioned otherwise, the signal-to-noise ratio (SNR) was set to twenty. Random forest regression was used as inversion method with parameters as in Chapter 4.3.4. 1000 random directions were drawn from the RKHS induced by the radial basis function (RBF) kernel with the random kitchen sink method. The σ value of the RBF was set to the approximate median sample distance and the B parameter of the KMM to ten. The L-BFGS-B minimizer [25] (scipy v0.19) was used to minimize the KMM objective function.

5.2.2. VALIDITY OF TISSUE MODEL

One of the prerequisites for covariate shift correction is that the support of the distribution of *in vivo* measurements is contained in the support of the simulated reflectances.

To investigate this for a range of different tissue types (cf. Chapter 5.2.1) a total of 57 images was collected, measurements from a 100x100 region of interest (ROI) were extracted and corrected by a flatfield and dark image. The first three principal components explain 98% = 81%+13%+4% of the simulated variance. For *in vivo* data, 97% = 89%+4%+4% of the variance lies on the first three principal components of the simulated data.

For a qualitative assessment the *in vivo* measurements were projected on the first two principal components of the simulated data. A selection can be seen in Fig. 5.2. Except for gallbladder, all of the *in vivo* data lie on the two dimensional manifold implied by the simulated data. Fig. 5.3 illustrates how changes in oxygenation and perfusion influence the distribution of the measurements.

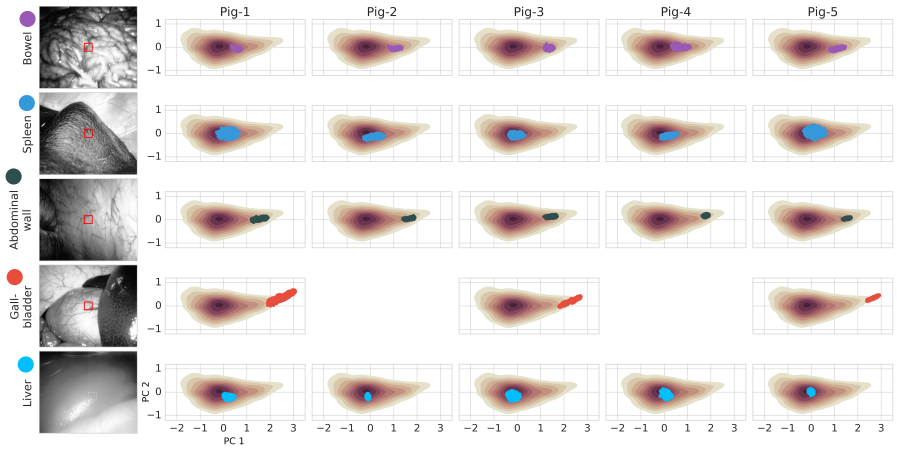


Figure 5.2: Five organs from five pigs projected onto the first two principal components of the simulated reflectance data plotted in brown. For some pigs, no Gallbladder measurements were performed. The images on the left show the 560nm band recorded for the first pig. The depicted measurements are taken from the red ROI. Except for gallbladder, organs lie on the non-zero density estimates of the simulated data (See Sect. 5.3).

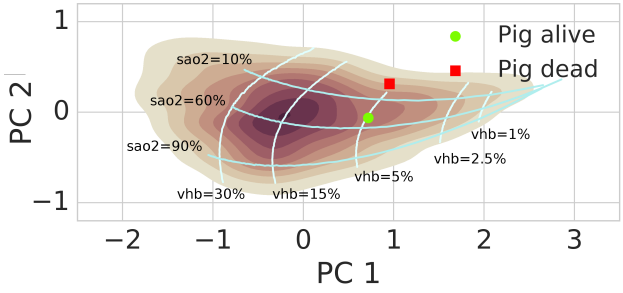


Figure 5.3: Liver tissue measurements before and after sacrificing a pig. The grid indicates how varying oxygenation (sao2) and blood volume fraction (vhb) changes the measurements in the space spanned by the first two principal components of the simulations. Note that these lines can not be directly interpreted as sao2 and vhb values for the two points, because other factors such as scattering will cause movement on this simplified manifold.

5.2.3. PERFORMANCE OF DOMAIN ADAPTATION

The domain adaptation was investigated under two points of view: (1) The *in silico* experiments focus on the benefit of DR DA and DA for oxygenation estimation. Due to the lack of ground truth data, these experiments were performed in silico. (2) The *in vivo* experiments investigated if the weighting method selects samples such that the distribution of simulations and measurements match more closely.

***In silico* experiments.** To validate the DA approach to physiological parameter estimation with reliable reference data, *in silico* experiments were performed with simulated colon tissue as target domain. For this purpose, 20,000 colon tissue samples with corresponding ground truth oxygenation from the last chapter at an SNR of 20 were used. 15,000 reflectance samples were selected for DA while the remaining 5,000 samples were used for testing. All experiments were run five times to reduce random influences.

A first experiment investigated the general effect of DA and DR DA compared to no DA. Additionally, the number of training samples was varied between 10^4 and $5 \cdot 10^5$ to investigate how the data set size influences results. The DR DA method reduced the median absolute oxygenation estimation error compared to the base estimator by 25-26% and by 9-12% compared to a version without the DR correction (Fig. 5.4). Training DR DA with 500,000 instead of 10,000 samples reduced the error by 7.1%.

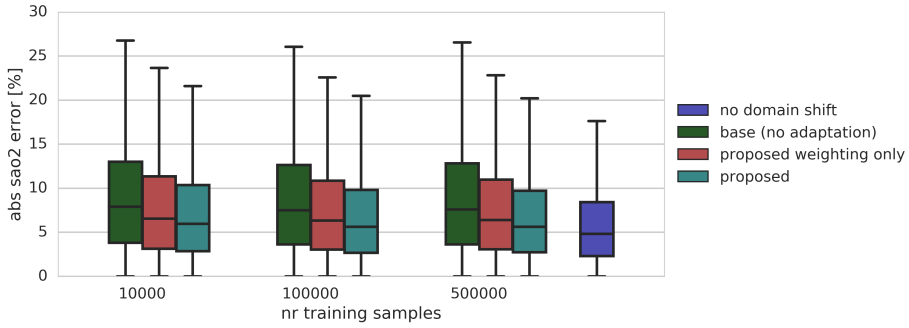


Figure 5.4: The generic data set was adapted with the proposed DR DA method to the *in silico* colon data set from the last chapter. Depicted is a boxplot of the absolute oxygenation estimation error for no adaptation, the proposed method without the DR correction and the proposed DR DA method. As a reference the results from training and testing on the colon data set are shown. This would be the best possible result. Boxplot, whiskers extend the low/high quartiles by 1.5 IQR, extreme values not shown.

Several weight determination algorithms have been proposed in literature [72, 87, 176]. The influence of the weight determination method has been investigated by comparing the chosen KMM method to the logistic regression classifier weight determination implementation from [72]. Figure 5.5 shows the oxygenation error results with and without the DR correction. The percentual median absolute oxygenation error for the KMM method is slightly higher (0.2 percentage points) without the DR correction and the same with the DR correction.

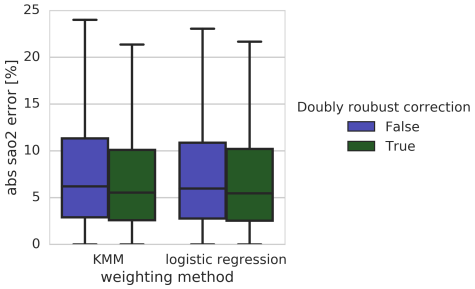


Figure 5.5: The generic data set was adapted with the proposed doubly robust domain adaptation (DR DA) method to the *in silico* colon data set from the last chapter. Depicted is the absolute oxygenation estimation error for two different weight determination algorithms. Both algorithms lead to similar results after applying the DR DA. Boxplot, whiskers extend the low/high quartiles by 1.5 IQR, extreme values not shown.

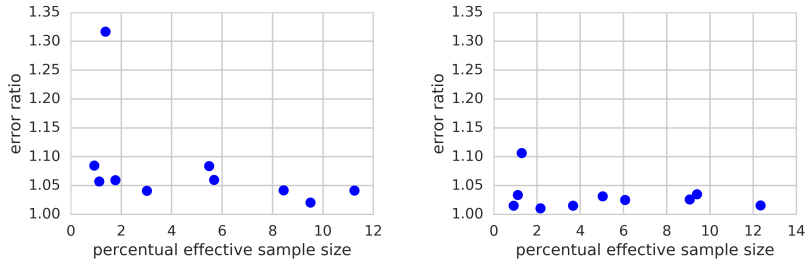
The effect of DR estimation should be greater for lower effective sample sizes, which can be approximated by $m_{\text{eff}} \approx \frac{\|\boldsymbol{\beta}\|_1^2}{\|\boldsymbol{\beta}\|_2^2}$ [111]. To investigate the effect of DR correction as function of effective sample size a second validation scheme was adopted from [159]. The complete data set was split in three disjoint subsets:

1. The first data set was chosen to be the source data set.
2. The second data set was resampled to form an artificial target distribution and used for weight determination.
3. The third data set was resampled to the same artificial target distribution and used for testing the regressor, trained on the weighted first data set.

To create an artificial target distribution the second and third data set were resampled by the following procedure: On the second data set, minimum $t_{\min}$, maximum $t_{\max}$ and the standard deviation σ_{pc} of the first principle component was determined. Bot target distributions were created from set two and three by sampling with replacement with $\mathcal{N}(t_{\min} + \alpha(t_{\max} - t_{\min}), 0.1\sigma_{\text{pc}})$ from the subsampling distribution. The α values controlled the location of the target distribution and as a consequence also varied m_{eff}.

Parameter α was varied from 0 to 1 in 10 equidistant increments. The B parameter was set to 100 to allow effective sample sizes smaller than 10%. The the root mean squared error for each α was determined for DR and DR DA. The ratio of DR DA to DA varies with the number of source training samples and effective sample size, ranging from no benefit to over 30% improvement. The results are summarized in Fig. 5.6.

***In vivo* experiments.** Because no oxygenation ground truth was available for the *in vivo* measurements, a different metric had to be employed. Thus the distance of the distributions of measurements to simulations weighted by the measurements was investigated. As notion of closeness served the Euclidean distance of the weighted simulation mean $\frac{1}{n}\sum_i^n \beta_i \mathbf{x}_i$ to the mean of the images. This distance was calculated for each organ.



(a) 10,000 training samples (b) 100,000 training samples

Figure 5.6: Investigation of doubly robust domain adaptation (DR DA) compared to DA without DR correction as a function of effective sample size. The y-axis show the quotient of domain adaptation and doubly robust domain adaptation RMSE.

The weighted distance was 34-91% (median 78%) smaller than the unweighted average. See Fig. 5.7 for a depiction in the principal component space.

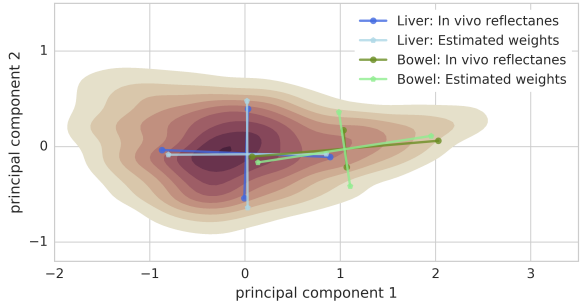


Figure 5.7: The distribution of *in vivo* measurements and adapted *in silico* reflectances in the principal component space of the simulations. For graphical clarity the distributions are visualized as their two principal axes in this space with lengths scaled by the square root of the corresponding eigenvalues.

Further the generalizability of the method to new subjects (here:pigs) was investigated. More specifically, simulations weighted by a set of subjects were compared to data from previously unseen subjects. The analysis was performed on bowel data from six pigs (C_{total}). All combinations of three pigs, C_{test}, were selected in a leave-three-out manner. The images from these pigs were used as test cases. From the corresponding remaining three pigs of each selection a second selection was made ($C_{\text{DA-i}} = \binom{3}{i}$ for i = [1, 2, 3]), and each $C_{\text{DA-i}}$ was used independently to determine the weights to be applied to the generalized simulation data, using the domain adaptation approach. This allowed investigation of the effect of additional *in vivo* samples, thus the generalization capability of the approach. To evaluate, for each C_{test}, the Euclidean distance was determined between the centroid of the reflectances of the test pigs from: (a) the weigthed centroid of the simulated reflectances (shown in blue in Figure 5.8) and, (b) the unweighted centroid

of the simulated reflectance; as a function of the number of pigs used in $C_{\text{DA-i}}$. Moving from one to two training pigs, the weighted sample distance decreases. Adding a third pig for training leads to similar median values, but smaller 75 percentile values.

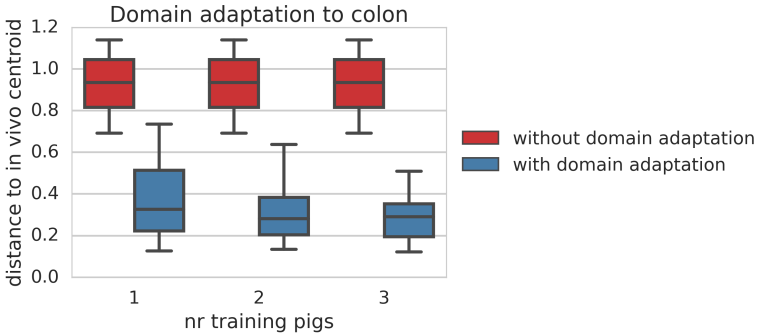


Figure 5.8: Investigation of effect of varying number of pigs used for domain adaptation. Using more pigs for training reduces median and/or 75 percentile distance. Boxplot, whiskers extend the low/high quartiles by 1.5 IQR, extreme values not shown.

5.3. DISCUSSION

This chapter introduced the first domain adaptation approach to functional parameter estimation from multispectral imaging data. It (a) does not rely on labelled images, (b) does not require specific prior knowledge of optical tissue properties and (c) is independent of camera model and corresponding optics. Thus it is broadly applicable to a wide range of clinical applications. Instabilities potentially caused by DA were addressed by a large training data set and incorporation of a DR correction method.

Data set. The method is built around a generic data set that can automatically be adapted to a given target anatomy based on samples of unlabelled *in vivo* data. Because it is capturing tissue in larger ranges than the model used in Chapter 4.1.1, the size was increased, forming the to my knowledge largest data set of MC simulated spectral tissue reflectances. Creating spectral reflectances at 2nm steps allows simulation of virtually any real multispectral camera by applying the transformations from Chapter 4.1.3.

According to porcine experiments with six different target structures, the first three principal components of the simulated data set capture 97% of measured *in vivo* variations. These three axes probably mainly represent the blood volume fraction, oxygenation and scattering. Visual inspection of the first two principal components showed the captured organ data lie within the simulated data, an important prerequisite for the subsequent DA to work. Gallbladder is the exception, most likely due to its distinctive green stain, caused by the bile shining through. Modelling the bile as another chromophore and extending the data set accordingly would be straightforward.

Domain adaptation. Instance-transfer DA was performed, weighting simulations to match the measurements using the KMM objective function (Equation 5.1). An important methodological component in this context was the integration of the recently proposed DR correction method to address the instabilities when few effective training samples are selected.

The *in silico* experiments demonstrate the potential performance boost when adapting the generalized model to a specific task using the presented DA technique. This performance boost was observed, even if the weight determination method was exchanged (Figure 5.5). As expected, the DR correction has the most significant impact if the number of training samples is low (Figure 5.6). Incorporating the DR correction almost always led to better results. These quantitative benefits in performance are especially encouraging, since they show benefit although $p_s(\mathbf{y}|\mathbf{x}) = p_t(\mathbf{y}|\mathbf{x})$ was not explicitly verified. This condition would be violated if several simulated functional parameter settings lead to similar measurements, but only some of these are present in real tissue. Gretton et al. [75] noted, that the method usually performs well even if the condition is not strictly fulfilled. To introduce as little arbitrary choices as possible, evaluation reflectances were normalized with Equation 4.6, skipping the additional, strictly not necessary normalization (Equation 4.8). Influence of differing kinds on normalization and their interplay with the domain adaptation remains an open research topic.

The *in vivo* experiments showed that weighted simulations resemble real measurements more closely their unweighted counterparts. This is quantified by the distance of the weighted and unweighted simulation mean to the measurement mean. The distance reduced for all tested organs, supporting the hypothesis that the method can be used to "tailor" the generic simulations to specific tasks. Further, both bias and variance of these means reduce when increasing the number of porcines used for weight determination from one to two. Adding a third pig further decreased the variance. This result requires some caution in interpretation. While the weighted distance of the centroids reduces, this does not necessarily imply better performance of the algorithm. Using only well oxygenated *in vivo* colon samples might e.g. lead to reduced weighted distances. However, focusing on oxygenated samples will likely lead to a regressor which cannot create correct estimates for deoxygenated tissue.

5.4. CONCLUSION

This chapter addressed the important bottleneck of lack of annotated MSI data with a novel domain adaptation-based method to functional parameter estimation. This allows use of highly accurate Monte Carlo based approaches in problem domains where specific knowledge about tissue is scarce.

A remaining issue to be addressed is the problem of 2.5Hz recording and analysis. While relatively fast for a multispectral system, it is far from the 30-60Hz of conventional RGB systems. To this end, the next chapter will focus on selecting a small subset of multispectral bands to speed up recording. Chapter 7 presents the first multispectral laparoscope able to both record and analyze multispectral images at video-rates.

6

MULTISPECTRAL BAND SELECTION

Parts of this chapter have been published in Computer-Assisted and Robotic Endoscopy **8899** (2014) [194].

WHILE the previous chapters focused on multispectral image analysis, this chapter investigates the question of which bands should be recorded. Band selection is related to two major challenges in multispectral imaging. Sequentially recording staring techniques suffer from long image recording times, often in the range of seconds (Chapter 3). Due to tissue and instrument movement, long acquisition times lead to misalignment of the recorded stack, thus rendering the measured spectra invalid. Reducing the number of recorded bands leads to faster image acquisition and thus less misalignment. The competing snapshot techniques record an image stack all at once, but spatial resolution decreases with increasing number of bands.

The central guiding hypothesis of this chapter is that a small but principled selection of bands can retain most of the information necessary for the subsequent analysis. Prior to the presented method [194], biomedical band selection approaches relied either on the availability of labelled training data or sophisticated knowledge of the tissue (Chapter 3.2.2). As discussed in the previous chapters, both these requirements are hard or even impossible to meet.

The next section of this chapter presents the Endoscopic Sheffield Index (ESI), a novel method for band selection in endoscopic environments. The experiments in Section 6.2 evaluate the approach both on phantoms and *in vivo* data from five pigs. The results are discussed in Section 6.3 and a conclusion and outlook is given in Section 6.4.

6.1. METHODS

The featured approach selects bands for surgical applications, requiring neither labelled training data nor a predefined light-tissue interaction model. The principle idea is to once carefully record a large number of bands and use them to select the subset of the recorded bands $\mathbf{b} \in B$ (see Fig. 6.1), based on the developed information-theoretic approach, which adapts a method known in the remote sensing community as the Sheffield index (SI) to work in a laparoscopic environment. Before the SI can be applied, several normalization steps remove noise sources which would introduce misinformation. The whole process is shown in Figure 6.1.

The multi-spectral imaging system used to acquire these high dimensional image stacks is described in the following Section 6.1.1. Section 6.1.2 explains how to model the measured image intensity as a function of illuminant, imaging system and underlying tissue. Section 6.1.3 presents the ESI.

6.1.1. MULTISPECTRAL IMAGING SYSTEM

A custom designed hardware set-up was developed by partners at Imperial College [33], depicted in Figure 6.2 (a). It combines a Karl Storz (Knittlingen, Germany) laparoscope and light source with a Varispec™(PerkinElmer, Waltham, MA, USA) VIS Liquid Crystal Tunable Filter (LCTF) and a Thorlabs (Newton, NJ, USA) DCU 223M monochrome camera to acquire the spectral images. The LCTF can be tuned to record bands of varying width in a spectral range of 400-720nm. The switching time between bands is 50ms.

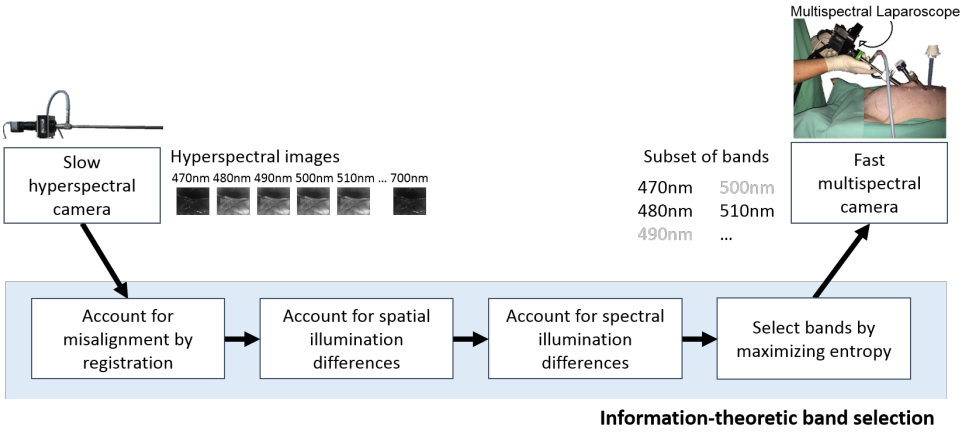


Figure 6.1: Band selection can aid in designing new cameras by selecting the most useful bands. For this purpose, some hyperspectral images are taken. After several normalizations are performed to account for uninformative noise sources, unsupervised band selection chooses bands which maximize differential entropy.

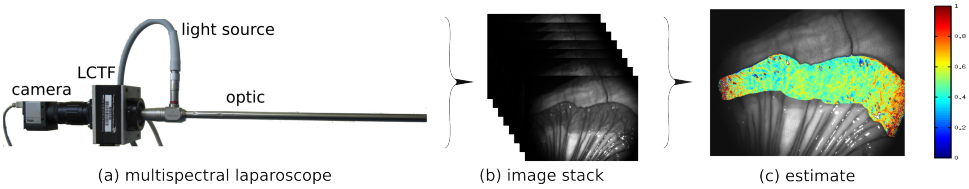


Figure 6.2: The hardware set-up, adapted from [33] and depicted in (a) used to collect spectral image stacks (b). The image stacks are used to estimate underlying tissue parameters such as hemoglobin oxygenation, visualized in (c). Red pixels represent high hemoglobin oxygenation. Reprinted with permission from Wirkert et al. [194].

6.1.2. FORMAL ENDOSCOPIC IMAGE ACQUISITION MODEL

The reasoning behind the band selection algorithm is based on the image acquisition model from chapter 4.1.3, briefly revisited here:

$$i_j(\mathbf{t}) = a(\mathbf{p}) \int_{\lambda_{\min}}^{\lambda_{\max}} l_j(\lambda) \cdot r(\lambda, \mathbf{t}) \cdot w_j \, \mathrm{d}\lambda. \tag{6.1}$$

where

- $a(\mathbf{p})$ subsumes the spatially dependent part of the illumination and imaging system specific constants as a scaling factor to account for the camera's detector range
- $l_j(\cdot)$ subsumes the wavelength dependent part of the illumination and wavelength specific parts of the imaging system, namely quantum efficiency of the camera, transmission of the lapararoscope, the jth filter response, acquisition time and gain
- $r(\lambda, \mathbf{t})$ is the light-tissue interaction model, relating tissue $\mathbf{t}$ to its reflectance.

$\mathbf{t}(\mathbf{p})$ is a parameter vector representing the composition of the tissue at $\mathbf{p}$, like e.g. hemoglobin concentration, diameters and refraction index of scattering parameters. With slight abuse of notation written as $\mathbf{t}$ in the following

w_j A noise term. Subsumes camera noise and noise due to image misalignment.

The integration over the spectrum accounts for the extension of the recorded band.

6.1.3. ENDOSCOPIC SHEFFIELD INDEX

In endoscopic applications spectral imaging is impeded by two main factors: firstly the stack of spectral images acquired is misaligned owing to the motion of the endoscope and the tissue. Secondly illumination in endoscopic procedures is highly inhomogeneous due to the radial attenuation of light, the varying organ morphology and properties of the imaging system. To compensate for motion, the spectral images are aligned with a feature driven approach proposed in [174]. Camera noise can be removed by applying Gaussian smoothing [30] or anisotropic diffusion filtering [133]. To eliminate the effects of the inhomogeneous lighting and the 3-dimensional structure of tissue $a(\mathbf{p})$ image quotients are used [175]. This is also known as discounting the illuminant and means all band intensities i_j are divided by one selected band's intensity $i_k \in B$, which was subsequently excluded from analysis $B' = B \backslash \{k\}$:

$$\frac{i_j(\mathbf{t})}{i_k(\mathbf{t})} = \frac{\int_{\lambda_{min}}^{\lambda_{max}} l_j(\lambda) r(\lambda, \mathbf{t}) d\lambda}{\int_{\lambda_{min}}^{\lambda_{max}} l_k(\lambda) r(\lambda, \mathbf{t}) d\lambda} \quad . \tag{6.2}$$

Subsequently, the image is divided by a reflectance standard with constant reflectance $1_j \propto \int_{\lambda_{min}}^{\lambda_{max}} l_j(\lambda) d\lambda$, measured at a central location. Doing this normalizes the integral w.r.t. the volume of $l_j(\lambda)$. This ensures that wavelength dependent lighting intensities, filter/camera quantum efficiencies and filter volume do not influence the band selection result:

$$\frac{i_j(\mathbf{t})}{i_k(\mathbf{t}) 1_j} \propto \frac{\int_{\lambda_{min}}^{\lambda_{max}} \frac{l_j(\lambda)}{\int_{\lambda_{min}}^{\lambda_{max}} l_j(\lambda) d\lambda} r(\lambda, \mathbf{t}) d\lambda}{\int_{\lambda_{min}}^{\lambda_{max}} l_k(\lambda) r(\lambda, \mathbf{t}) d\lambda} \quad . \tag{6.3}$$

As a final step we apply the logarithm to both sides:

$$m_j(\mathbf{t}) := \log \frac{i_j(\mathbf{t})}{i_k(\mathbf{t}) 1_j} \quad . \tag{6.4}$$

This can be motivated by the exponential relation of chromophore absorption and measured intensity, which is the essence of Beer-Lambert's law (See Chapter 3.3.2) and by experimental analysis of the distribution of intensity values which were negatively skewed.

Now that the data is prepared to be as independent to tissue movement and lighting as possible, the subset of bands $\hat{\mathbf{b}} \subset B'; |\hat{\mathbf{b}}| = p; p < |B'|$ can be chosen in a way to retain maximum of information on $m_j(\mathbf{t})$. This can be done by finding the $b \subset B'; |\mathbf{b}| = p$ with

the maximal determinant for the covariance matrix (sometimes referred to as Sheffield Index):

$$\mathrm{SI} = \det(\mathrm{Cov}(m_{\mathbf{b}}(\mathbf{t}))). \tag{6.5}$$

where $m_{\mathbf{b}}(\mathbf{t})$ is the random vector formed by concatenation of $m_j(\mathbf{t})$, $j \in \mathbf{b}$. This is equal to finding the bands with the maximum entropy assuming $m_{B'}(\mathbf{t})$ has an underlying multivariate normal distribution. For a complete derivation please refer to [167].

6.2. EXPERIMENTS AND RESULTS

The experimental evaluation is performed for phantom data and *in vivo* surgical images. As reflectance standard a Labsphere (North Sutton, NH, USA) Spectralon reflectance standard with near constant reflectance in the visible and near infra-red spectrum was chosen. For evaluation purposes, the modified Beer-Lambert's (chapter 3.3.2) law was used to determine relative chromophore concentration.

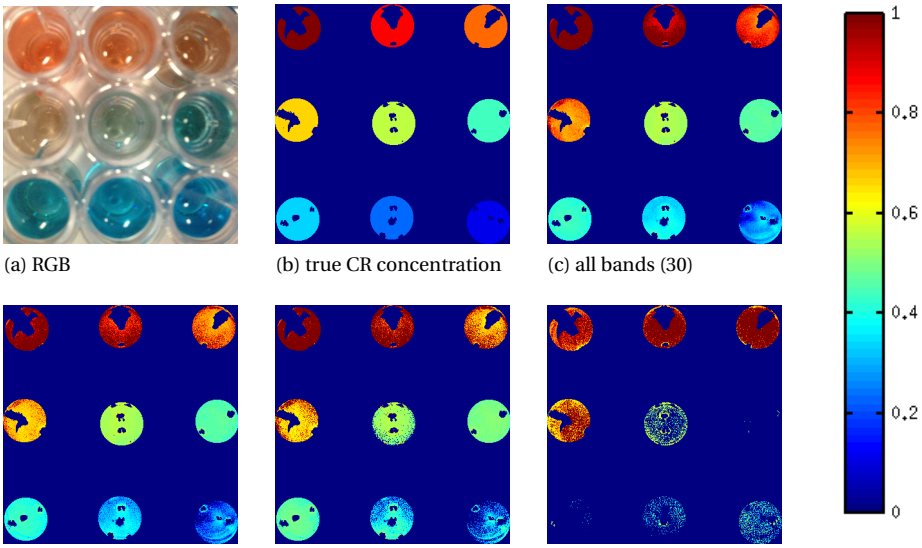
6.2.1. PHANTOM EXPERIMENTS

Combinations of oxygenated and de-oxygenated hemoglobin are often the only notable absorbers in human tissue. The fraction of oxygenated hemoglobin to total hemoglobin concentration is referred to as *oxygenation* and is highly relevant for many diagnoses as mentioned in chapter 3.1. The phantom experiments investigate how the number of selected bands influences the systems capability to estimate oxygenation.

Because oxygenation is hard to reproduce artificially it was emulated by mixing two chromophores, methylene blue (MB) and Congo red (CR). Five 3×3 grids of wells were built (see Fig. 6.3 (a)), containing mixtures starting with 90% CR and 10% MB descending in 10% steps to the final 10% CR, 90% MB well. To maximize the evaluated combinations of CR and MB concentrations, the two chromophores were acquired at molar concentrations of 20µM, 30µM and 50µM. Five grids were constructed according to the above scheme for the combinations (20µM CR, 50µM MB), (30µM CR, 50µM MB), (30µM CR, 30µM MB), (50µM CR, 30µM MB) and (50µM CR, 20µM MB).

With the set-up described in Section 6.1.1 31 spectra were recorded, ranging from 500nm to 650nm in 5nm steps. The 540nm band was used to discount the illuminant (see Section 6.1.3). Relative CR concentration was estimated using the modified Beer-Lambert. Evaluations were performed with leave one grid of wells out cross-validation.

The relative CR concentration estimation as a function of selected bands was determined. Figure 6.4 shows that seven bands achieve estimation results similar to 30 bands, using RMSE as metric. A detailed look at the estimation result of one exemplary grid of wells is available in Fig. 6.3. It can be seen that the results for three bands are notably less robust than the results for seven or 30 bands. The chosen baseline method was to select the bands with the maximum variance [79]. The result is significantly worse than the proposed approach, showing the importance of accounting for inter-band covariances.



(a) RGB (b) true CR concentration (c) all bands (30)

(e) our method seven bands (f) our method three bands (g) seven max variance bands

Figure 6.3: Comparison of the (30µM MB, 30µM CR) grid of wells estimation results with different numbers of bands. The missing pixels in the wells are saturated pixels and were excluded from evaluation. Red means high Congo red (CR) concentration. Reprinted with permission from Wirkert et al. [194].

6.2.2. *In vivo* EXPERIMENTS

Eventually the purpose of this work was to find the best filter for the fast filter wheel based multispectral laparoscope (Section 4.3.1) with eight filter slots. Because the fast filter wheel technique allows shorter integration and acquisition times the effect of motion is greatly reduced and near-real-time operation is possible. One filter is needed to discount the illuminant, leaving seven slots. The *in vivo* experiments thus investigate if the band selection algorithm is able robustly determine the seven most relevant bands by analyzing several image stacks taken from different animals. Furthermore oxygenation estimation accuracy is evaluated similar to the phantom experiments and compared with several baseline methods.

For five different pigs, three *in vivo* spectral image stacks each of the serosal porcine small bowel surface were analyzed. Twenty-one bands were recorded from 470 to 670nm in ten nm steps, with a total acquisition time of 10.5s per image stack. The images were aligned using the algorithm described in [174]. The 580nm band was chosen to discount the illuminant as this band shows very similar extinction coefficients for oxy- and deoxyhemoglobin and thus is not likely to be chosen as a relevant band. Band selection was performed doing leave one porcine out cross validation in order to not train on image stacks coming from the test animal.

The optimal band combination determined was (470 480 490 520 550 590 670), with one

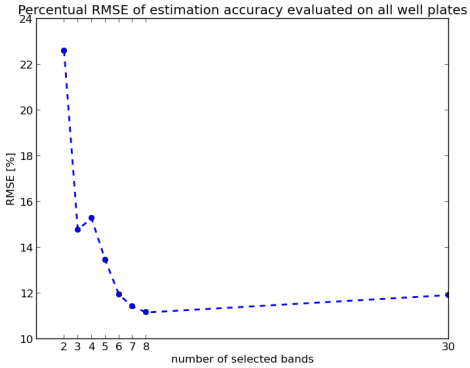


Figure 6.4: The Root Mean Square Error (RMSE) of the relative Congo red (CR) concentration estimation of all analyzed spectra in the phantom experiments as a function of the number of selected bands. Reprinted with permission from Wirkert et al. [194].

validation choosing band 510 instead of 520 and one validation choosing band 620 instead of 590. The chosen bands are depicted in Figure 6.5 in combination with the reference spectra of oxygenated and de-oxygenated hemoglobin.

As in the phantom experiments, the oxygenation estimation performance was compared for seven bands and 20 bands estimate. Again, the modified Beer-Lambert's law was used for estimation. The 20 band results were defined as reference values, reasoning that the selected bands should be able to produce stable, similar results. Figure 6.6 shows how the median absolute errors and inter quartile ranges (IQRs) are distributed on each of the acquired data sets for the three and seven band ESI estimate. As baselines the ESI was compared with the seven maximum variance bands and seven equidistant bands, keeping the normalizations and registration described in 6.1.3 in place. The results for all image stacks combined can be found in Fig. 6.7 In Fig. 6.8, exemplary oxygenation estimates are shown.

6.3. DISCUSSION

The ESI presented in this chapter identifies the most relevant spectral bands. It adapts a popular information theoretic approach stemming from the remote sensing community to the endoscopic environment. The overall approach and the results are discussed in the following.

ESI. The ESI enables band selection without the need to obtain labelled training data or simulated data. By maximizing the differential entropy, the selected bands retain a maximium of information provided in the images, taking into account inter-band covariances. In endoscopic environments, the plain SI suffers from misinformation introduced by the imaging system, lighting and tissue motion/structure. To this end, im-

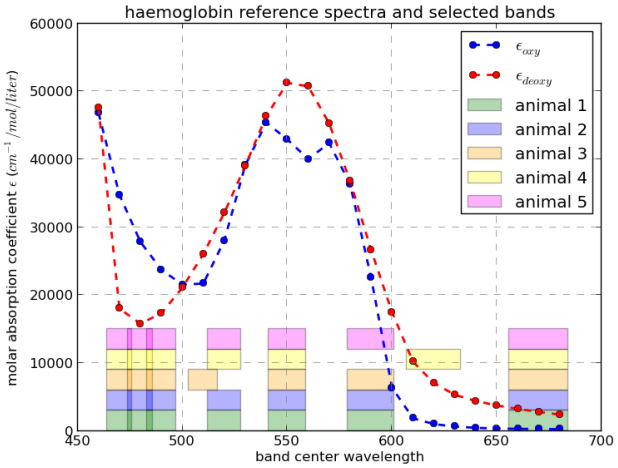


Figure 6.5: The oxy- and deoxy-hemoglobin absorption curves with LCTF responses taken into account. Also indicated are the chosen bands with their respective center wavelength and full widths at half maximum (FWHM). The legend denotes which animal has been left out from training. Reprinted with permission from Wirkert et al. [194].

6

age registration and illumination normalization were introduced as preprocessing steps forming the ESI in Section 6.1.3. The ESI is easy to implement and selects seven bands out of 20 in less than a minute using an unoptimized Octave implementation.

One limitation of the study could be seen in the fact that the method requires the underlying data to be multivariate normal distributed. Non-linearities could be caused by light scattering on particles as collagen and wavelength dependent tissue penetration depths of light. Preliminary analyses of intra-band variances and inter-band scatter plots showed no gross deviations from results expected by normal distributions. The optimal selection of i bands out of n needs evaluation of $\binom{n}{i}$ subdeterminants. While this is feasible for 20 to 30 bands, this number will explode when band selection should be performed from hundreds of bands. In these cases, heuristics as done by Han et al. [82] can be applied.

The information maximized by the ESI algorithm can come from oxygenation, but might also be generated by amount of blood in tissue or variations in scattering. A further investigation could investigate if the domain adaptation method developed in Chapter 5 can be used to find the best task and domain specific bands. One approach could be to first generate simulations weighted by real data as in 5.1.2. From these weighted simulations the bands can be selected by a wrapper method with respect to a metric like RMSE. This would provide the best bands for a specific task, such as cancer hypoxia detection, without the need for labeled *in vivo* data. Contrary to the approach presented in this chapter, a sophisticated light-tissue interaction model would be needed. A further interesting topic would be studying band selection in context of the optimal wave-

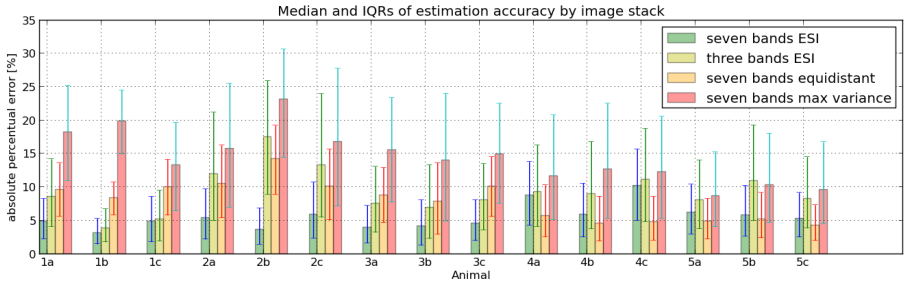


Figure 6.6: Comparison of the median absolute estimation errors and Inter Quartile Ranges (IQRs) for each *in vivo* image stack. As baselines to the endoscopic Sheffield Index (ESI) the result for the seven bands with the highest variance and the result for seven equidistant bands is depicted. Reprinted with permission from Wirkert et al. [194].

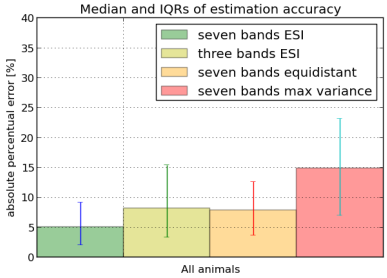
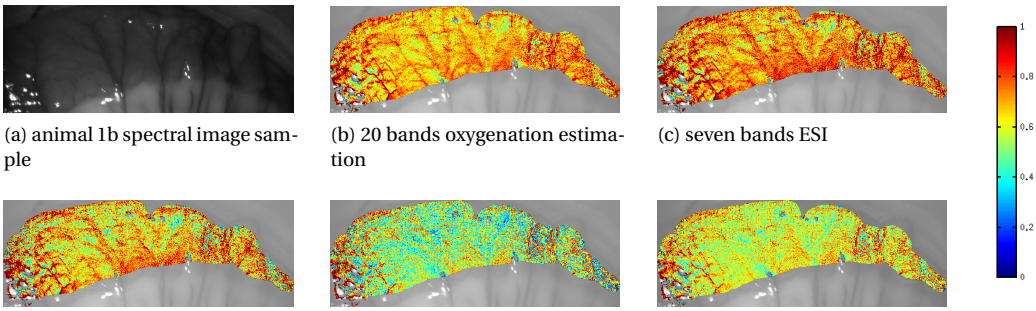


Figure 6.7: Comparison of the overall median absolute estimation errors and Inter Quartile Ranges (IQRs). As baselines to the endoscopic Sheffield Index (ESI) the result for the seven bands with the highest variance and the result for seven equidistant bands is depicted. Reprinted with permission from Wirkert et al. [194].

length range for MSI. This thesis focused on visible light, because standard endoscopic light sources operate in this range. However, there there are indications that the infrared wavelength range is better suited for certain tasks such as ureter discrimination [142].

Experiments. The phantom experiments showed that relative CR concentration estimation done with seven bands yields results comparable to using all 30 bands. The *in vivo* hemoglobin oxygenation estimates also confirmed that seven bands show results similar to all 20 bands. Using only three bands yields median errors up to a factor of two larger and higher variance compared to the seven band estimate. Additional experiments using the ESI without the normalization, which is the SI with motion compensation led to results worse than the worst baseline method. For animal 4, the 620nm band was chosen opposed to the 590nm band chosen in all the other cross-validations. The results for this animal are noticeably worse. Re-evaluation with the 590nm band showed errors similar to the other animals. This suggests that the higher errors for animal four are caused by the suboptimal selection of the 620nm band.

The three band ESI performs comparably to the seven equidistant band selection. Overall, the seven band ESI error is 36% lower than the best baseline method in the *in vivo*



(a) animal 1b spectral image sample

(b) 20 bands oxygenation estimation

(c) seven bands ESI

(e) three bands ESI

(f) seven bands max variance

(g) seven bands equidistant

Figure 6.8: The estimated hemoglobin oxygenation with various band numbers and methods for animal 1b. Red means high oxygenation. Reprinted with permission from Wirkert et al. [194].

experiments. The seven bands estimate had a median deviation of only 5% from the 20 bands estimate. The band selection almost always chooses the same bands with only twelve spectral image stacks as training data in the *in-vivo* study. Interestingly, while the 670nm band was always chosen as relevant band, its neighbouring bands are not chosen in the seven band estimate, although they show similarly high differences between oxy- and deoxy-hemoglobin (see Figure 6.5). This is most likely due to high dependencies between these bands.

In this study image stacks from four different animals served as training data. Future work should investigate how much diversity in animals and number of sample images is necessary to obtain robust estimation. Ideally, a single surgery should be sufficient for training. The presented results encouraged swapping the LCTF with a fast filter wheel, using the filters determined in this chapter. This system was used in chapters 4 and 5. The technology change reduced image acquisition times from 10.5s to 0.4s. Due to the higher frame-rate this system is less susceptible to tissue motion and thus more robust.

6.4. CONCLUSION

This chapter described the Endoscopic Sheffield Index (ESI), which can be used to determine the most informative spectral bands. To my knowledge this was the first adoption [194] of an unsupervised spectral band selection technique for endoscopic applications. It enables band selection without the need to obtain labelled training data or simulated data. The experiments include a porcine study and showed that the selected bands are highly suited for oxygenation estimation.

In conclusion, the ESI presented in this work can to be used for band selection in endoscopic procedures and could thus enable highly time resolved spectral tissue classification. For single-shot recording snapshot cameras the technique can increase resolution: if the 16 off-the-shelf bands from the camera used in the next chapter were to be replaced with the eight bands proposed in this chapter the relatively low resolution would double.

7

First fluent, in-patient, multispectral functional imaging

Figure 7.1: The video-rate multispectral laparoscope (left), compared to a standard RGB laparoscope (right).

The vision that has driven this thesis is a laparoscope that looks and feels familiar to the surgeon; and yet; at the press of a button gives immediate access to functional information. The system presented in chapters 4 and 5 records and analyses images at a speed of approximately 2.5Hz. This low frame rate leads to misaligned multispectral image stacks and hinders widespread adoption by surgeons as these generally prefer seamless feedback. Furthermore, the system is too bulky and heavy to be held steadily during interventions.

While the last chapters focused on algorithms and camera design, this chapter is devoted to a entire system ready to be used by the physician. It presents the first multispectral laparoscope capable of recording and analyzing images at video rates (25Hz). An important feature of this system is that it boasts a form factor similar to a conventional laparoscope (see Figure 7.1). To asses the true clinical benefit of the new device, it was used for detection of ischemia during an in-patient partial nephrectomy.

With its focus on a holistic system, this chapter touches upon all the main topics of multispectral imaging: recording, analysis and application. The new laparoscopic recording system is presented in the next Section. Section 7.2 describes how the analysis method from chapters 4 and 5 can be extended to give immediate feedback. This system was used for recording during in-patient partial nephrectomy, an application discussed in Section 7.3. The results of the experiments, presented in Section 7.4 are discussed in Section 7.5; concluding thoughts can be found in Section 7.6.

7.1. VIDEO-RATE LAPAROSCOPE: IMAGE RECORDING

The laparoscope builds upon the Ximea (Muenster, Germany) MQ022HG-IM-SM4X4-VIS multispectral snapshot camera. The small (26x26x31mm) and light (32g) multispectral camera records images with the imec (Leuven, Belgium) mosaic snapshot sensor with a frame rate of up to 170Hz for image cubes of 512×272×16 resolution and a bit depth of 10. The sensor acquires the complete multispectral image at a single snapshot, using a 4x4 repeating mosaic pattern. Figure 7.2 shows the mosaic pattern when zoomed into a single image.

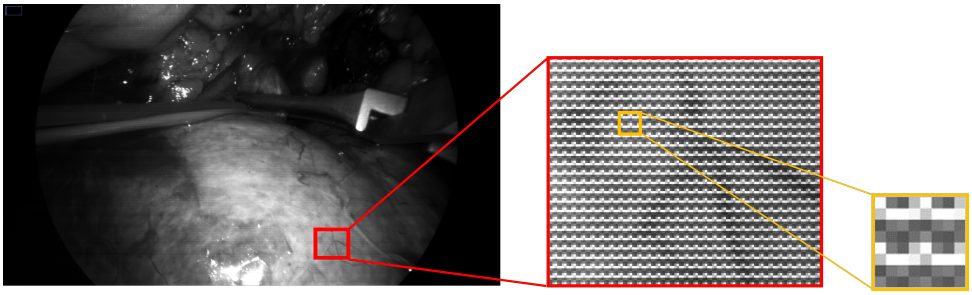


Figure 7.2: *In vivo* image recorded with the Ximea camera. The zooms in the right images reveal the mosaic pattern.

The Ximea was connected via Universal Serial Bus (USB) 3.0 with a MSI GT73VR 7RF-297 Titan Pro 4K Gaming Notebook with Nvidia GeForce GTX1080 8GB graphics card. A C-mount to laparoscope adapter, provided by Richard Wolf (Knittlingen, Germany), connected the Storz (Tuttlinge, Germany) 26003BA 30°laparoscope lens to the camera. Light was provided with the Storz D-light P Xenon light source. The system is depicted in Figure 7.3, for a comparison to the previous Spectrocam system, refer to Figure 7.4.

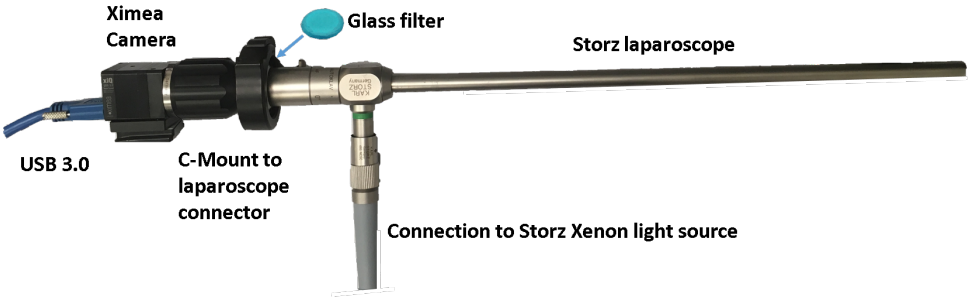


Figure 7.3: The proposed system, which connects the Ximea snapshot camera to a Storz laparoscope. .

As tissue reflects red brighter than blue, a 335 - 610nm bandpass filter[1] was placed between connector and laparoscope. This ensures both information from red and blue are recorded at more balanced camera counts, and thus similar noise levels.

[1] Thorlabs (Newton, NJ, USA), FGB37 - Ø25 mm BG40 Colored Glass Bandpass Filter, 335 - 610 nm

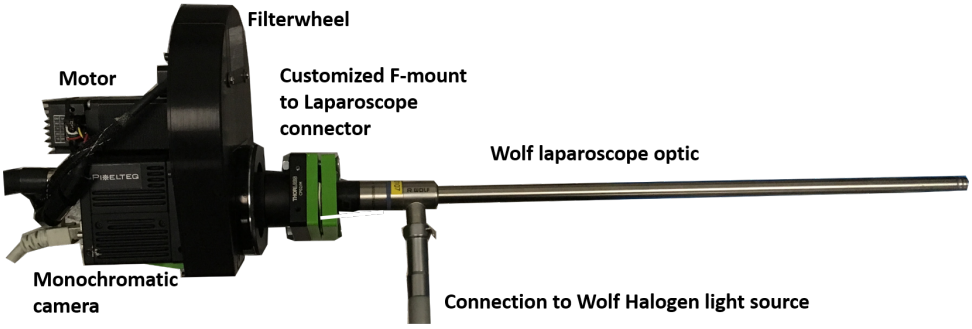


Figure 7.4: The slower and bulkier Spectrocam fast filter wheel system presented in Chapter 4. .

7.2. VIDEO-RATE LAPAROSCOPE: IMAGE ANALYSIS

The random forest approach proposed in chapter 4 needs about 0.14s to generate the functional images for this camera, thus not fulfilling the 25Hz requirement. To match the fast recording with equally fast analysis, the random forests were replaced with convolutional neural networks (CNN) [116]. CNNs were chosen for three reasons:

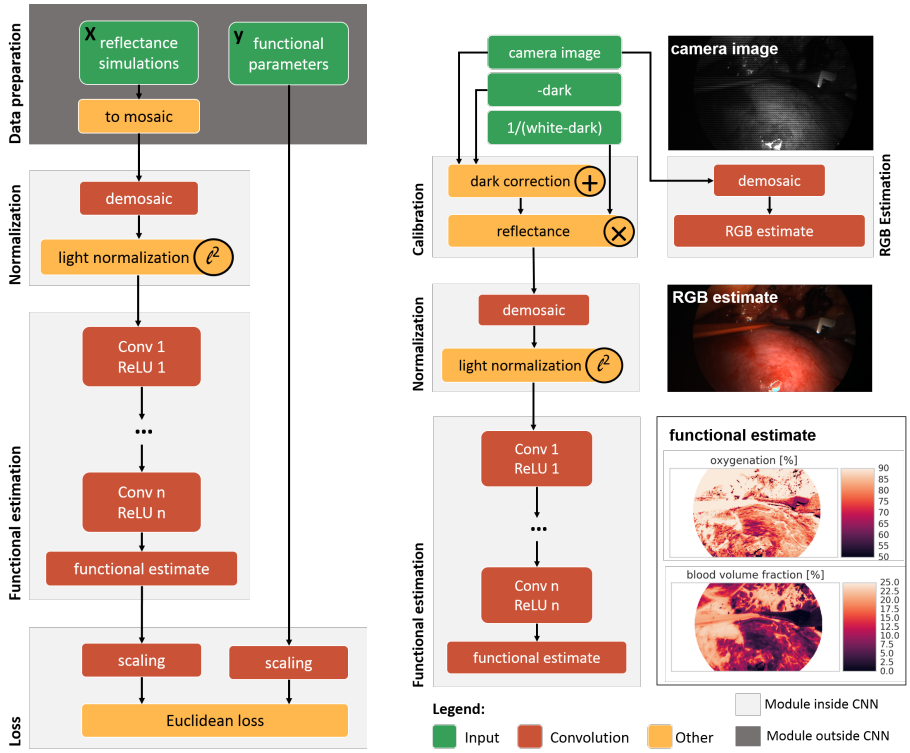
1. modern CNN libraries natively offer GPU acceleration, thus provide rapid inference.
2. Image demosaicing and calibration can naturally be formulated as (convolutional) operations within the CNN architecture.
3. CNNs can operate on whole images and image patches, thus taking advantage of neighbourhood information for contextual information processing.

Two networks were created. The offline training network (Figure 7.5a) determines weights and biases for functional estimation. The trained weights and biases are put into the live evaluation network (Figure 7.5b), which runs within a C++ application on the notebook and processes new images during the intervention. The networks can be separated into several modules, explained in the following.

DATA PREPARATION

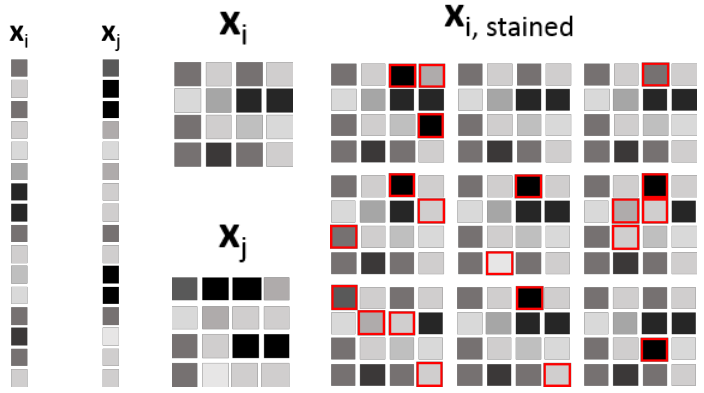
The data preparation module is for training only. It's propose is to convert the simulations of spectral reflectance into data which as closely as possible resembles measurements from the Ximea camera.

The generic spectral reflectance simulations specified in Chapter 5, Table 5.1 were reused for this purpose. Through camera quantum efficiency and filter transmission measurements provided by the manufacturer and light source relative irradiance, laparoscope transmission and glass filter transmission measurements from a spectrometer, the spectral reflectance simulations were transformed to normalized reflectance space (Chapter 4, Equations 4.5 and 4.7).



(a) Training network (b) Live evaluation network

Figure 7.5: The principle flowgraph of the developed CNN, with modules described in more detail in the manuscript text. (a) The network used for training, whose weights and biases determined for the convolution blocks are copied into the live evaluation network. (b) The live evaluation network is used during the intervention. The raw camera image is preprocessed within the CNN framework in the calibration and normalization modules. The learned Conv/ReLU blocks estimate both oxygenation and blood volume fraction. A second branch simultaneously estimates an RGB image to provide a reference familiar to the surgeon.



(a) Two simulated spectra (b) Reshaped to mosaic (c) Expanded to 3x3 and stained $\mathbf{x}_i$ with $\mathbf{x}_j$ (marked red)

Figure 7.6: Instead of performing regression on single pixels, the presented approach takes into account small neighborhoods. Within this neighborhood the tissue can change. (a) To simulate these spatial changes in tissue, for each simulated sample $\mathbf{x}_i$ a random "staining sample" $\mathbf{x}_j$ was selected. (b) Both are reshaped into the mosaic pattern recorded by the sensor. (c) Elements of $\mathbf{x}_i$ were replaced by elements of $\mathbf{x}_j$ with a certain probability. Replaced elements are indicated with a red border in this example. The process of rearrangement and staining was repeated 3x3=9 times to form a small simulated patch of data. Figure 7.15 shows the effect of this procedure.

7

These pixel-wise normalized reflectance simulations were used as basis for functional estimation in the previous chapters. However, the mosaic pattern of the imec sensor has a certain spatial extent. Changes in tissue composition within the mosaic could thus lead to abrupt changes of the measured spectra at these locations. To account for this effect, mosaics were simulated by rearranging the simulated normalized reflectances $\mathbf{x}_i$ in the mosaic pattern. A second, randomly drawn, disturbing simulation $\mathbf{x}_j$ was arranged in the same manner. The pixels of the original mosaic are exchanged with pixels of the disturbing mosaic with a certain probability (e.g. 20%).

To make the results more robust, a small patch of 3x3 mosaics was simulated, so that one training sample consist of 9=3x3 mosaics with the same base $\mathbf{x}_i$ and staining $\mathbf{x}_j$ reflectances. The whole procedure is shown in Figure 7.6. To model sensor noise, zero mean Gaussian noise was added to the simulations as in previous chapters. Oxygenation and blood volume fraction were selected to be the functional parameters for regression. These were set to the mean values within the first 250 μm within the artificial tissue. The simulated mosaic patches served as input to the normalization module, which is part of the training and of the live evaluation network.

NORMALIZATION

The mosaic images are demosaiced to separate spectral and spatial domains into individual axis. This can for example be realized with bilinear interpolation [205] or by

learning the interpolation within the CNN framework [65]. Because demosaicing is not the focus of this chapter, it was implemented by stacking each mosaic of 4x4 to a 16 dimensional vector. This operation was implemented by shifting a 4x4 kernel with 16 output channels and a stride of four over the image. Each of the 16 channels had a 1 entry in the extracted mosaic pixel and zeros elsewhere. This simple way of demosaicing introduces a small spatial misalignment.

Both real measurements and simulations need to be normalized to account for constant changes in illumination. These changes are caused by varying light source intensities or laparoscope poses (Chapter 3.3.1). In this chapter, this was done by dividing each of the demosaiced reflectances by its ℓ^2 norm. The further normalization to absorbance done in Chapter 4 did not show any benefit for CNNs and thus was omitted. The normalized reflectances served as input for the functional estimation module.

FUNCTIONAL ESTIMATION

The functional estimation module is "heart" of the network, in which the relationship between input data and functional parameters is established.

It was composed of several convolution/rectified linear units (ReLU) [136] blocks. In the most simple realization, all blocks are 1x1 convolutions, denoting the same result as a fully connected, pixel-wise network applied to each image pixel. If patch-wise 3x3 mosaic input is generated as described in Section 7.2, the first two layers are 2x2 convolutions, followed by 1x1 feature transformations. Because of the fully convolutional architecture, patches and images of arbitrary sizes can serve as input to the network. The number of convolutional layers and filters depend on the implementation, weights and biases are learned with respect to the loss described next.

LOSS

The loss is the objective function to be minimized during training. An Euclidean loss was chosen, defined as

$$\frac{1}{2N}\sum_{i}^{N}||\mathbf{y}_{\text{true},i}-\mathbf{y}_{\text{pred},i}||_2^2, \tag{7.1}$$

with N being the elements in the training batch. Functional parameters come in differing ranges. In the employed data set, oxygenation varied from 0-100%, while blood volume fraction only ranged from 0-30%. To ensure that different ranges of the estimated parameters do not influence the loss, the real $\mathbf{y}_{\text{true}}$ and estimated $\mathbf{y}_{\text{pred}}$ were scaled. This scaling was determined on the training data so that the training labels range from 0 to 1 for each parameter. The trained weights and biases were copied to the *in vivo* evaluation network.

CALIBRATION

Before the normalization and functional estimation can be applied to the *in vivo* network, the camera images are transformed to reflectance (see Chapter 3.3.1).

Two images were recorded before the intervention for this purpose: a dark image with no light and a white image, taken from a white reflectance standard. They were stored in the network to be simply added (−dark) and multiplied (1/(white − dark)).

RGB ESTIMATION

To show the physician a familiar image as a reference during the intervention, an RGB image was simultaneously estimated.

This calculation took place in a separate path and aimed to reconstruct an idealized RGB filter response $F_{\text{RGB}} \in \mathbb{R}^{3xm}$ from the multispectral systems response $F_{\text{multispectral}} \in \mathbb{R}^{|\mathbf{b}|xm}$ using a linear transformation $T \in \mathbb{R}^{3x|\mathbf{b}|}$:

$$F_{\text{RGB}} = T \cdot F_{\text{multispectral}} \tag{7.2}$$

with m being the number of parsed wavelengths and $|\mathbf{b}|$, the number of multispectral bands. Filter matrix $F_{\text{multispectral}}$ can incorporate light source irradiance and transmission of the laparoscope to make the estimation inherently white balanced. The linear transformation T was found by least squares regression and put into the network. Here the transformation were represented by three 1x1 filters, one for each R, G and B, with no biases.

7.3. VIDEO-RATE LAPAROSCOPE: APPLICATION

The target application is partial nephrectomy, in which a kidney tumor is surgically removed. During these interventions, often the renal artery has to be clamped to prevent bleeding [179]. Verification of correct clamping is not straightforward; especially if the preferable selective clamping of a segmental artery is performed, in which ischemia is induced only in the cancerous part of the kidney [21, 131]. One possibility to ensure correct clamping is to check perfusion with indocyanine green (ICG) fluorescence: after ICG is injected in the blood stream, it binds to the plasma. The bound ICG travels through the blood stream and accumulates in the internal organs, especially in the kidney and liver within a minute [59, 181]. No fluorescent signal thus corresponds to no perfusion. Due to long washout periods of about 30 minutes, this test is not repeatable if the wrong segment has been clamped [59]. Multispectral imaging (MSI) could be a possible substitute for the only once applicable fluorescence method. Likewise, ICG fluorescence can serve as gold standard perfusion estimate for MSI.

7.4. EXPERIMENTS AND RESULTS

The previous section described a translational solution for MSI in laparoscopy. This section aims at analyzing the proposed system with respect to recording, analysis and the performance in a first human subject.

Section 7.4.1 presents the experimental setup in the operating room (OR), the CNN parameters and metrics. The in-patient results are summarized in Section 7.4.2, including

a comparison to baseline methods. The employed CNN analysis method is evaluated in Section 7.4.3 and the novel recording system setting is assessed in Section 7.4.4, with a special focus on the imec mosaic sensor.

7.4.1. GENERAL SETTING

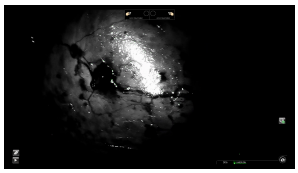
This section gives an overview on the context in which the recordings were conducted, the standard CNN settings and training procedure and general metrics.

***In vivo* recording.** In the scope of this thesis, multispectral recordings were taken from a patient undergoing partial nephrectomy. To ensure a sterile environment, a sterile laparoscope was connected with the rest of the system via a sterile, disposable camera cover. To guarantee successful renal clamping, ICG fluorescence measurements were taken, using a da Vinci® Si (Intuitive Surgical, Sunnyvale, CA, USA) surgical system. The white and dark images as well as relative irradiance measurement were taken after the intervention.

CNN training procedure. For training the inversion, the simulation sets described in Chapter 5 were used. The batch size for training was set to 500 (out of 500,000) for all experiments, testing was done on the separate set of 50,000 simulations. One epoch is defined as one forward/backward pass through all the training examples (i.e. 1,000 training steps). To enable inference in C++, the Caffe [99] deep learning framework was used. Minimization was guided by the Adam optimizer [110] with standard parameter settings and a base learning rate of 0.005. The learning rate was multiplied with 0.1 every 300 epochs. To prevent overfitting to the simulations, a L2 regularization term for the weights was added by setting the weight decay parameter to 0.001. Weights were initialized with the Xavier initialization [69], biases were initialized to 0.1.

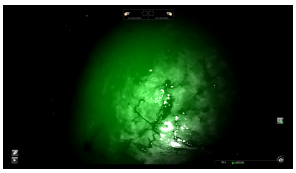
Baseline methods. Two baseline methods were investigated: the random forest approach, with parameters and normalizations as described in Chapter 4 and the standard Beer-Lambert regression (Chapter 3.3.2). Beer-Lambert is not applicable directly, because the measured bands are broad and show second order peaks (see Section 7.4.4). To still enable Beer-Lambert estimation, the 16 band data was transformed to 12 narrow virtual bands. The linear transformation to these virtual bands was determined by an algorithm of the camera manufacturer.

Standard metrics. Oxygenation and blood volume fraction were estimated. Metrics were calculated on the absolute percentual errors of these functional parameters with ground truth provided by the separate test set. These are referred to as oxygenation and blood volume fraction error in the *in silico* evaluations. All *in silico* experiments were conducted five times to reduce random influences. If not mentioned otherwise, results refer to the concatenation of these five runs (e.g. the median error of all five runs).



(a) Before

(b) During



(c) After

Figure 7.7: Fluorescent signal recorded before, during and after applying the clamps. (a) Before injection, no fluorescent background signal was detected. (b) After clamps had been applied and ICG was injected, no signal was detected in the liver. (c) Release of the clamp and thus reperfusion of the kidney was immediately followed by a clear green fluorescent signal.

7.4.2. *In patient* APPLICATION

Recordings were taken from a patient undergoing partial nephrectomy in collaboration with the Department for Urology of the Heidelberg University Hospital. The experiment was conducted in three phases:

0. A baseline non-fluorescing image was taken to spot potential background signals. A video was taken from the unclamped kidney with the da Vinci®. Baseline measurements were also acquired with the MSI system (sequence 1).
1. The renal artery was clamped and the kidney was monitored at video rate by the MSI system for about 90s. ICG was prepared by mixing ICG PULSION 50mg powder with 10ml of distilled water. After applying a second clamp 2ml of the ICG solution were admitted to the central venous catheter. A second video from the da Vinci® was taken to ensure the clamping successfully stopped perfusion.
2. The clamp was removed and the reperfusion was monitored with the MSI system (sequence 2). An video of the perfused kidney was acquired with the da Vinci®.

The fluorescent videos taken before, during and after clamping confirmed perfusion was stopped (see Figure 7.7). The data acquired from the MSI system is analyzed in the following. First functional estimation with the proposed approach is presented, next it is compared to two baseline methods for oxygenation estimation.

OXYGENATION AND BLOOD VOLUME FRACTION TRACKING

Two candidates for characterizing ischemia are blood volume fraction and oxygenation. This experiments investigates if the parameters correlate with perfusion and hence how suited they are for ischemia monitoring.

A template matching based[2] ROI tracker was developed to follow an ROI through the videos. Due to substantial movement and thus perspective change, the ROI was reinitialized at approximately the same position after (1) applying the first clamp (2) applying the second clamp (3) before releasing the clamps.

[2] Using opencv-python [97] v3.2.0.6 method matchTemplate, with TM_CCOEFF_NORMED parameter

Figure 7.8a depicts the mean oxygenation and blood volume fraction after the artery was clamped (sequence 1). It can be observed that after the clamp was applied, both parameters reduced. Blood volume fraction dropping by more than 50% in less than 15s. Figure 7.8b shows some key frames from the sequence.

The approximately same ROI was tracked for the sequence 2, in which the clamps were released. The results in Figure 7.9a show both blood volume fraction and oxygenation sharply increasing after the clamp was released. Figure 7.9b shows four key frames.

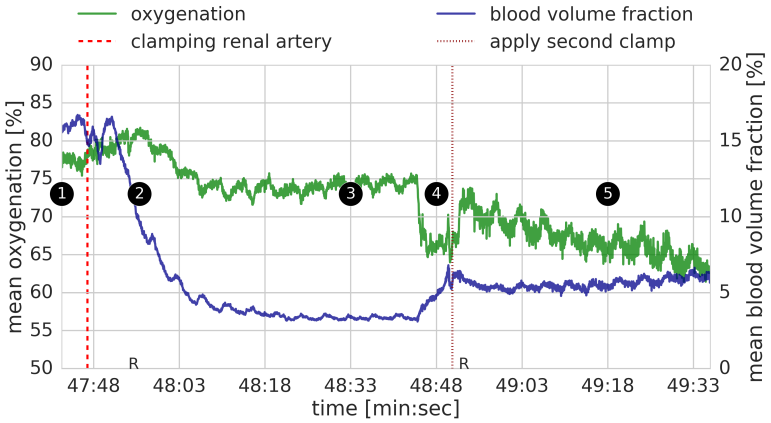
METHOD COMPARISON

To get a qualitative feel for the performance of the proposed method, results for oxygenation estimation were compared to the standard Beer-Lambert (BL) approach and to the random forest (RF) method presented in Chapter 4. Further, the proposed simulation of small neighborhoods was compared to other possible simulation schemes.

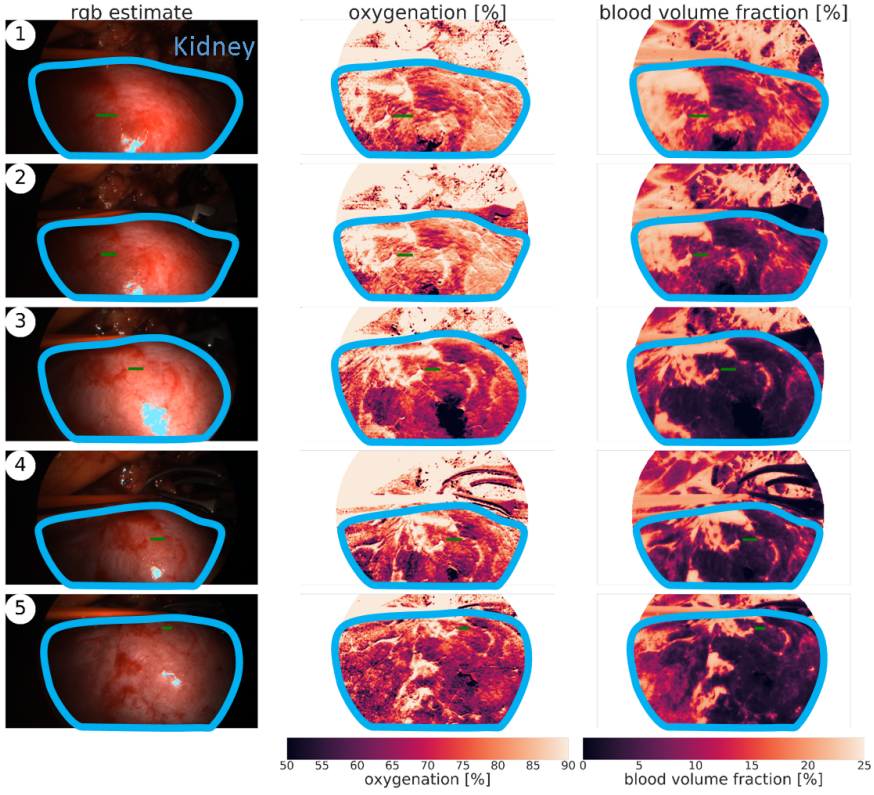
Comparison to BL and RF. Figures 7.10 and 7.11 show the oxygenation as estimated by the different methods for the same ROI as before. The standard deviations of the oxygenation within the ROI are also depicted. The mean standard deviations measured in sequence 1 were 4.6 (CNN), 9.2 (RF) and 20 (BL) percentage points. The mean standard deviations measured in sequence 2 were 3.7 (CNN), 8.5 (RF) and 13.4 (BL) percentage points. Figure 7.12 shows example oxygenation images of sequence 1, estimated by the different methods.

Comparison to pixel-wise training and no staining. Contrary to conventional pixel-level analysis, Section 7.2 proposed processing a small 3x3 neighborhood. To model tissue changes within the small region, some pixels were exchanged by other spectra, referred to as “staining”. To investigate the effect of staining and patch-wise learning, four networks were trained, one for each combination of patch-wise v pixel-wise and staining v no staining.

Figures 7.13 and 7.14 show the oxygenation as estimated by the different methods for the same ROI as before. The standard deviations of the oxygenation within the ROI are also depicted. The mean standard deviations measured in sequence 1 were 4.3 (patch + staining), 8.8 (patch + no staining), 5.9 (pixel + staining) 7.4 (pixel + no staining) percentage points. The mean standard deviations measured in sequence 2 were 3.6 (patch + staining), 8.8 (patch + no staining), 5.8 (pixel + staining) and 7 (pixel + no staining) percentage points. Figure 7.15 shows an oxygenation image of sequence 1, estimated by the different methods. Blood volume fraction estimates were similar for all compared methods.

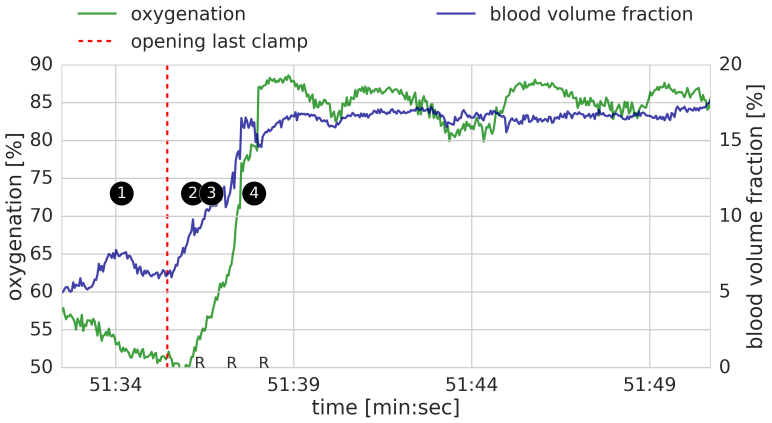


(a) Mean functional parameters monitored while applying clamps to renal artery. "R" indicates reinitialization of the ROI. Example frames are depicted in Fig. 7.8b.

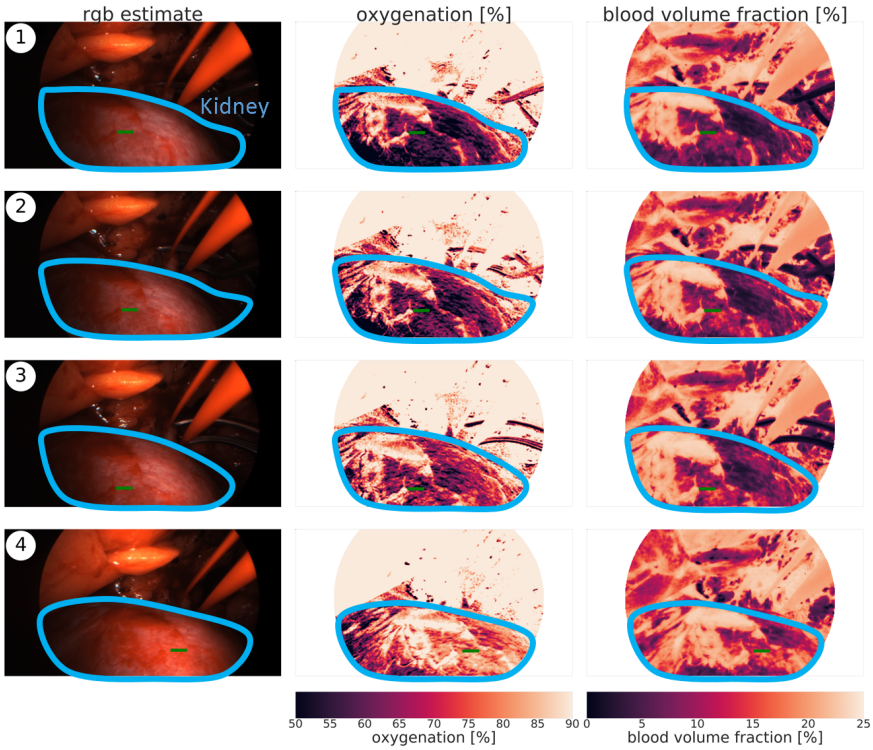


(b) Example frames corresponding to Fig. 7.8a. The small green lines mark the ROI.

Figure 7.8: Sequence 1: two clamps were applied to the renal artery. A drop both in oxygenation and blood volume fraction was measured in the ROI.

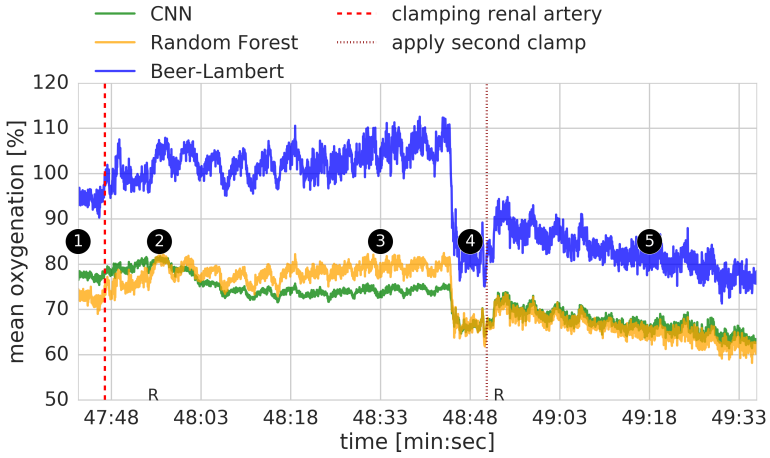


(a) Mean functional parameters monitored while blood flow was restored. "R" indicates reinitialization of the ROI. Example frames are depicted in Fig. 7.9b.

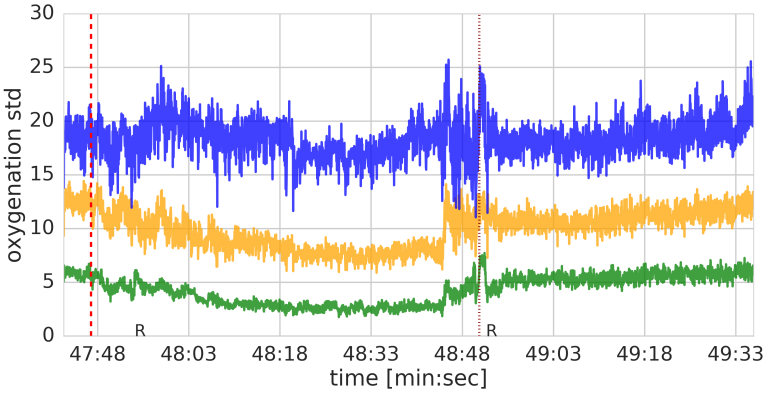


(b) Example frames corresponding to Fig. 7.8a. The small green lines mark the ROI.

Figure 7.9: Sequence 2: when releasing the last clamp, a sharp increase in both parameters was measured in the ROI.

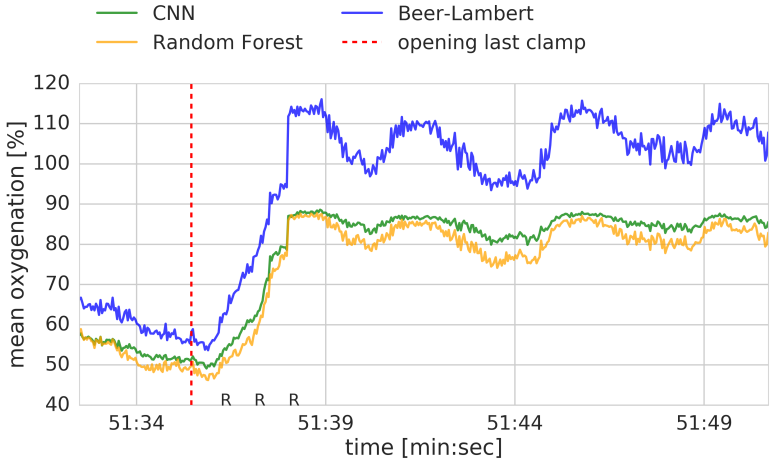


(a) Mean oxygenation estimated by different methods while applying clamps to renal artery. Example frames are depicted in Figure 7.12.

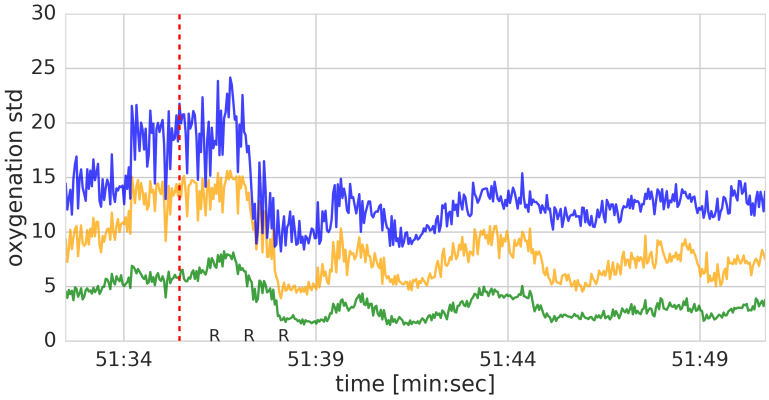


(b) Standard deviation of oxygenation estimated by different methods while applying clamps to renal artery.

Figure 7.10: Sequence 1: two clamps were applied to the renal artery. (a) shows the oxygenation estimate for the proposed CNN method and the baseline Beer-Lambert and random forest approaches. The small dip of oxygenation within the first minute after clamping could only be measured with the proposed CNN approach. Beer-Lambert frequently estimates impossible oxygenation values greater than 100%. (b) The standard deviation of the different methods within the ROI. The other methods furthermore produce more noisy estimates. “R” indicates reinitialization of the ROI.



(a) Mean oxygenation estimated by different methods while releasing clamps to renal artery.



(b) Standard deviation of oxygenation estimated by different methods while releasing clamps to renal artery.

Figure 7.11: Sequence 2: last clamp is released from renal artery. (a) shows the oxygenation estimate for the proposed CNN method and the baseline Beer-Lambert and random forest approaches. Beer-Lambert frequently estimates impossible oxygenation values greater than 100%. All methods detect the reoxygenation after releasing the clamp. (b) The standard deviation of the different methods within the ROI. While all methods seem to follow the same trend, the CNN estimate is substantially less noisy. "R" indicates reinitialization of the ROI.

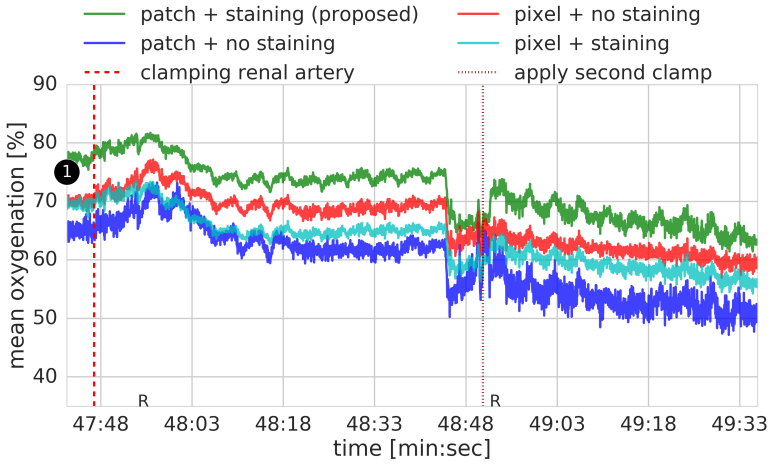
Figure 7.12: Sequence 1: example frames from the proposed CNN method and the two baseline methods corresponding to Figure 7.10. The CNN results are substantially less noisy; the Beer-Lambert approach repeatedly estimated physically impossible oxygenation values greater than 100%. The small green lines mark the ROI analyzed in Figure 7.10.

7.4.3. EVALUATION OF ANALYSIS METHOD

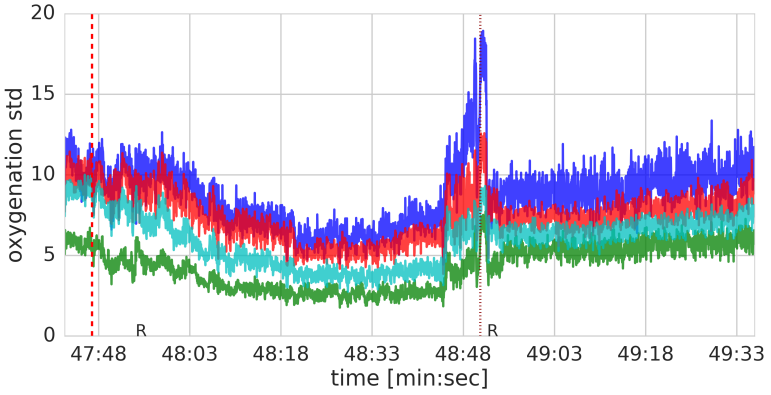
In this chapter, the previously used random forests were exchanged by CNNs. These experiments evaluate the execution speed, the convergence of the training procedure, the differences caused by changes in network architecture and compares CNNs to the random forest and conventional Beer-Lambert approach. Except for the run time analysis the experiments were conducted in the *in silico* setting described in Section 7.4.1.

VIDEO-RATE ANALYSIS

One of the main advantages of the CNN approach is that images can be processed end-to-end on the GPU. This means new camera images are directly processed by the CNN architecture depicted in 7.5b, which incorporates image calibration, normalization and functional and RGB estimation.

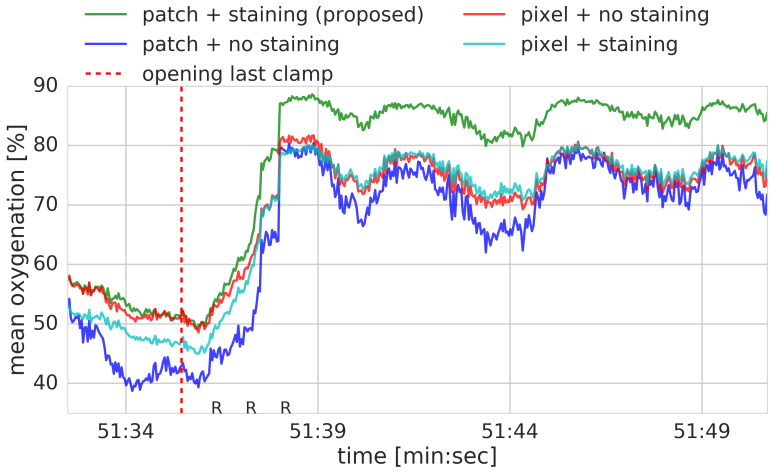


(a) Mean oxygenation estimated by different methods while applying clamps to renal artery. Example frames are depicted in Figure 7.12.

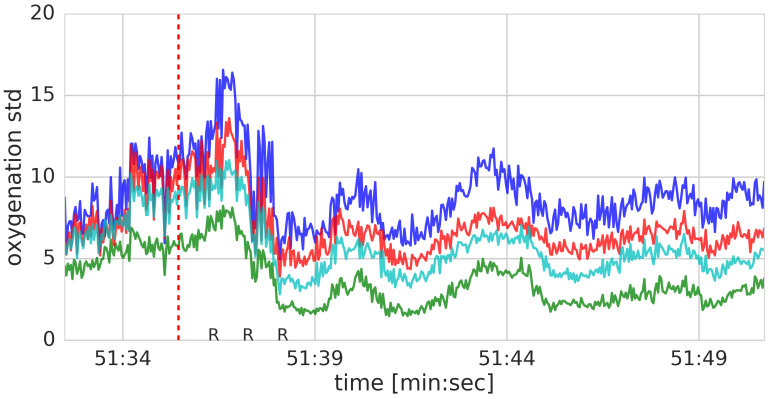


(b) Standard deviation of oxygenation estimated by different methods while applying clamps to renal artery.

Figure 7.13: Sequence 1: two clamps were applied to the renal artery. Different training methods for the CNN are compared: either the training is based on a small neighborhood (patch) or a single pixel (pixel). For both these methods additional noise coming from nearby tissues can be simulated (staining) or not (no staining). (a) All methods show the same general trend. The proposed approach estimates higher non-ischemic oxygenation. (b) The standard deviation of the different methods within the ROI. While all methods seem to follow the same trend, the proposed method is substantially less noisy. An example frame marked as circled 1 is depicted in Fig. 7.15. "R" indicates reinitialization of the ROI.

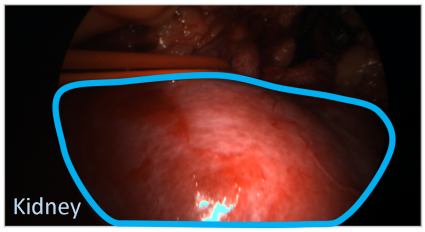


(a) Mean oxygenation estimated by different methods while releasing clamps to renal artery.

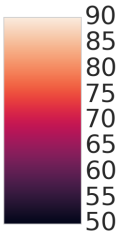


(b) Standard deviation of oxygenation estimated by different methods while releasing clamps to renal artery.

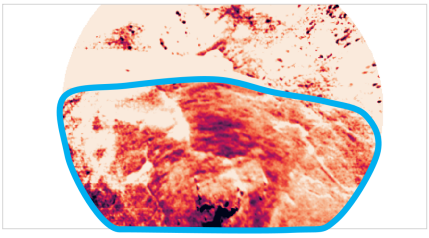
Figure 7.14: Sequence 2: last clamp is released from renal artery. Different training methods for the CNN are compared: either the training is based on a small neighborhood (patch) or a single pixel (pixel). For both these methods additional noise coming from nearby tissues can be simulated (staining) or not (no staining). (a) All methods show the same general trend. The proposed approach estimates higher non-ischemic oxygenation, which is physiologically more plausible. (b) The standard deviation of the different methods within the ROI. While all methods seem to follow the same trend, the proposed method is substantially less noisy. “R” indicates reinitialization of the ROI.



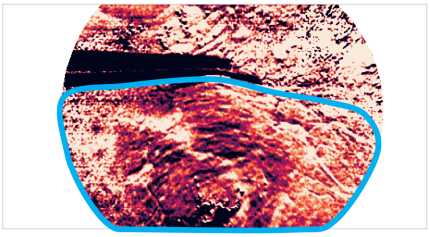
(a) RGB



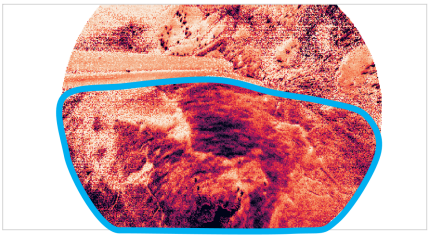
(b) oxygenation [%] color code



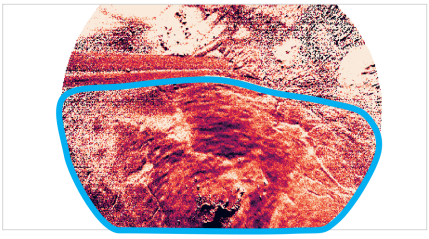
(c) patch wise + staining (proposed)



(d) patch wise + no staining



(e) pixel wise + staining



(f) pixel wise + no staining

Figure 7.15: Example oxygenation estimates using different simulated input data for CNN training, corresponding to Figure 7.13. (a) RGB image for context. The lower two thirds of the images show a well perfused kidney. (c) The proposed network qualitatively performs best. It trains on a small simulated neighborhood (patch). To account for changes in tissue within this region, the simulated patch was "stained" by random other tissue samples as described in Section 7.2. The method detects vessels as highly oxygenated with slightly smaller values estimated for surrounding tissue. This indicates the vessels are arteries, which deliver oxygen to the adjacent tissue. (d) Without adding staining pixels in the training input, the patch wise network seems to overfit, estimating mostly high/low values. (e + f) As expected, estimation done for each multispectral pixel individually leads to noisier results.

To evaluate the exact run time of the model, it was executed on over 1,000 images, on a notebook[3]. The worst time for a complete forward pass of the network was 6.4ms (median 5.4ms), transferring the camera image to the GPU took a maximum of .9ms (median .8ms). This leads to a total run time of less than 7.5ms, or more than 133Hz. The goal of analysis with at least 25Hz was thus realized.

TEST ERROR CONVERGENCE

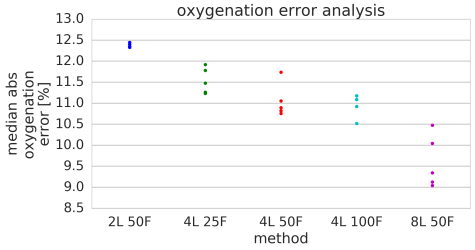
This experiment should ensure that the performance of the network has converged before training was aborted.

Comparing oxygenation estimation error on the test set after ten and 1000 epochs of training, the worst of the five runs showed 4.4 percentage points lower performance for the ten epoch result. After 100 epochs the difference reduced to 2.9 percentage points, after 700 epochs the difference was 0 percentage points. Differences in blood volume fraction errors were smaller. For all further experiments the training was conducted for 1000 epochs.

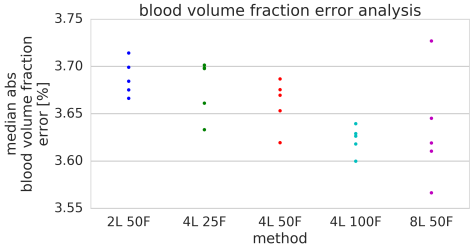
CNN SETUP

This experiment investigated the influence of CNN architectures on the function estimation performance.

Six architectures for CNN regression were tested. They vary the number of convolution/ReLU blocks and the number of filters learned by in each of these blocks. The investigated architectures were 2 Layer (L) 50 filters (F), 4L 25F, 4L 50f, 4L 100F, 8L 50F. The 8 layer network was an exception with 50 filters in the first 4 layers and 10 filters in the last 4 layers. Figure 7.16 shows the oxygenation and blood volume fraction results for the 5 runs of each network.



(a) Median oxygenation estimation results for each of the five runs



(b) Median blood volume fraction estimation results for each of the five runs

Figure 7.16: Influence of network architecture.

A Welch's T-test for independent sample of scores[4] with a 5% significance threshold compared the standard 4L 50F network to the other architectures. Significant difference in

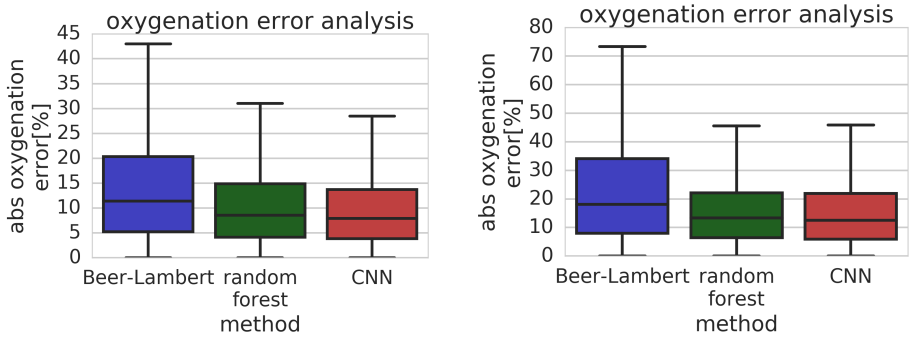
[3] MSI GT73VR 7RF-297 Titan Pro 4K Gaming Notebook with Nvidia GeForce GTX1080 8GB graphics card.
[4] Using scipy.stats.ttest_ind v19.0 [101]

oxygenation estimation performance were determined comparing to the 2 layer network (p-value=0.002) and to the eight layer network (p-value=0.004). For blood volume fraction, only the difference to the 4L 100F network was found significant.

CNN COMPARED TO OTHER ESTIMATORS

To ensure the CNN approach is suited for functional estimation it was compared to RF regression and BL estimation.

The evaluation was limited to oxygenation estimation performance, since BL cannot determine blood volume fraction. Further, since BL can only process single input spectra, only a single mosaic was simulated, contrasting the 3x3 mosaic input used in the other experiments. The evaluation was done with a staining probability of 20% and with no staining, to give results more comparable to the previous chapters. The oxygenation error was 7.9 (12.5)% for CNN, 8.5 (13.3) % for RF and 11.4 (18.1)% for the BL method (stained results are in brackets) See Figure 7.17 for a boxplot of these results.



(a) Estimation methods, evaluated with "unstained" data

(b) Estimation methods, evaluated with "stained" data

Figure 7.17: Comparison of Beer-Lambert, random forests and the proposed CNN approach. "Staining" refers to a method to inject additional shot noise described in Section 7.2. Boxplot, whiskers extend the low/high quartiles by 1.5 IQR, extreme values not shown.

7.4.4. EVALUATION OF IMAGE RECORDING

In chapters 4 and 5 the Spectrocam, a fast filter wheel based camera, was used (Chapter 4.3.1). This slow sequentially recording camera was switched with the Ximea snapshot camera; the system now further uses a Storz laparoscope and Xenon light source, in contrast to the previously employed Wolf laparoscope and Halogen light source. Aside from the speed of image recording, the most significant difference between the two systems are the cameras. The recording speed of the Ximea camera and the "second order" peaks specific to the new sensor were investigated in the following. Other properties of the new system have been analyzed in Appendix A.

VIDEO-RATE RECORDING

The Ximea camera can record up to 170 multispectral images per second. This easily surpasses the 2.5Hz recording of the Spectrocam. However, this value is more of theoretical nature, because proper illumination and thus sufficient camera counts have to be ensured.

The Xenon light source provided brighter and spectrally more even light than the previously used Halogen light source. During the patient study, the exposure time was set to 40ms, recording images with 25Hz. This exposure time was enough to ensure a well lit scene at typical laparoscopic distances with the Xenon light source. Thus video-rate recording is achieved.

CAMERA EVALUATION

The sensor records 16 multispectral bands. Spectral response of these bands are provided by the camera manufacturer and depicted in Figure 7.18. Several of the bands show two peaks in spectral response. They are caused by the measurement principle of the Fabry-Péro filters, which lead to so called "second order" peaks depending on cavity height of the filter, refractive index of material and the cosine of the incident light. These second order peaks can be quite pronounced (see bands 5-12) and are a challenge working with the sensor: as the two peaks are in the red and blue, changes in oxygenation will be less strongly measured, as illustrated in Fig. 7.19. The consequences are explored in the following *in silico* evaluation.

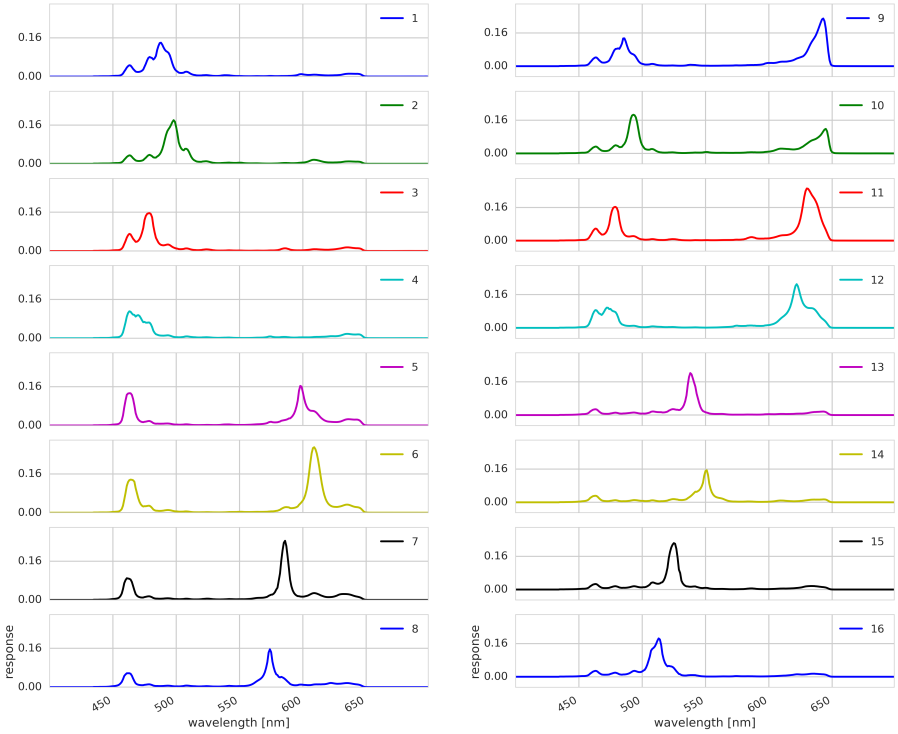
To investigate the influences of the second order peaks, three setups were compared:

1. the real 16 band imec filters,
2. idealized imec filters with no second orders and a gaussian response at the sensors central wavelength and a broadness defined by a Full-Width at Half Maximum (FWHM) of 10nm,
3. The Spectrocam from chapters 4 and 5 with the filters at 470, 480, 511, 560, 580, 600, 660, 700nm center wavelengths and 20nm FWHM, except for the 480nm band with 25nm FWHM.

As Figure 7.20 shows, second orders led to a 1.7 percentage points higher oxygenation error and higher IQRs compared to the idealized system. The error of the Spectrocam is 2.5 percentage points lower than the imec system. Median blood volume fraction errors were within 3.5-3.6% for all systems.

SENSITIVITY TO SETUP CHANGES

This experiment evaluated robustness of the imec camera to setup changes. It mimics changes in the OR, such as use of a different light source or laparoscope, by simulating different setups for training and testing. This corresponds to changing the setup followed by a new white balance.



(a) Bands 1-8

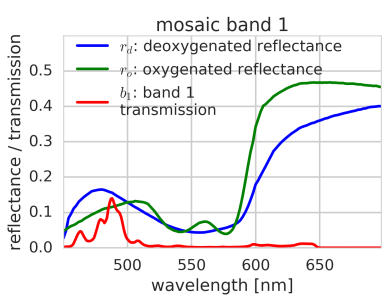
(b) Bands 9-16

Figure 7.18: The filter responses of the 16 imec bands. Some bands, such as 5-12 show two peaks in spectral response, which in this thesis is referred to as "second orders".

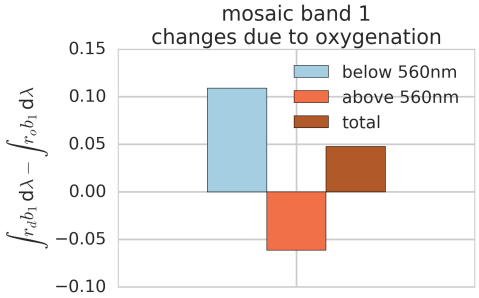
A summary of results for the imec and Spectrocam system is depicted in Figure 7.21. While the imec system was rather robust to smaller changes in optical system (different laparoscope with similar, but not same transmission), performance was reduced when switching from Xenon to Halogen light. The Spectrocam system was robust to both changes.

7.5. DISCUSSION

This chapter introduced the first laparoscopic multispectral system which is able to fluently (≥ 25Hz) record and analyze multispectral images. To enable this fluent recording, the sequential fast filter wheel used in chapters 4 and 5 was exchanged by a snapshot multispectral camera, which acquires spatial and spectral information within one image. Switching the random forest regression from the previous chapters with a CNN based regression approach allowed fast, end-to-end processing of the multispectral images. The results acquired during a partial nephrectomy suggest that multispectral imaging could replace the more invasive ICG method in this context.



(a) Transmission of example band and oxygenated/deoxygenated reflectances



(b) Sensor integration from oxygenated to deoxygenated

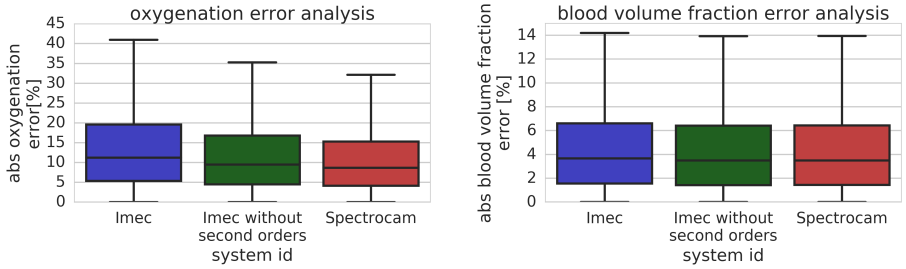
Figure 7.19: By having a second, small peak in the red wavelengths, the sensor measures at high and low wavelengths at the same time. As spectral reflectance is higher at wavelengths above 560nm, even the small peak has an influence. This is further investigated in (b), which summarizes contributions to the camera measurement from below and above 560nm when changing from complete oxygenation to complete deoxygenation. Deoxygention will cause a higher measurement in the blue wavelengths and a lower wavelengths in the red wavelengths. Thus two peaks work against each other, leaving a less pronounced change in measurement.

In the following the in-patient results are discussed, followed by a review of the novel recording and analysis methods in context of the *in vivo* and *in silico* experiments.

Application: in-patient partial nephrectomy. Quantitative evaluation of multispectral images is not straightforward, as was discussed in Section 3.3. What can be tested however is if the technique meets a clinical need. In this chapter, this clinical need was perfusion monitoring during partial nephrectomy. ICG fluorescence served as a reference for the presented technique.

The in-patient results show the first video rate monitoring of oxygenation and blood volume fraction during partial nephrectomy. Before clamping the oxygenation in the ROI was about 80%, a value comparable to in-human measurements from [68]. The blood supply to the kidney was stopped before resection of a tumor by applying two clamps. After clamping, blood volume fraction dropped from 15 to 5% in the tracked ROI within 15s. Oxygenation reduced at a slower pace from 80 to 65% in about 90s (Figure 7.8). When the blood flow was restored, both parameters restored to their pre-clamping baseline within 2s (Figure 7.9). This is a hint that blood volume fraction is the better parameter to characterize perfusion changes. The results are in line with the ICG reference method, a first indication it could be replaced with the less invasive and quantitative MSI.

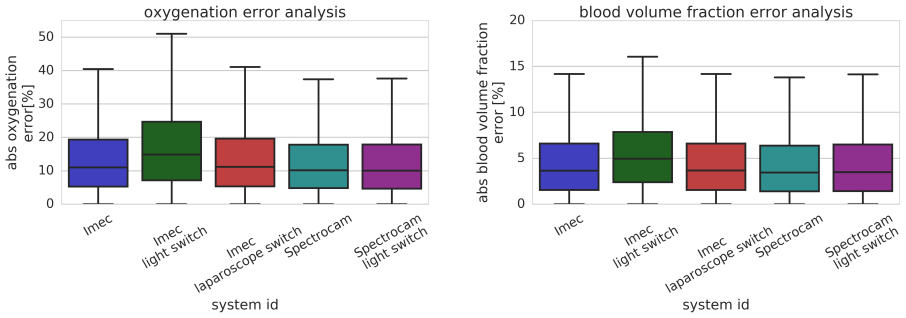
The slower reduction of oxygenation could be explained by most of the blood immediately drained from the organ after the clamp was applied, but the oxygen contained in the remaining blood being only slowly consumed. Further, a small increase in oxygenation was observed in the first ten seconds after the clamp was applied, followed by a larger drop in the next 10s. This effect can also be seen in the example images, Fig.



(a) Oxygenation performance

(b) Blood volume fraction performance

Figure 7.20: Comparison of the imec sensor, an idealized imec sensor and the Spectrocam from previous chapters. Boxplot, whiskers extend the low/high quartiles by 1.5 IQR, extreme values not shown.



(a) Oxygenation performance

(b) Blood volume fraction performance

Figure 7.21: Consequences of switching light source or laparoscope during the intervention on estimation performance. Results based on simulations of the imec sensor and the Spectrocam system from previous chapters. Boxplot, whiskers extend the low/high quartiles by a 1.5 IQR, extreme values not shown.

7.8b images 1-3. A possible reason could be that the venous blood is transported from the kidney faster than the arterial blood, which loses pressure due to clamping. This brief imbalance might lead to a small increase of oxygenation until most of the venous blood has left the kidney. In both recorded sequences (see figures 7.8 and 7.9) periodic oscillations of oxygenation and blood volume fraction occur with a frequency of approximately 3-4 per 15s. These correspond with the respiratory cycle of the patient. However, it is unlikely that the oscillations correspond to real changes in physiology but are caused by tracking of the 2D roi template on the 3D kidney surface. This could also be the reason for the dip in oxygenation shortly before the second clamp was applied, which was preceded by abrupt changes of camera pose. To reduce perspective effects caused by the high curvature of the kidney in the up-down direction of the image, the ROI was chosen to be wide (20-30 pixel) but small in height (10 pixel). An alternative hypothesis for the observed fluctuations would be that the pose influences the measurements. This is unlikely, because inspection of the complete images (see Figure 7.8b images 3 and 4) seems to indicate that the estimation results are unaffected by the change in pose.

Further animal and in-patient studies is currently planned with the Department for Urology of the Heidelberg University Hospital. The animal trials will investigate the influence of pose and monitor the functional parameter estimates on a longer timescale in a controlled environment. The clinical study will further test the hypothesis that ICG can be replaced by MSI for detecting ischemia during partial nephrectomies. Further studies could also determine if oxygenation can serve as an outcome predictor for postoperative renal function, as studied by Best et al. [16]. If this were possible, the surgeon could adapt the surgery to the current state of the kidney, e.g. by early unclamping if otherwise long-term damage could be foreseen due to low oxygenation readings.

Recording: new snapshot system. The fast filter wheel used in the previous chapter had narrow (20nm FWHM) filters, which were recorded sequentially. Due to this sequential, 400ms long, recording, the image stacks can be misaligned by often inevitable motion. This misalignment leads to invalid measurements and is a serious difficulty when working with the sensor, which eventually led to switching the Spectrocam with the Ximea camera, which relies on an imec sensor. The main advantage of the imec snapshot sensor is it records 16 bands simultaneously, within 40ms through a repeating 4x4 mosaic pattern (Fig. 7.2). Together with to the light weight and small form factor (Fig. 7.1) the camera meets important speed and ergonomic constraints for clinical translation. In the following the specifics of the imec sensor are discussed in context of functional imaging.

A challenge in working with the camera are the second order peaks in the spectral bands (Fig. 7.18). The two peaks are mainly located in red (>580nm) and blue (<500nm). A theoretical analysis (Fig. 7.19) revealed a potential problem: reduction in oxygenation leads to a lower measurement in the red, but a higher measurement in the blue. By integrating over two peaks at the same time, the measured effect of oxygenation changes is small compared to measuring only a single peak. This theoretical result was confirmed by the *in silico* experiments in Chapter 7.4.4, where a theoretical system without the second peaks performed 1.7 percentage points better in oxygenation estimation. These results also show that the eight Spectrocam bands, which were selected with the method from the last chapter, perform better than even the idealized 16 band setup. This establishes further support for the band selection results.

Another problem caused by the two peaks is that the system is more sensitive to changes in the optical system during the intervention. While *in silico* experiments in Chapter 7.4.4 showed that small changes like exchanging the laparoscope lens can be neglected, switching from Xenon to Halogen light leads to worse results. The previously used Spectrocam system remains unaffected. The likely reason being, that if the response $f_j(\lambda)$ of the filters is narrow as in the Spectrocam, changes in lighting $l(\lambda)$, optical system $o(\lambda)$ and spectral reflectance $r(\lambda)$ are close to constant. They are therefore not expected to have strong impact on band reflectance r_j:

$$r_j = \frac{\int_{\lambda_{\min}}^{\lambda_{\max}} o(\lambda) l(\lambda) f_j(\lambda) r(\lambda)\,\mathrm{d}\lambda}{\int_{\lambda_{\min}}^{\lambda_{\max}} o(\lambda) l(\lambda) f_j(\lambda)\,\mathrm{d}\lambda} \approx \frac{\int_{\lambda_{\min}}^{\lambda_{\max}} \alpha \cdot f_j(\lambda) r(\lambda)\,\mathrm{d}\lambda}{\int_{\lambda_{\min}}^{\lambda_{\max}} \alpha \cdot f_j(\lambda)\,\mathrm{d}\lambda} \approx \frac{\alpha \int_{\lambda_{\min}}^{\lambda_{\max}} f_j(\lambda) \beta\,\mathrm{d}\lambda}{\alpha \int_{\lambda_{\min}}^{\lambda_{\max}} f_j(\lambda)\,\mathrm{d}\lambda} = \beta. \quad (7.3)$$

The broad second order peaks of the imec chip make this approximation not applicable. This could complicate clinical use of the sensor, because a switch in light source would require retraining the algorithm. One straightforward solution would be to precompute several regressors for common light sources. The white balance could be used to detect the most likely light source and switch to the respective regressor.

Because the sensor acquires images through the 4x4 mosaic pattern, the resolution of the image is low compared to conventional HD laparoscopic cameras. The low resolution functional image is probably sufficient for clinical practice. However, the RGB image needs to be available at HD resolutions. Several solutions can be imagined:

1. The multispectral laparoscope can be a standalone system, only providing functional information. This would require switching of laparoscopes during the intervention and thus severely interfere with the standard workflow.
2. The RGB image estimated from the MSI can be interpolated to higher resolutions. Relying on the MSI for RGB estimation alone is dangerous, because the MSI does not capture the full spectral information (e.g. due to a bandpass filter, the imec sensor is restricted to the 450nm to 650nm spectral range).
3. Similarily to 3D laparoscopes, a two chip (MSI + RGB) solution can be envisioned. The image would be taken from the currently active sensor. This would lead to light lost, due to a prism that splits light before the two sensors. However, this might be a practical solution, especially if further developments of the imec sensor are more quantum efficient.

Another theoretical limitation of the sensor is the spatial extent of the mosaics. Changes in tissue within one mosaic could lead to misaligned spectra. In the investigated case of kidney perfusion monitoring this was not viewed as critical, because the organ surface is relatively homogeneous and the vessels could be delineated clearly. However, this high resolution might be a requirement for other tasks, like searching for small spatial changes in metabolism, as caused by micro-metastases. Interpolation to higher spatial resolution might be done within the CNN framework discussed in the next paragraph.

Analysis: CNN regression. Random forests were exchanged with CNNs, mainly because of their inherent GPU acceleration, capability of end-to-end-processing and the innate ability to incorporate spatial regularization by analyzing the pixel neighborhood. The CNNs can perform the regression on the whole image in 130Hz, compared to 7Hz of the random forest (RF) regression used in Chapter 4. During this time the CNN additionally performs white/dark balancing, normalization and computes an RGB estimate.

The RF method and the conventional Beer-Lambert (BL) were evaluated in-patient to serve as a qualitative baseline for the CNN approach (Section 7.4.2). Because BL cannot estimate blood volume fraction, only oxygenation was considered. While it seemed all methods can estimate the general trend, the RF and especially the BL produced very noisy estimates (see Figure 7.12). Probably due to this reason they miss the initial oxygenation drop within the first minute after applying the clamps in the ROI. Qualitative analysis of the example images show that the CNN method can delineate small vessels,

despite delivering much smoother results. To be applicable, the broad filter responses have to be transformed to narrow virtual bands for BL. This is an additional source of error and is not necessary for the presented estimation method, which directly learns dependencies from simulated band reflectances. The BL repeatedly estimates physically impossible oxygenation values higher than 100%. These impossible values could be remedied by a non-negativity constraint (Section 3.3.2). Such constraints prevent a closed form solution of the problem and thus fast estimation, the strongest asset of BL regression.

Instead of evaluating the functional parameters for each pixel individually, a small region of interest was taken into account (Chapter 7.2). The simulated homogeneous tissue patches were stained randomly with pixels from another tissue, to account for changing tissue states within the covered neighborhood. This technique was compared in Section 7.4.2 to conventional pixel-wise regression and regression without this staining procedure (see Figure 7.15). Incorporating neighborhood delivered stabler results. If staining was omitted, the CNNs seemed to overfit to the simulations, producing less plausible *in vivo* and more than twice as noisy results. The proposed, simple procedure can serve as a baseline for more involved, future approaches. One idea is to simulate a small patch with one base tissue and a second tissue overlayed as simple geometric shape, such as lines or circles. Using such simplified artificial structures to learn the CNN was successful in optic flow applications [41]. However, evaluating different patch simulation techniques needs more *in vivo* recordings, as *in silico* comparisons are not meaningful.

As shown in the comparison of stained and unstained simulation approaches, overfitting can be an issue. In this chapter, L2 regularization and the four layer network was chosen for *in vivo* application, opposed to the better performing, but more parameter-heavy eight layer network. Future architectures should explore a higher numbers of filters in combination with dropout regularization [171]. The principle issue remains, that overfitting to simulations cannot be spotted during the *in silico* network training. In future work elaborate phantoms could validate the CNN results. Furthermore, recordings from healthy organs can be used to check if the oxygenation results are within plausible limits.

The presented network architectures provide a foundation for future, more explorational studies. The Conv/ReLU blocks could be amended by residual connections [85]. More involved, learned demosaicing and denoising [65] could provide cleaner, high resolution imaging and remove the small spatial misalignment caused by the current demosaicing.

7.6. CONCLUSION

This chapter implemented important steps required for moving multispectral imaging from a research environment into a more application oriented phase. Both recording and analysis of the multispectral images is now live and fluent, giving immediate and precise feedback to the surgeon. The system successfully detected and monitored ischemia during partial nephrectomy. If supported by further studies, the complete and usable system presented here could enable the transfer of MSI into the clinical environment.

The most important future task is thus the evaluation of this system in a larger study. Future work should identify further areas of application in which multispectral imaging can provide benefits. Our next field of investigation is detection of otherwise often missed flat adenoma in colon cancer screening. Due to the broad applicability and the large potential clinical benefit of the presented functional estimation approach, I am confident the descendents of this system are an integral part of the next generation OR.

8

Concluding Thoughts

Unanfechtbare Wahrheiten gibt es überhaupt nicht, und wenn es welche gibt, so sind sie langweilig.

Dubslav von Stechlin in Theodor Fontane's "Der Stechlin"

When I first learned of multispectral functional imaging (MSI), the possibilities seemed endless: with this technique one might monitor ischemia, detect cancer and other diseases, help clinicians in identifying tumor margins or verify surgery success; surely, it seemed to me, the future of interventional care would be bright and multispectral. However, soon dark spots became apparent. While enlightening from a research perspective, the chance that an actual physician would want to use one of the systems presented in the literature were low. The systems were bulky, slow devices, recording images that were analyzed with basic, limited methods. My vision became to develop a system that clinicians would want to use. I imagined it to be lightweight and fast, backed up by powerful image processing to interpret the information.

The main achievement of this thesis is an image analysis method which premiers live (25Hz) calculation of functional information without resorting to restrictive simplifying assumptions. At the heart of this method lies an approach that combines sophisticated Monte Carlo based simulations of light-tissue interaction (Chapter 4) with convolutional neural networks (CNNs, Chapter 7), one of the most powerful machine learning techniques. The networks relate multispectral measurements to underlying functional parameters such as oxygenation. They employ simulations, which are generated offline (and thus pre-intervention), as examples to learn this complicated relationship. Once the relationship is established, the method enables rapid (7.5ms) determination of functional information from images acquired during surgery. Throughout this work, a strong focus was put on an accurate understanding of multispectral image formation. However, learning from simulations has to be done with care, as they never fully mimic reality. To more closely resemble real measurements the first transfer learning method in this field (Chapter 5) was presented. By incorporating *in vivo* measurements, it adapts masses of generic simulations of visceral tissue to the desired application, such as liver surgery. This technique is key to applying the parameter-hungry Monte Carlo methods to various domains. A V-model like validation chain investigated the simulation process, the correct understanding of image formation and the functional estimation performance. The results gained by employing *in silico* evaluations, tissue mimicking phantoms, lab measurements and a porcine study involving six pigs supported the hypothesis that rapid and accurate functional estimation can be learned from simulations.

The second main achievement is the incorporation of the analysis method in a novel device which matches the speed of analysis with a corresponding speed of image recording. This device is the first laparoscope to use a multispectral snapshot camera (Chapter 7). Its sensor records at several multi-peaked bands, which prevents use of common analysis approaches. By employing the CNN approach it was possible to use this camera for the first ever fluent display of oxygenation during an intervention with MSI. In these in-human-patient measurements, ischemia was successfully detected and monitored. By presenting a device which can record and analyze multispectral images at video rates, one of the main limitations of MSI was overcome.

One remaining challenge is the low resolution of the acquired images. By studying the hypothesis that most of the information is captured by few spectral bands (Chapter 6) a small but discriminative set of bands was identified. Equipping the off-the-shelf snapshot camera with these selected bands would be a step towards high resolution MSI.

Together, the in-patient results, the novel device and the powerful but rapid algorithm for analysis realized my initial vision of a truly usable system. Therefore the chance of further clinical translation of this system is high. For future researchers, this thesis sketches a map that guides towards successful application of MSI. On this map there are areas that are left for exploration; new minds will need to fill this uncharted territory:

Computer scientists have two choices when analyzing multispectral images. They can use physical models to derive the functional information or directly learn from training images, for example to spot a tumor. If MSI devices become more standardized, and thus more training data is available, the direct approach might become feasible. On this exciting path all the powerful tools provided by modern machine learning techniques would be usable. The presented functional approach could be extended to work outside of the controlled laparoscopic environment. Generalization of the method to arbitrary lighting scenarios would allow use in the vast applications of open surgery. Other researchers might investigate what else is hidden in the spectral signature. Can we go beyond oxygenation and blood volume fraction? Can we incorporate other optical effects like fluorescence or polarization in the presented method?

Physicists would be needed to develop devices capable of recording such multimodal information. Imagine a system recording both multispectral and fluorescence signals concurrently. The ability of such devices to monitor perfusion with MSI and visualize tumors with targeted fluorescence could unlock new clinical workflows. To evaluate these systems, biophysical experts are needed to create highly realistic tissue mimicking phantoms. Recent advances in bioprinting of 3D tissue [135] might close this in my opinion most important gap: the unavailability of realistic looking and properly controlled gold standards.

The medical doctors finally should test the nuts and bolts of the devices. Only they as end users can judge the benefit. This can only be determined in larger studies, which evaluate MSI in context of perfusion monitoring and early cancer detection and screening. Going further, MSI might complement non-invasive imaging modalities: intraoperative measurements from MSI could be fused with preoperative perfusion MRI measurements to identify tumor margins. Ultimately, the most exciting consequences of widespread MSI might be the unforeseen findings. What will we learn from visualizing oxygenation at a large scale?

With the combined ideas from the mentioned talents and the seal of approval from the physicians the medical companies will build the new functional devices. Afterwards we can take care of what really counts, the final and often forgotten member of the chain: **the patient**.

A

IMEC NOISE, LINEARITY AND IMAGING SYSTEM TRANSFORMATION

The linearity and noise of the Ximea camera were evaluated, two important parameters for working with the imec sensor. Further, it was tested how well the transformation from spectral reflectance to camera reflectance is understood with this sensor.

A.1. SENSOR NOISE AND LINEARITY

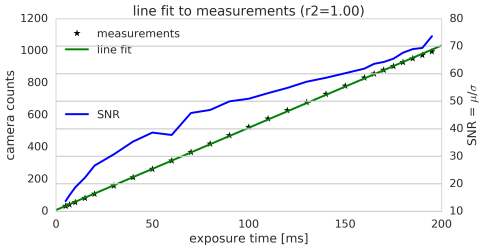
One prerequisite of the employed method is a linear camera response (see chapter 4.1.3). To check, the laparoscope was placed above a white tile and the integration time was varied to record images at differing intensity levels. One hundred recordings each were taken for each integration time to get a stable estimate for the mean value and the possibility to estimate the variance of the values.

Figure A.1a shows the line fit to the mean recordings (taken from a 10x10 central ROI on the images) and the SNR for different exposure times. Figure A.1b shows the worst fit of these recordings to a normal distribution, according to a D'Agostino and Pearson's based test[1]. The median p-value was 0.003%.

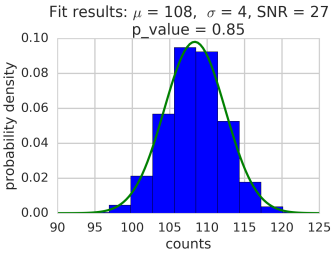
A.2. IMAGING SYSTEM

Light exiting from the light source is reflected on the tissue surface, passes through the laparoscope, the filter and finally reaches the camera, to be transformed into electric

[1] Tested with the scipy.stats.mstats.normaltest function of scipy v0.19



(a) Mean measurements, line fit and the SNR



(b) Worst normal fit. Blue measurement histogram, green the Gaussian induced by mean and standard deviation.

Figure A.1: (a) Estimating linearity for the dynamic range of the sensor. The line fits to the measurements is almost perfect, indicating the camera is linear. SNRs, determined from a 100 repeat measurements each, seem to grow approximately linear as well with camera brightness. (b) shows the measurements which fit worst to a normal distribution (20ms exposure time).

current. When simulating camera reflectance, all the elements on the path of light need to be considered.

To judge the overall understanding of the imaging system transformation, color tiles with known reflectance were measured with a spectrometer and the camera as in Chapter 4.3.3. For each color tile, spectrometer and camera measurements were transformed to camera reflectance. Figure A.2 shows an overview on these matching for all color tiles. The mean error of all color tile measurements is 0.2% and the standard deviation is 2.9%.

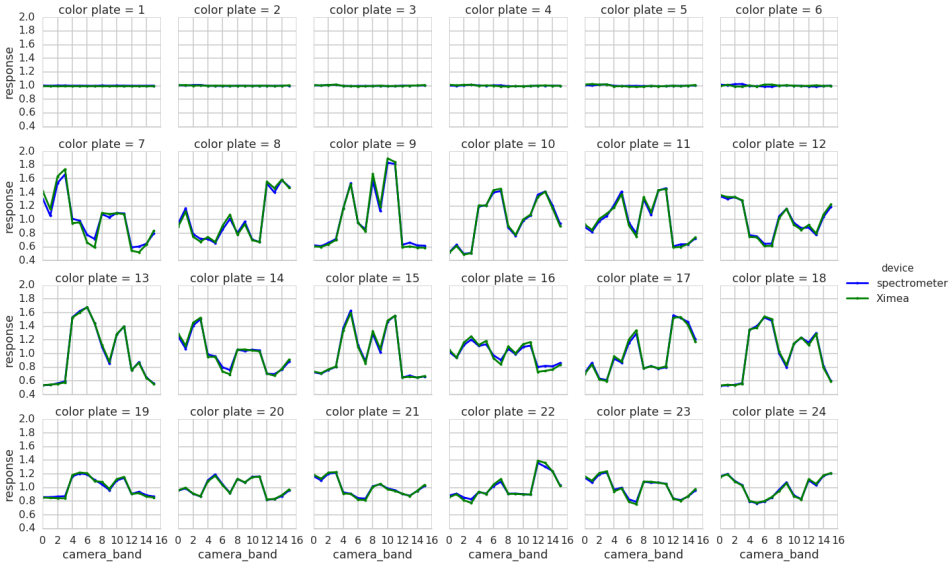


Figure A.2: Color tiles measured with spectrometer and camera. The artificial optical system was applied to the spectrometer measurements to transform them into camera reflectance space.

BIBLIOGRAPHY

REFERENCES

[1] A. P. Adams, J. O. Morgan-Hughes, and M. K. Sykes. pH and blood—gas analysis. *Anaesthesia*, 22(4):575–597 (1967).

[2] H. Akbari, L. V. Halig, D. M. Schuster, A. Osunkoya, V. Master, P. T. Nieh, G. Z. Chen, and B. Fei. Hyperspectral imaging and quantitative analysis for prostate cancer detection. *Journal of Biomedical Optics*, 17(7):0760051–07600510 (2012).

[3] H. Akbari and Y. Kosugi. Hyperspectral Imaging: a New Modality in Surgery. In G. R. Naik, editor, *Recent Advances in Biomedical Engineering*. InTech (2009). ISBN 978-953-307-004-9.

[4] H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka. Blood vessel detection and artery-vein differentiation using hyperspectral imaging. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 1461–1464. IEEE (2009).

[5] H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka. Cancer detection using infrared hyperspectral imaging. *Cancer Science*, 102(4):852–857 (2011).

[6] J. T. Alander, I. Kaartinen, A. Laakso, T. Pätilä, T. Spillmann, V. V. Tuchin, M. Venermo, and P. Välisuo. A review of indocyanine green fluorescent imaging in surgery. *International Journal of Biomedical Imaging*, 2012:940585 (2012).

[7] E. Alerstam, W. C. Yip Lo, T. D. Han, J. Rose, S. Andersson-Engels, and L. Lilge. Next-generation acceleration and code optimization for light transport in turbid media using GPUs. *Biomedical Optics Express*, 1(2):658–675 (2010).

[8] J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, T. Aso, E. Bagli, A. Bagulya, S. Banerjee, G. Barrand, et al. Recent developments in Geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186–225 (2016).

[9] W. S. Atkin, R. Edwards, I. Kralj-Hans, K. Wooldrage, A. R. Hart, J. M. A. Northover, D. M. Parkin, J. Wardle, S. W. Duffy, J. Cuzick, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet (London, England)*, 375(9726):1624–1633 (2010).

[10] R. Autorino, H. Zargar, W. M. White, G. Novara, F. Annino, S. Perdonà, M. De Angelis, A. Mottrie, F. Porpiglia, and J. H. Kaouk. Current Applications of Near-infrared

Fluorescence Imaging in Robotic Urologic Surgery: A Systematic Review and Critical Analysis of the Literature. *Urology*, 84(4):751–759 (2014).

[11] I. Barman, N. C. Dingari, N. Rajaram, J. W. Tunnell, R. R. Dasari, and M. S. Feld. Rapid and accurate determination of tissue optical properties using least-squares support vector machines. *Biomedical Optics Express*, 2(3):592–599 (2011).

[12] K. Basak, M. Manjunatha, and P. K. Dutta. Review of laser speckle-based analysis in medical imaging. *Medical & Biological Engineering & Computing*, 50(6):547–558 (2012).

[13] M. Beauchemin and K. B. Fung. On statistical band selection for image visualization. *PE & RS- Photogrammetric Engineering and Remote Sensing*, 67(5):571–574 (2001).

[14] N. Bedard, R. A. Schwarz, A. Hu, V. Bhattar, J. Howe, M. D. Williams, A. M. Gillenwater, R. Richards-Kortum, and T. S. Tkaczyk. Multimodal snapshot spectral imaging for oral cancer diagnostics: a pilot study. *Biomedical Optics Express*, 4(6):938 (2013).

[15] A. G. Bell. Upon the production and reproduction of sound by light. *Journal of the Society of Telegraph Engineers*, 9(34):404–426 (1880).

[16] S. L. Best, A. Thapa, N. Jackson, E. Olweny, M. Holzer, S. Park, E. Wehner, K. Zuzak, and J. A. Cadeddu. Renal Oxygenation Measurement During Partial Nephrectomy Using Hyperspectral Imaging May Predict Acute Postoperative Renal Function. *Journal of Endourology*, 27(8):1037–1040 (2013).

[17] T. Binzoni, A. Vogel, A. H. Gandjbakhche, and R. Marchesini. Detection limits of multi-spectral optical imaging under the skin surface. *Physics in Medicine and Biology*, 53(3):617–636 (2008).

[18] A. Bjorgan, M. Milanic, and L. L. Randeberg. Estimation of skin optical parameters for real-time hyperspectral imaging applications. *Journal of Biomedical Optics*, 19(6):066003–066003 (2014).

[19] D. A. Boas and A. K. Dunn. Laser speckle contrast imaging in biomedical optics. *Journal of Biomedical Optics*, 15(1) (2010).

[20] C. F. Bohren, E. E. Clothiaux, and D. R. Huffman. *Absorption and scattering of light by small particles*. Wiley-VCH, Weinheim, Bergstr, 2., vollst. überarb. u. aktualis. aufl. edition (2009). ISBN 978-3-527-40664-7.

[21] M. S. Borofsky, I. S. Gill, A. K. Hemal, T. P. Marien, I. Jayaratna, L. S. Krane, and M. D. Stifelman. Near-infrared fluorescence imaging to facilitate super-selective arterial clamping during zero-ischaemia robotic partial nephrectomy. *BJU international*, 111(4):604–610 (2013).

[22] K. Bousmalis, G. Trigeorgis, N. Silberman, D. Krishnan, and D. Erhan. Domain separation networks. In *Advances in Neural Information Processing Systems*, pages 343–351 (2016).

[23] L. Breiman. Random Forests. *Machine Learning*, 45(1):5–32 (2001).

[24] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, et al. Using Raman spectroscopy to characterize biological materials. *Nature Protocols*, 11(4):664–687 (2016).

[25] R. Byrd, P. Lu, J. Nocedal, and C. Zhu. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208 (1995).

[26] H. J. Byrne, M. Baranska, G. J. Puppels, N. Stone, B. Wood, K. M. Gough, P. Lasch, P. Heraud, J. Sulé-Suso, and G. D. Sockalingum. Spectropathology for the next generation: quo vadis? *The Analyst*, 140(7):2066–2073 (2015).

[27] C.-I. Chang and S. Wang. Constrained band selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6):1575–1585 (2006).

[28] Y.-W. Chen and S.-H. Tseng. Efficient construction of robust artificial neural networks for accurate determination of superficial sample optical properties. *Biomedical Optics Express*, 6(3):747 (2015).

[29] Z. Chen, A. Taflove, and V. Backman. Concept of the equiphase sphere for light scattering by nonspherical dielectric particles. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 21(1):88–97 (2004).

[30] N. T. Clancy, S. Arya, D. Stoyanov, M. Singh, G. B. Hanna, and D. S. Elson. Intraoperative measurement of bowel oxygen saturation using a multispectral imaging laparoscope. *Biomedical Optics Express*, 6(10):4179 (2015).

[31] N. T. Clancy, M. Ebner, J. S. Crane, R. Corbett, N. Duncan, C. Caro, and D. S. Elson. Multispectral imaging using a fast filter wheel system during vascular surgery. In *Proceedings of The Hamlyn Symposium on Medical Robotics*, pages 69–70 (2013).

[32] N. T. Clancy, S. Saso, D. Stoyanov, V. Sauvage, D. J. Corless, M. Boyd, D. E. Noakes, M.-Y. Thum, S. Ghaem-Maghami, J. R. Smith, et al. Multispectral imaging of organ viability during uterine transplantation surgery in rabbits and sheep. *Journal of Biomedical Optics*, 21(10):106006 (2016).

[33] N. T. Clancy, V. Sauvage, S. Saso, D. Stoyanov, D. J. Corless, M. Boyd, D. Noakes, G.-Z. Yang, S. Ghaem-Maghami, J. R. Smith, et al. Registration and Analysis of Multispectral Images Acquired During Uterine Transplantation Surgery. In *Biomedical Optics and 3-D Imaging*, page BSu3A.73. Optical Society of America (2012).

[34] N. T. Clancy, D. Stoyanov, D. R. C. James, A. D. Marco, V. Sauvage, J. Clark, G.-Z. Yang, and D. S. Elson. Multispectral image alignment using a three channel endoscope in vivo during minimally invasive surgery. *Biomedical Optics Express*, 3(10):2567–2578 (2012).

[35] E. Claridge, S. Cotton, P. Hall, and M. Moncrieff. From colour to tissue histology: physics-based interpretation of images of pigmented skin lesions. *Medical Image Analysis*, 7(4):489–502 (2003).

[36] E. Claridge and D. Hidovic-Rowe. Model based inversion for deriving maps of histological parameters characteristic of cancer from ex-vivo multispectral images of the colon. *IEEE Trans Med Imaging* (2013).

[37] J.-A. Collins, A. Rudenski, J. Gibson, L. Howard, and R. O'Driscoll. Relating oxygen partial pressure, saturation and content: the haemoglobin-oxygen dissociation curve. *Breathe (Sheffield, England)*, 11(3):194–201 (2015).

[38] D. Cosgrove and N. Lassau. Imaging of perfusion using ultrasound. *European Journal of Nuclear Medicine and Molecular Imaging*, 37(S1):65–85 (2010).

[39] D. J. Cuccia, F. Bevilacqua, A. J. Durkin, F. R. Ayers, and B. J. Tromberg. Quantitation and mapping of tissue optical properties using modulated imaging. *Journal of Biomedical Optics*, 14(2):024012 (2009).

[40] R. S. Dacosta, B. C. Wilson, and N. E. Marcon. New optical technologies for earlier endoscopic diagnosis of premalignant gastrointestinal lesions. *Journal of gastroenterology and hepatology*, 17(s1) (2002).

[41] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. van der Smagt, D. Cremers, and T. Brox. Flownet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2758–2766 (2015).

[42] 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*, 34(1):28–36 (2001).

[43] Z. Du, M. K. Jeong, and S. G. Kong. Band Selection of Hyperspectral Images for Automatic Detection of Poultry Skin Tumors. *IEEE Transactions on Automation Science and Engineering*, 4(3):332–339 (2007).

[44] A. K. Dunn. Laser Speckle Contrast Imaging of Cerebral Blood Flow. *Annals of Biomedical Engineering*, 40(2):367–377 (2012).

[45] K. L. Eales, K. E. R. Hollinshead, and D. A. Tennant. Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis*, 5(1):e190 (2016).

[46] W. Egan. *Optical properties of Inhomogeneous materials: Applications to geology, astronomy chemistry, and engineering*. Elsevier (2012). ISBN 978-0-323-15914-2.

[47] H. K. Eltzschig and T. Eckle. Ischemia and reperfusion—from mechanism to translation. *Nature medicine*, 17(11) (2011).

[48] M. Ewerloef, M. Larsson, and E. G. Salerud. Spatial and temporal skin blood volume and saturation estimation using a multispectral snapshot imaging camera. In *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XV*, volume 10068, page 1006814. International Society for Optics and Photonics (2017).

[49] H. Fabelo, S. Ortega, R. Guerra, G. Callicó, A. Szolna, J. F. Piñeiro, M. Tejedor, S. López, and R. Sarmiento. A Novel Use of Hyperspectral Images for Human Brain Cancer Detection Using in-Vivo Samples. In *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies*, BIOSTEC 2016, pages 311–320. SCITEPRESS - Science and Technology Publications, Lda, Portugal (2016). ISBN 978-989-758-170-0.

[50] T. J. Farrell, M. S. Patterson, and B. Wilson. A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo. *Medical Physics*, 19(4):879–888 (1992).

[51] Y. Fawzy, S. Lam, and H. Zeng. Rapid multispectral endoscopic imaging system for near real-time mapping of the mucosa blood supply in the lung. *Biomedical Optics Express*, 6(8):2980 (2015).

[52] Y.-Z. Feng and D.-W. Sun. Application of hyperspectral imaging in food safety inspection and control: a review. *Critical Reviews in Food Science and Nutrition*, 52(11):1039–1058 (2012).

[53] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars. Unsupervised Visual Domain Adaptation Using Subspace Alignment. In *2013 IEEE International Conference on Computer Vision*, pages 2960–2967 (2013).

[54] L. K. Ferris and R. J. Harris. New Diagnostic Aids for Melanoma. *Dermatologic Clinics*, 30(3):535–545 (2012).

[55] C. Fischer and I. Kakoulli. Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*, 51(sup1):3–16 (2006).

[56] R. D. Frostig, S. A. Masino, M. C. Kwon, and C. H. Chen. Using light to probe the brain: Intrinsic signal optical imaging. *International Journal of Imaging Systems and Technology*, 6(2-3):216–224 (1995).

[57] J. Galeano, R. Jolivot, Y. Benezeth, F. Marzani, J. F. Emile, and D. Lamarque. Analysis of Multispectral Images of Excised Colon Tissue Samples Based on Genetic Algorithms. In *SITIS*, pages 833–838. IEEE (2012). ISBN 978-1-4673-5152-2.

[58] J. Galeano, S. Perez, Y. Montoya, D. Botina, and J. Garzón. Blind source separation of ex-vivo aorta tissue multispectral images. *Biomedical Optics Express*, 6(5):1589 (2015).

[59] G. Gandaglia, P. Schatteman, G. De Naeyer, F. D'Hondt, and A. Mottrie. Novel Technologies in Urologic Surgery: a Rapidly Changing Scenario. *Current Urology Reports*, 17(3):19 (2016).

[60] Y. Ganin and V. Lempitsky. Unsupervised Domain Adaptation by Backpropagation. *arXiv:1409.7495 [cs, stat]* (2014).

[61] L. Gao, R. T. Kester, and T. S. Tkaczyk. Compact Image Slicing Spectrometer (ISS) for hyperspectral fluorescence microscopy. *Optics express*, 17(15):12293–12308 (2009).

[62] L. Gao and R. T. Smith. Optical hyperspectral imaging in microscopy and spectroscopy – a review of data acquisition. *Journal of Biophotonics*, 8(6):441–456 (2015).

[63] B. Geelen, C. Blanch, P. Gonzalez, N. Tack, and A. Lambrechts. A tiny VIS-NIR snapshot multispectral camera. In *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII*, volume 9374, page 937414. International Society for Optics and Photonics (2015).

[64] E. A. Genina and V. V. Tuchin. *Handbook of biomedical optics*. CRC press (2011).

[65] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand. Deep Joint Demosaicking and Denoising. *ACM Trans. Graph.*, 35(6):191:1–191:12 (2016).

[66] N. Ghosh and I. A. Vitkin. Tissue polarimetry: concepts, challenges, applications, and outlook. *Journal of Biomedical Optics*, 16(11):110801 (2011).

[67] I. S. Gill, M. B. Patil, A. L. de Castro Abreu, C. Ng, J. Cai, A. Berger, M. S. Eisenberg, M. Nakamoto, O. Ukimura, A. C. Goh, et al. Zero Ischemia Anatomical Partial Nephrectomy: A Novel Approach. *The Journal of Urology*, 187(3):807–815 (2012).

[68] S. Gioux, A. Mazhar, B. T. Lee, S. J. Lin, A. M. Tobias, D. J. Cuccia, A. Stockdale, R. Oketokoun, Y. Ashitate, E. Kelly, et al. First-in-human pilot study of a spatial frequency domain oxygenation imaging system. *Journal of Biomedical Optics*, 16(8):086015 (2011).

[69] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In *PMLR*, pages 249–256 (2010).

[70] G. H. Glover. Overview of Functional Magnetic Resonance Imaging. *Neurosurgery clinics of North America*, 22(2):133–139 (2011).

[71] A. F. Goetz, G. Vane, J. E. Solomon, and B. N. Rock. Imaging spectrometry for Earth remote sensing. *Science (New York, N.Y.)*, 228(4704):1147–1153 (1985).

[72] M. Goetz, C. Weber, F. Binczyk, J. Polanska, R. Tarnawski, B. Bobek-Billewicz, U. Koethe, J. Kleesiek, B. Stieltjes, and K. H. Maier-Hein. DALSA: Domain Adaptation for Supervised Learning From Sparsely Annotated MR Images. *IEEE Transactions on Medical Imaging*, 35(1):184–196 (2016).

[73] B. Gong, K. Grauman, and F. Sha. Connecting the Dots with Landmarks: Discriminatively Learning Domain-Invariant Features for Unsupervised Domain Adaptation. In *ICML (1)*, pages 222–230 (2013).

[74] A. Goto, J. Nishikawa, S. Kiyotoki, M. Nakamura, J. Nishimura, T. Okamoto, H. Ogihara, Y. Fujita, Y. Hamamoto, and I. Sakaida. Use of hyperspectral imaging technology to develop a diagnostic support system for gastric cancer. *Journal of Biomedical Optics*, 20(1):016017–016017 (2015).

[75] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schölkopf. Covariate shift by kernel mean matching. *Dataset shift in machine learning*, 3(4):5 (2009).

[76] M. D. Grossberg and S. K. Nayar. What is the space of camera response functions? In *2003 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages II–602. IEEE (2003).

[77] X. Gu, Z. Han, L. Yao, Y. Zhong, Q. Shi, Y. Fu, C. Liu, X. Wang, and T. Xie. Image enhancement based on *in vivo* hyperspectral gastroscopic images: a case study. *Journal of Biomedical Optics*, 21(10):101412 (2016).

[78] B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson. Band Selection for Hyperspectral Image Classification Using Mutual Information. *IEEE Geoscience and Remote Sensing Letters*, 3(4):522–526 (2006).

[79] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157–1182 (2003).

[80] N. Hagen and M. W. Kudenov. Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9):090901–090901 (2013).

[81] M. Halicek, G. Lu, J. V. Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen, and B. Fei. Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging. *Journal of Biomedical Optics*, 22(6):060503–060503 (2017).

[82] Z. Han, A. Zhang, X. Wang, Z. Sun, M. D. Wang, and T. Xie. In vivo use of hyperspectral imaging to develop a noncontact endoscopic diagnosis support system for malignant colorectal tumors. *Journal of Biomedical Optics*, 21(1):016001–016001 (2016).

[83] D. Hanahan and R. A. Weinberg. Hallmarks of Cancer: The Next Generation. *Cell*, 144(5):646–674 (2011).

[84] D. Hattery, M. Hassan, S. Demos, and A. Gandjbakhche. Hyperspectral imaging of Kaposi's Sarcoma for disease assessment and treatment monitoring. In *Applied Imagery Pattern Recognition Workshop, 2002. Proceedings.*, pages 124–130 (2002).

[85] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778 (2016).

[86] G. E. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(3):267–276 (1994).

[87] T. Heimann, P. Mountney, M. John, and R. Ionasec. Real-time ultrasound transducer localization in fluoroscopy images by transfer learning from synthetic training data. *Medical Image Analysis*, 18(8):1320–1328 (2014).

[88] R. Hennessy, S. L. Lim, M. K. Markey, and J. W. Tunnell. Monte Carlo lookup table-based inverse model for extracting optical properties from tissue-simulating phantoms using diffuse reflectance spectroscopy. *Journal of Biomedical Optics*, 18(3):037003–037003 (2013).

[89] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the Galaxy. *The Astrophysical Journal*, 93:70–83 (1941).

[90] D. Heresbach, T. Barrioz, M. G. Lapalus, D. Coumaros, P. Bauret, P. Potier, D. Sautereau, C. Boustiere, J. C. Grimaud, C. Barthelemy, et al. Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. *Endoscopy*, 40(4):284–290 (2008).

[91] D. Hidovic-Rowe and E. Claridge. Modelling and validation of spectral reflectance for the colon. *Phys Med Biol*, 50(6):1071–1093 (2005).

[92] E. M. C. Hillman. Optical brain imaging in vivo: techniques and applications from animal to man. *Journal of Biomedical Optics*, 12(5):051402 (2007).

[93] M. Hohmann, R. Kanawade, F. Klämpfl, A. Douplik, J. Mudter, M. F. Neurath, and H. Albrecht. *In*-vivo multispectral video endoscopy towards *in-vivo* hyperspectral video endoscopy. *Journal of Biophotonics*, 10(4):553–564 (2017).

[94] M. S. Holzer, S. L. Best, N. Jackson, A. Thapa, G. V. Raj, J. A. Cadeddu, and K. J. Zuzak. Assessment of Renal Oxygenation During Partial Nephrectomy Using Hyperspectral Imaging. *The Journal of Urology*, 186(2):400–404 (2011).

[95] J. Huang, A. J. Smola, A. Gretton, K. M. Borgwardt, B. Schölkopf, and others. Correcting sample selection bias by unlabeled data. *Advances in neural information processing systems*, 19:601 (2007).

[96] S. Hwang, J. Park, N. Kim, Y. Choi, and I. S. Kweon. Multispectral pedestrian detection: Benchmark dataset and baseline. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1037–1045 (2015).

[97] Itseez. Open Source Computer Vision Library. `https://github.com/itseez/opencv` (2015). Accessed: 2017-08-21.

[98] S. L. Jacques. Optical properties of biological tissues: a review. *Physics in Medicine and Biology*, 58(11):R37–R61 (2013).

[99] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional Architecture for Fast Feature Embedding. *arXiv:1408.5093 [cs]* (2014).

[100] R. Jolivot, P. Vabres, and F. Marzani. Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system. *Computerized Medical Imaging and Graphics*, 35(2):85–88 (2011).

[101] E. Jones, T. Oliphant, P. Peterson, and others. SciPy: Open source scientific tools for Python. http://www.scipy.org/ (2001). Accessed: 2017-08-21.

[102] G. Jones, N. Clancy, Y. Helo, S. Arridge, D. Elson, and D. Stoyanov. Bayesian Estimation of Intrinsic Tissue Oxygenation and Perfusion from RGB Images. *IEEE Transactions on Medical Imaging*, pages 1–1 (2017).

[103] G. Jones, N. T. Clancy, S. Arridge, D. S. Elson, and D. Stoyanov. Inference of Tissue Haemoglobin Concentration from Stereo RGB. In *International Conference on Medical Imaging and Virtual Reality*, pages 50–58. Springer (2016).

[104] A. Jubran. Pulse oximetry. *Critical Care*, 3(2):R11–R17 (1999).

[105] K. Kaneko, H. Yamaguchi, T. Saito, T. Yano, Y. Oono, H. Ikematsu, S. Nomura, A. Sato, M. Kojima, H. Esumi, et al. Hypoxia Imaging Endoscopy Equipped with Laser Light Source from Preclinical Live Animal Study to First-In-Human Subject Research. *PLoS ONE*, 9(6):e99055 (2014).

[106] A. Karliczek, N. J. Harlaar, C. J. Zeebregts, T. Wiggers, P. C. Baas, and G. M. van Dam. Surgeons lack predictive accuracy for anastomotic leakage in gastrointestinal surgery. *International Journal of Colorectal Disease*, 24(5):569–576 (2009).

[107] J. Katrarnik, F. Pernur, and B. Likar. Radiometric calibration and noise estimation of acousto-optic tunable filter hyperspectral imaging systems. *Applied Optics*, 52(15):3526–3537 (2013).

[108] W. E. Kelley. The Evolution of Laparoscopy and the Revolution in Surgery in the Decade of the 1990s. *JSLS : Journal of the Society of Laparoendoscopic Surgeons*, 12(4):351–357 (2008).

[109] R. T. Kester, N. Bedard, L. Gao, and T. S. Tkaczyk. Real-time snapshot hyperspectral imaging endoscope. *Journal of Biomedical Optics*, 16(5):056005 (2011).

[110] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[111] L. Kish. *Survey Sampling*. Wiley-Interscience, New York, revised edition edition (1995). ISBN 978-0-471-10949-5.

[112] F. Knieling, C. Neufert, A. Hartmann, J. Claussen, A. Urich, C. Egger, M. Vetter, S. Fischer, L. Pfeifer, A. Hagel, et al. Multispectral Optoacoustic Tomography for Assessment of Crohn's Disease Activity. *New England Journal of Medicine*, 376(13):1292–1294 (2017).

[113] A. Krogh. On the Mechanism of the Gas-Exchange in the Lungs1. *Skandinavisches Archiv Für Physiologie*, 23(1):248–278 (1910).

[114] J. Kuula, I. Pölönen, H.-H. Puupponen, T. Selander, T. Reinikainen, T. Kalenius, and H. Saari. Using VIS/NIR and IR spectral cameras for detecting and separating crime scene details. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI*, volume 8359, page 83590P. International Society for Optics and Photonics (2012).

[115] J. H. Lambert. *Lamberts Photometrie: hft. 1.th. Das directe licht. 2.th. Die schwächung des lichts durch durchsichtige körper, besonders durch glas*. W. Engelmann (1892).

[116] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324 (1998).

[117] R. Leitner, M. D. Biasio, T. Arnold, C. V. Dinh, M. Loog, and R. P. W. Duin. Multispectral video endoscopy system for the detection of cancerous tissue. *Pattern Recognition Letters*, 34(1):85 – 93 (2013).

[118] Q. Li, X. He, Y. Wang, H. Liu, D. Xu, and F. Guo. Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10):100901 (2013).

[119] H. Liang. Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A*, 106(2):309–323 (2012).

[120] H.-T. Lim and V. M. Murukeshan. A four-dimensional snapshot hyperspectral video-endoscope for bio-imaging applications. *Scientific Reports*, 6:24044 (2016).

[121] G. S. Litynski. Laparoscopy-The Early Attempts: Spotlighting Georg Kelling and Hans Christian Jacobaeus. *JSLS: Journal of the Society of Laparoendoscopic Surgeons*, 1(1):83 (1997).

[122] Z. Liu, H. Wang, and Q. Li. Tongue Tumor Detection in Medical Hyperspectral Images. *Sensors (Basel, Switzerland)*, 12(1):162–174 (2011).

[123] G. Lu and B. Fei. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):10901 (2014).

[124] G. Lu, L. Halig, D. Wang, X. Qin, Z. G. Chen, and B. Fei. Spectral-spatial classification for noninvasive cancer detection using hyperspectral imaging. *Journal of Biomedical Optics*, 19(10):106004–106004 (2014).

[125] G. Lu, D. Wang, X. Qin, L. Halig, S. Muller, H. Zhang, A. Chen, B. W. Pogue, Z. G. Chen, and B. Fei. Framework for hyperspectral image processing and quantification for cancer detection during animal tumor surgery. *Journal of Biomedical Optics*, 20(12):126012–126012 (2015).

[126] S. A. Maadeed, S. Kunhoth, A. Bouridane, and R. Peyret. Multispectral imaging and machine learning for automated cancer diagnosis. In *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1740–1744 (2017).

[127] A. Mansouri, F. S. Marzani, and P. Gouton. Development of a protocol for CCD calibration: application to a multispectral imaging system. *International Journal of Robotics and Automation*, 20(2):94–100 (2005).

[128] A. Mansouri, F. S. Marzani, and P. Gouton. Neural networks in two cascade algorithms for spectral reflectance reconstruction. In *IEEE International Conference on Image Processing, 2005. ICIP*, volume 2, pages II–718. IEEE (2005).

[129] S. E. Martinez-Herrera, Y. Benezeth, M. Boffety, J.-F. Emile, F. Marzani, D. Lamarque, and F. Goudail. Multispectral Endoscopy to Identify Precancerous Lesions in Gastric Mucosa. In *Image and Signal Processing*, Lecture Notes in Computer Science, pages 43–51. Springer, Cham (2014). ISBN 978-3-319-07997-4 978-3-319-07998-1.

[130] K. Matthes. Untersuchungen über die Sauerstoffsättigung des menschlichen Arterienblutes. *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie*, 179(6):698–711 (1935).

[131] T. R. McClintock, M. A. Bjurlin, J. S. Wysock, M. S. Borofsky, T. P. Marien, C. Okoro, and M. D. Stifelman. Can selective arterial clamping with fluorescence imaging preserve kidney function during robotic partial nephrectomy? *Urology*, 84(2):327–332 (2014).

[132] G. Mie. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen der Physik*, 330(3):377–445 (1908).

[133] S. Moccia, S. J. Wirkert, H. Kenngott, A. S. Vemuri, M. Apitz, B. Mayer, E. De Momi, L. S. Mattos, and L. Maier-Hein. Uncertainty-Aware Organ Classification for Surgical Data Science Applications in Laparoscopy. *arXiv preprint arXiv:1706.07002* (2017).

[134] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of Machine Learning*. The MIT Press (2012). ISBN 978-0-262-01825-8.

[135] S. V. Murphy and A. Atala. 3d bioprinting of tissues and organs. *Nature Biotechnology*, 32(8):nbt.2958 (2014).

[136] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814 (2010).

[137] B. S. Nichols, N. Rajaram, and J. W. Tunnell. Performance of a lookup table-based approach for measuring tissue optical properties with diffuse optical spectroscopy. *Journal of Biomedical Optics*, 17(5):0570011–0570018 (2012).

[138] S. P. Nighswander-Rempel, R. A. Shaw, V. V. Kupriyanov, J. Rendell, B. Xiang, and H. H. Mantsch. Mapping tissue oxygenation in the beating heart with near-infrared spectroscopic imaging. *Vibrational Spectroscopy*, 32(1):85 – 94 (2003).

[139] I. Nishidate, T. Ishizuka, K. Yoshida, S. Kawauchi, S. Sato, and M. Sato. Visualization of hemodynamics and light scattering in exposed brain of rat using multispectral image reconstruction based on Wiener estimation method. In *Advanced Microscopy Techniques IV; and Neurophotonics II*, volume 9536, page 95360Z. International Society for Optics and Photonics (2015).

[140] I. Nishidate, T. Maeda, K. Niizeki, and Y. Aizu. Estimation of Melanin and Hemoglobin Using Spectral Reflectance Images Reconstructed from a Digital RGB Image by the Wiener Estimation Method. *Sensors*, 13(6):7902–7915 (2013).

[141] D. Nouri, Y. Lucas, and S. Treuillet. Efficient tissue discrimination during surgical interventions using hyperspectral imaging. In *International Conference on Information Processing in Computer-Assisted Interventions*, pages 266–275. Springer (2014).

[142] D. Nouri, Y. Lucas, and S. Treuillet. Hyperspectral interventional imaging for enhanced tissue visualization and discrimination combining band selection methods. *International Journal of Computer Assisted Radiology and Surgery*, 11(12):2185–2197 (2016).

[143] V. Ntziachristos. Going deeper than microscopy: the optical imaging frontier in biology. *Nature Methods*, 7(8):603–614 (2010).

[144] O. J. Old, L. M. Fullwood, R. Scott, G. R. Lloyd, L. M. Almond, N. A. Shepherd, N. Stone, H. Barr, and C. Kendall. Vibrational spectroscopy for cancer diagnostics. *Analytical Methods*, 6(12):3901 (2014).

[145] E. O. Olweny, S. Faddegon, S. L. Best, N. Jackson, E. F. Wehner, Y. K. Tan, K. J. Zuzak, and J. A. Cadeddu. First Place: Renal Oxygenation During Robot-Assisted Laparoscopic Partial Nephrectomy: Characterization Using Laparoscopic Digital Light Processing Hyperspectral Imaging. *Journal of Endourology*, 27(3):265–269 (2013).

[146] G. M. Palmer and N. Ramanujam. Monte Carlo-based inverse model for calculating tissue optical properties. Part I: Theory and validation on synthetic phantoms. *Applied optics*, 45(5):1062–1071 (2006).

[147] S. J. Pan and Q. Yang. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359 (2010).

[148] A. Paul, A. Dey, D. P. Mukherjee, J. Sivaswamy, and V. Tourani. Regenerative Random Forest with Automatic Feature Selection to Detect Mitosis in Histopathological Breast Cancer Images. In N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, number 9350 in Lecture Notes in Computer Science, pages 94–102. Springer International Publishing (2015). ISBN 978-3-319-24570-6 978-3-319-24571-3.

[149] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830 (2011).

[150] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238 (2005).

[151] B. W. Pogue and M. S. Patterson. Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry. *Journal of Biomedical Optics*, 11(4):041102 (2006).

[152] S. Prahl. Optical Absorption of Hemoglobin. `http://omlc.org/spectra/hemoglobin/` (1999). Accessed: 2017-08-21.

[153] S. J. Preece and E. Claridge. Spectral filter optimization for the recovery of parameters which describe human skin. *IEEE Trans Pattern Anal Mach Intell*, 26(7):913–922 (2004).

[154] J. Qi and D. S. Elson. A high definition Mueller polarimetric endoscope for tissue characterisation. *Scientific Reports*, 6:25953 (2016).

[155] A. Rahimi, B. Recht, and others. Random Features for Large-Scale Kernel Machines. In *NIPS*, volume 3, page 5 (2007).

[156] D. Ramsoekh, J. Haringsma, J. W. Poley, P. van Putten, H. van Dekken, E. W. Steyerberg, M. E. van Leerdam, and E. J. Kuipers. A back-to-back comparison of white light video endoscopy with autofluorescence endoscopy for adenoma detection in high-risk subjects. *Gut*, 59(6):785–793 (2010).

[157] L. L. Randeberg, E. L. P. Larsen, and L. O. Svaasand. Characterization of vascular structures and skin bruises using hyperspectral imaging, image analysis and diffusion theory. *Journal of Biophotonics*, 3(1-2):53–65 (2009).

[158] L. L. Randeberg, E. L. P. Larsen, and L. O. Svaasand. Hyperspectral imaging of blood perfusion and chromophore distribution in skin. In *Photonic Therapeutics and Diagnostics V*, volume 7161, page 71610C. International Society for Optics and Photonics (2009).

[159] S. J. Reddi, B. Poczos, and A. Smola. Doubly Robust Covariate Shift Correction. In *Twenty-Ninth AAAI Conference on Artificial Intelligence* (2015).

[160] G. S. d. Santos, E. Maneas, D. Nikitichev, A. Barburas, A. L. David, J. Deprest, A. Desjardins, T. Vercauteren, and S. Ourselin. A Registration Approach to Endoscopic Laser Speckle Contrast Imaging for Intrauterine Visualisation of Placental Vessels. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pages 455–462. Springer, Cham (2015). ISBN 978-3-319-24552-2 978-3-319-24553-9.

[161] E. Sarhrouni, A. Hammouch, and D. Aboutajdine. Dimensionality Reduction and Classification Feature Using Mutual Information Applied to Hyperspectral Images: A Wrapper Strategy Algorithm Based on Minimizing the Error Probability Using the Inequality of Fano. *arXiv:1211.0055 [cs]* (2012).

[162] A. Sassaroli and S. Fantini. Comment on the modified Beer-Lambert law for scattering media. *Phys Med Biol*, 49(14):N255–257 (2004).

[163] Schools. *Human Biology and Health*. Pearson Prentice Hall, Englewood Cliffs, N.J. (1993). ISBN 978-0-13-981176-0.

[164] I. Seo, C. K. Hayakawa, and V. Venugopalan. Radiative transport in the delta-P1 approximation for semi-infinite turbid media. *Medical physics*, 35(2):681–693 (2008).

[165] J. W. Severinghaus. The invention and development of blood gas analysis apparatus. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 97(1):253–256 (2002).

[166] M. Sharma, R. Hennessy, M. K. Markey, and J. W. Tunnell. Verification of a two-layer inverse Monte Carlo absorption model using multiple source-detector separation diffuse reflectance spectroscopy. *Biomedical Optics Express*, 5(1):40–53 (2013).

[167] C. Sheffield. Selecting Band Combinations from Multispectral Data. In *Photogrammetric Engineering and Remote Sensing* (1985).

[168] S. A. Skinner, G. M. Frydman, and P. E. O'Brien. Microvascular structure of benign and malignant tumors of the colon in humans. *Dig. Dis. Sci.*, 40(2):373–384 (1995).

[169] S. A. Skinner and P. E. O'Brien. The microvascular structure of the normal colon in rats and humans. *The Journal of Surgical Research*, 61(2):482–490 (1996).

[170] B. S. Sorg, B. J. Moeller, O. Donovan, Y. Cao, and M. W. Dewhirst. Hyperspectral imaging of hemoglobin saturation in tumor microvasculature and tumor hypoxia development. *Journal of Biomedical Optics*, 10(4):044004 (2005).

[171] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958 (2014).

[172] staff Blausen.com. Medical gallery of blausen medical 2014. *WikiJournal of Medicine*, 1(2) (2014). Accessed: 2017-08-21.

[173] M. M. Stephen, J. L. Jayanthi, N. G. Unni, P. E. Kolady, V. T. Beena, P. Jeemon, and N. Subhash. Diagnostic accuracy of diffuse reflectance imaging for early detection of pre-malignant and malignant changes in the oral cavity: a feasibility study. *BMC cancer*, 13(1):278 (2013).

[174] D. Stoyanov, A. Rayshubskiy, and E. Hillman. Robust registration of multispectral images of the cortical surface in neurosurgery. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1643–1646. IEEE (2012).

[175] I. B. Styles, A. Calcagni, E. Claridge, F. Orihuela-Espina, and J. M. Gibson. Quantitative analysis of multi-spectral fundus images. *Med Image Anal*, 10(4):578–597 (2006).

[176] M. Sugiyama and M. Kawanabe. *Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation*. The MIT Press (2012). ISBN 978-0-262-01709-1.

[177] M. Sugiyama, S. Nakajima, H. Kashima, P. V. Buenau, and M. Kawanabe. Direct importance estimation with model selection and its application to covariate shift adaptation. In *Advances in neural information processing systems*, pages 1433–1440 (2008).

[178] B. Sun, J. Feng, and K. Saenko. Return of Frustratingly Easy Domain Adaptation. In *AAAI*, volume 6, page 8 (2016).

[179] R. H. Thompson, I. Frank, C. M. Lohse, I. R. Saad, A. Fergany, H. Zincke, B. C. Leibovich, M. L. Blute, and A. C. Novick. The impact of ischemia time during open nephron sparing surgery on solitary kidneys: a multi-institutional study. *The Journal of Urology*, 177(2):471–476 (2007).

[180] H. J. Thomson, A. Busuttil, M. A. Eastwood, A. N. Smith, and R. A. Elton. The submucosa of the human colon. *Journal of Ultrastructure and Molecular Structure Research*, 96(1-3):22–30 (1986).

[181] S. Tobis, J. K. Knopf, C. R. Silvers, J. Marshall, A. Cardin, R. W. Wood, J. E. Reeder, E. Erturk, R. Madeb, J. Yao, et al. Near infrared fluorescence imaging after intravenous indocyanine green: initial clinical experience with open partial nephrectomy for renal cortical tumors. *Urology*, 79(4):958–964 (2012).

[182] C. R. Tracy, J. D. Terrell, R. P. Francis, E. F. Wehner, J. Smith, M. Litorja, D. L. Hawkins, M. S. Pearle, J. A. Cadeddu, and K. J. Zuzak. First Prize: Characterization of Renal Ischemia Using DLP® Hyperspectral Imaging: A Pilot Study Comparing Artery-Only Occlusion Versus Artery and Vein Occlusion. *Journal of endourology*, 24(3):321–325 (2010).

[183] C. Tsui, R. Klein, and M. Garabrant. Minimally invasive surgery: national trends in adoption and future directions for hospital strategy. *Surgical Endoscopy*, 27(7):2253–2257 (2013).

[184] P. K. Upputuri and M. Pramanik. Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review. *Journal of Biomedical Optics*, 22(4):041006 (2016).

[185] G. M. van Dam, G. Themelis, L. M. A. Crane, N. J. Harlaar, R. G. Pleijhuis, W. Kelder, A. Sarantopoulos, J. S. de Jong, H. J. G. Arts, A. G. J. van der Zee, et al. Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor-alpha targeting: first in-human results. *Nature Medicine*, 17(10):1315–1319 (2011).

[186] M. van de Giessen, J. P. Angelo, and S. Gioux. Real-time, profile-corrected single snapshot imaging of optical properties. *Biomedical Optics Express*, 6(10):4051 (2015).

[187] R. Vecchio, B. V. MacFayden, and F. Palazzo. History of laparoscopic surgery. *Panminerva Medica*, 42(1):87–90 (2000).

[188] 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*, pages 524–525. IEEE (2013).

[189] 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 S. S. Shen and P. E. Lewis, editors, *Proc. SPIE 8390, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, 83901B*, pages 83901B–83901B–7 (2012).

[190] L. Wang and S. L. Jacques. Monte Carlo modeling of light transport in multilayered tissues in standard C. *The University of Texas, MD Anderson Cancer Center, Houston* (1992).

[191] L. V. Wang and H.-I. Wu. *Biomedical Optics: Principles and Imaging*. John Wiley & Sons, Inc. (2012).

[192] X. Wang, Y. Pang, G. Ku, X. Xie, G. Stoica, and L. V. Wang. Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain. *Nature biotechnology*, 21(7):803–806 (2003).

[193] J. G. Webster. *The Physiological Measurement Handbook*. CRC Press (2014). ISBN 978-1-4398-0848-1.

[194] S. J. Wirkert, N. T. Clancy, D. Stoyanov, S. Arya, G. B. Hanna, H.-P. Schlemmer, P. Sauer, D. S. Elson, and L. Maier-Hein. Endoscopic Sheffield Index for Unsupervised In Vivo Spectral Band Selection. In X. Luo, T. Reichl, D. Mirota, and T. Soper, editors, *Computer-Assisted and Robotic Endoscopy*, volume 8899, pages 110–120. Springer International Publishing, Cham (2014). ISBN 978-3-319-13409-3 978-3-319-13410-9.

[195] S. J. Wirkert, H. Kenngott, B. Mayer, P. Mietkowski, M. Wagner, P. Sauer, N. T. Clancy, D. S. Elson, and L. Maier-Hein. Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression. *International Journal of Computer Assisted Radiology and Surgery*, 11(6):909–917 (2016).

[196] S. J. Wirkert, A. S. Vemuri, H. G. Kenngott, S. Moccia, M. Götz, B. F. B. Mayer, K. H. Maier-Hein, D. S. Elson, and L. Maier-Hein. Physiological Parameter Estimation from Multispectral Images Unleashed. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017*, Lecture Notes in Computer Science, pages 134–141. Springer, Cham (2017). ISBN 978-3-319-66178-0 978-3-319-66179-7.

[197] T. C. Wood, S. Thiemjarus, K. R. Koh, D. S. Elson, and G. Z. Yang. Optimal feature selection applied to multispectral fluorescence imaging. *Med Image Comput Comput Assist Interv*, 11(Pt 2):222–229 (2008).

[198] S. Wug Oh, M. S. Brown, M. Pollefeys, and S. Joo Kim. Do It Yourself Hyperspectral Imaging With Everyday Digital Cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2461–2469 (2016).

[199] M. W. Wukitsch, M. T. Petterson, D. R. Tobler, and J. A. Pologe. Pulse oximetry: Analysis of theory, technology, and practice. *Journal of Clinical Monitoring*, 4(4):290–301 (1988).

[200] K. Yoshida, I. Nishidate, T. Ishizuka, S. Kawauchi, S. Sato, and M. Sato. Multispectral imaging of absorption and scattering properties of in vivo exposed rat brain using a digital red-green-blue camera. *Journal of Biomedical Optics*, 20(5):051026–051026 (2015).

[201] C.-C. Yu, C. Lau, G. O'Donoghue, J. Mirkovic, S. McGee, L. Galindo, A. Elackattu, E. Stier, G. Grillone, K. Badizadegan, et al. Quantitative spectroscopic imaging for non-invasive early cancer detection. *Optics express*, 16(20):16227–16239 (2008).

[202] D. Yudovsky, A. Nouvong, and L. Pilon. Hyperspectral imaging in diabetic foot wound care. *J Diabetes Sci Technol*, 4(5):1099–1113 (2010).

[203] S. H. Yun and S. J. J. Kwok. Light in diagnosis, therapy and surgery. *Nature Biomedical Engineering*, 1(1):0008 (2017).

[204] Y. Zhang, S. J. Wirkert, J. Iszatt, H. Kenngott, M. Wagner, B. Mayer, C. Stock, N. T. Clancy, D. S. Elson, and L. Maier-Hein. Tissue classification for laparoscopic image understanding based on multispectral texture analysis. *Journal of Medical Imaging*, 4(1):015001–015001 (2017).

[205] R. Zhen and R. L. Stevenson. Image Demosaicing. In *Color Image and Video Enhancement*, pages 13–54. Springer, Cham (2015). ISBN 978-3-319-09362-8 978-3-319-09363-5.

[206] C. Zhu and Q. Liu. Review of Monte Carlo modeling of light transport in tissues. *Journal of Biomedical Optics*, 18(5):050902–050902 (2013).

[207] G. Zonios, L. T. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M. S. Feld. Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo. *Appl Opt*, 38(31):6628–6637 (1999).

[208] K. Zuzak, E. Livingston, J. Cadeddu, R. Ufret-Vincenty, and R. Francis. Digital Light Processing Hyperspectral Imaging Apparatus (2010).

[209] K. J. Zuzak, R. P. Francis, E. F. Wehner, M. Litorja, J. A. Cadeddu, and E. H. Livingston. Active DLP Hyperspectral Illumination: A Noninvasive, in Vivo, System Characterization Visualizing Tissue Oxygenation at Near Video Rates. *Analytical Chemistry*, 83(19):7424–7430 (2011).

[210] K. J. Zuzak, S. C. Naik, G. Alexandrakis, D. Hawkins, K. Behbehani, and E. H. Livingston. Characterization of a Near-Infrared Laparoscopic Hyperspectral Imaging System for Minimally Invasive Surgery. *Analytical Chemistry*, 79(12):4709–4715 (2007).

LIST OF PUBLICATIONS

Parts of this thesis have been published in journals and were presented at conferences during the scientific work. The author's work was awarded with the "Science Sparks Startups - Best Pitch" award (2500 Euro) and the KUKA Best Paper Award MICCAI 2014 (2nd place). Furthermore, the author was involved in related work that was published before or during the work on this thesis. The list of the author's publications, divided into both, (I.) publications related to this thesis and (II.) other publications, is presented in the following.

I. PUBLICATIONS RELATED TO THIS THESIS

i) **S. J. Wirkert**, A. Vemuri, H. Kenngott, S. Moccia, M. Goetz, B. Mayer, K. Maier-Hein, D. S. Elson, and L. Maier-Hein. Physiological Parameter Estimation from Multispectral Images Unleashed. Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2017. Pages 134-141

ii) **S. J. Wirkert**, H. Kenngott, B. Mayer, P. Mietkowski, M. Wagner, P. Sauer, N. T. Clancy, D. S. Elson, and L. Maier-Hein. Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression. International Journal of Computer Assisted Radiology and Surgery, Special Issue IPCAI. 11(6):909-917, 2016

iii) **S. J. Wirkert**, N. T. Clancy (contr. equally), S. Danail, A. Shobhit, G. B. Hanna, H.-P. Schlemmer, P. Sauer, D. S. Elson, and L. Maier-Hein. Endoscopic Sheffield Index for Unsupervised In Vivo Spectral Band Selection. In Computer-Assisted and Robotic Endoscopy 2014. Pages 110-120

iv) Y. Zhang, **S. J. Wirkert**, I. Justin, H. Kenngott, N. T. Clancy, D. S. Elson, and L. Maier-Hein. Tissue classification for laparoscopic image understanding based on multispectral texture analysis. Journal of Medical Imaging. 4(1):015001, 2017

v) M. Sara, **S. J. Wirkert**, H. Kenngott, A. Vemuri, M. Apitz, B. Mayer, E. De Momi, L. S. Mattos, and L. Maier-Hein. Uncertainty-Aware Organ Classification for Surgical Data Science Applications in Laparoscopy. arXiv preprint arXiv:1706.07002, 2017

II. OTHER PUBLICATIONS

i) **S. J. Wirkert**, E. Dellandrea, and C. Liming. Bayesian GOETHE Tracking. In 20th International Conference on Pattern Recognition (ICPR), 2010

ii) L. Maier-Hein, T. ROss, J. Groehl, B. Glocker, S. Bodenstett, C. Stock, E. Heim, M. Goetz, **S. J. Wirkert**, H. Kenngott, S. Speidel, and K. Maier-Hein. In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2016. Pages 616–623

iii) R. Altendorfer, and **S. J. Wirkert**. Why the association log-likelihood distance should be used for measurement-to-track association. In IEEE Intelligent Vehicles Symposium (IV), 2016

iv) R. Altendorfer, and **S. J. Wirkert**. Path assignment techniques for vehicle tracking. In IEEE Intelligent Vehicles Symposium Proceedings, 2014. Pages 1451-1456

v) R. Altendorfer, **S. J. Wirkert**, and S. Heinrichs-Bartscher. Sensor Fusion as an Enabling Technology for Safety-critical Driver Assistance Systems. SAE Int. J. Passeng. Cars - Electron. Electr. Syst. 3(2):183-192, 2010

Sebastian Wirkert

Eight years of experience solving real-world problems in cancer research and automotive with machine learning.

Professional Experience

07/2013–07/2017 **Research Associate/PhD Student**, *German Cancer Research Center (DKFZ)*, Heidelberg.

I am currently developing a novel multispectral endoscopy system for early cancer detection during minimally invasive surgeries within the scope of my ERC-funded doctoral thesis.

- Created the worldwide first multispectral laparoscopic setup which allows recording and analysis in video rates
- Developed techniques involve machine learning [1], transfer learning [2] and information theoretic multispectral band selection [3]
- Supervised students - bachelor student's thesis won the Thomas Gessman Prize (2000€)

02/2010–06/2013 **Automotive Software Engineer**, *ZF-TRW - Driver Assistance Systems*, Koblenz.

Developed advanced driver assistant systems like automatic cruise control and emergency braking using mid-range radar and video camera information.

- Interim project lead of advanced development project ($>$ 500,000€) with German car supplier, client is currently evaluating a possible translation to mass production
- Lead developer of path fusion module, extensively validated in driving trials, now part of production code and already in use on vehicles in real-life operation

2006–2008 **Internships**.

- Daimler AG: tested software in the Xentry Diagnosis project ($5\,1/2$ months)
- Universität Karlsruhe: lecturer in computer science (3 months)
- Universität Karlsruhe: Java developer in the KeY automatic prover project (9 months)

University Studies

10/2003–10/2009 **Computer Science**, *University of Karlsruhe (TH)*.

- Grade: "very good" (1.2, top 10% of class)
- Specialization: cognitive systems and robotics
- Complementary subjects: automotive engineering, project management, physics
- Prediploma: 1.6

Master Thesis, *Counting Persons in Videos*, Ecole Centrale de Lyon.

- Developed particle filter which handles occlusions in O($nlogn$) compared to O(n^2) baseline by efficiently modeling occlusion using state vector [4]

Herrmann-Hampe Weg 3 – 69126 Heidelberg

0049/160/91669335 • *sebastian.wirkert@googlemail.com*

Student Research Project, *Automatic Weather Classification in Traffic*.
- Enabled automatic classification of weather situations in traffic using a SVM classifier

Student Research Project, *Sub-Zero Temperature Analysing Baton (S.T.A.B.)*.
- Developed from scratch in team of three: baton for recording temperature distribution over time to support the detection of potential avalanches in skiing areas

Organization

- Member of the *2nd international workshop on cognition guided surgery* organization committee
- Organizer of the symposium on *Radiological and Surgical Data Science 2016*
- Organizer of the *1st, 2nd* and *3rd annual symposium on practical medical imaging*
- Organizer of the divisions internal seminar 2014-2016
- Member of the DKFZ students social events team 2015

Selected Honors and Awards

- "Best Pitch" at Science Sparks Start-ups (2,500€)
- Graduated top 10% of class at the KIT
- KUKA Best Paper Award MICCAI 2014 (2nd place)
- Was awarded a position in the Helmholtz International Graduate School for Cancer Research as a result of the 2012/2013 Winter Selection

Selected Publications

[1] Sebastian Wirkert et al. Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse monte carlo and random forest regression. In *Int J Cars (Special Issue: IPCAI)*, 2016.

[2] Sebastian Wirkert et al. Physiological parameter estimation from multispectral images unleashed. In *MICCAI*, 2017.

[3] Sebastian Wirkert et al. Endoscopic Sheffield Index for Unsupervised In Vivo Spectral Band Selection. In *Computer-Assisted and Robotic Endoscopy*. 2014.

[4] Sebastian Wirkert et al. Bayesian goethe tracking. In *IEEE 20th International Conference On Pattern Recognition (ICPR)*, 2010.

[5] Richard Altendorfer and Sebastian Wirkert. Why the association log-likelihood distance should be used for measurement-to-track association. In *IEEE Intelligent Vehicles Symposium*, 2016.

Herrmann-Hampe Weg 3 – 69126 Heidelberg

0049/160/91669335 • *sebastian.wirkert@googlemail.com*