Towards quantitative imaging using coded light

Andersson, David

2025

Towards quantitative imaging using coded light

DAVID ANDERSSON
DEPARTMENT OF PHYSICS | FACULTY OF ENGINEERING | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Faculty of Engineering
Department of Physics

ISBN 978-91-8104-663-2

9 789181 046632

Towards quantitative imaging using coded light

Towards quantitative imaging using coded light

by David Andersson

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Ass. Prof. Elias Kristensson
Faculty opponent: Prof. Zeyad Alwahabi

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism in Rydbergsalen (Fysicum) on Friday, the 3rd of October 2025 at 9:15.

Organization
LUND UNIVERSITY

Department of Physics
Box 118
SE–221 00 LUND
Sweden

Document name
DOCTORAL THESIS

Date of disputation
2025-10-03

Sponsoring organization

Author(s)
David Andersson

Title and subtitle
Towards quantitative imaging using coded light

Abstract

Spectral imaging allows for the simultaneous visualization of multiple different spectral bands of a scene, and is a powerful tool used in a range of contexts - from the color cameras in smart phones, to scientific imaging applications such as remote sensing and fluorescence microscopy. The progress within these fields of research is therefore intrinsically linked to the advances made for more efficient spectral capture.

The most common solution to achieve spectral sensitivity in snapshot is to use a filter array which is fixed to the sensor at a manufacturing stage. This provides a compact, stable and permanent solution, suitable for uses where the spectral bands never need to be changed. However, in many scientific measurement scenarios, spectral flexibility - the ability to alter the spectral bands captured - is crucial in order for efficient capture to be possible.

This thesis presents and discusses the camera-independent image multiplexing technique FRAME, Frequency Recognition Algorithm for Multiple Exposures, and demonstrates its ability to be used for quantitative measurements. FRAME, among other things, enables snapshot multispectral capture using a monochrome focal-plane array. Thanks to its camera independence, spectral bands can be tailored to fit the specific conditions of the measurement. Furthermore, the technique can be used in conjunction with specialized high capture-rate and intensified cameras, broadening the use of such technology. The work presented demonstrates these uses, as well as quantitative applications within temperature imaging and spray characterization that utilize intensity and spatial information respectively. The impact of image quality on neural network-based processing is also discussed, using data captured with and without coded light to suppress scattering in images, and the compatibility of multispectral FRAME and microscopic imaging is demonstrated.

These advances within the field of coded light pave the way for the continued development of FRAME for quantitative imaging applications, assisting the progress of adjacent, applied research.

Key words
Multiplexed imaging, spectral imaging, structured light, quantification, spatial frequency lock-in.

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title
1102-8718

ISBN
978-91-8104-663-2 (print)
978-91-8104-664-9 (pdf)

Recipient's notes

Number of pages
212

Price

Security classification

Signature _____ Date 2025-08-25

Towards quantitative imaging using coded light

by David Andersson

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration front: The optical setup I have designed and worked with during my PhD-studies.

Cover illustration back: A real "Kumbaya" moment.

Funding information: The thesis work was financially supported by the Swedish Research Council, Grant Number 2019-05183, and Kungliga Fysiografiska Sällskapet.

Faculty of Engineering, Department of Physics

isbn: 978-91-8104-663-2 (print)
isbn: 978-91-8104-664-9 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2025

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

Printed matter 3041 0903

MADE IN SWEDEN

This work is dedicated to five people

My grandfather Bengt, the most curious and inventive 90-year old I know

My grandmother Kerstin, who not always understood what it was that I did in my studies, but always supported it whole-heartedly

My parents Christina and Sven, who have provided everything a son could ever need

The love of my life Frida, my guiding light who improves me in every way, and without whom this would not exist

This is as much your work as it is mine

Contents

Abstract

Spectral imaging allows for the simultaneous visualization of multiple different spectral bands of a scene, and is a powerful tool used in a range of contexts - from the color cameras in smart phones, to scientific imaging applications such as remote sensing and fluorescence microscopy. The progress within these fields of research is therefore intrinsically linked to the advances made for more efficient spectral capture.

The most common solution to achieve spectral sensitivity in snapshot is to use a filter array which is fixed to the sensor at a manufacturing stage. This provides a compact, stable and permanent solution, suitable for uses where the spectral bands never need to be changed. However, in many scientific measurement scenarios, spectral flexibility - the ability to alter the spectral bands captured - is crucial in order for efficient capture to be possible.

This thesis presents and discusses the camera-independent image multiplexing technique FRAME, Frequency Recognition Algorithm for Multiple Exposures, and demonstrates its ability to be used for quantitative measurements. FRAME, among other things, enables snapshot multispectral capture using a monochrome focal-plane array. Thanks to its camera independence, spectral bands can be tailored to fit the specific conditions of the measurement. Furthermore, the technique can be used in conjunction with specialized high capture-rate and intensified cameras, broadening the use of such technology. The work presented demonstrates these uses, as well as quantitative applications within temperature imaging and spray characterization that utilize intensity and spatial information respectively. The impact of image quality on neural network-based processing is also discussed, using data captured with and without coded light to suppress scattering in images, and the compatibility of multispectral FRAME and microscopic imaging is demonstrated.

These advances within the field of coded light pave the way for the continued development of FRAME for quantitative imaging applications, assisting the progress of adjacent, applied research.

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **A light-efficient and versatile multiplexing method for snapshot spectral imaging**
D. Andersson, Yupan Bao, Vassily Kornienko, Dean Popović, Elias Kristensson
Scientific Reports **14**, 2024, Article Number 16116

II **Multispectral coded light for time-resolved temperature imaging of aluminum combustion**
D. Andersson, Z. Wu, Z. Li, E. Kristensson
To be submitted.

III **Simultaneous multiple time scale imaging for kHz–MHz high-speed accelerometry**
V. Kornienko, **D. Andersson**, M. Stiti, J. Ravelid, S. Ek, A. Ehn, E. Berrocal, E. Kristensson
Photonics Research **10.7**, 2022, pp. 1712-1722

IV **Using SILMAS to improve machine learning-assisted quantification of pathology**
D. Andersson, D. Frantz, D. Kirik, E. Berrocal, E. Kristensson
Optics Letters **50.10**, *2025, 3465-3468.*

V **Snapshot multicolor fluorescence imaging using double multiplexing of excitation and emission on a single detector**
K. Dorozynska, S. Ek, V. Kornienko, **D. Andersson**, A. Andersson, A. Ehn, E. Kristensson
Scientific Reports **11**, 2021, Article Number 20454

VI **Simultaneous multispectral microscopy using passive FRAME multiplexing with seven spectral channels**
D. Andersson, S. Ek, V. Kornienko, J. Beech, A. Merdasa, E. Kristensson
Submitted.

Acknowledgements

And so, the time has come for me to write the part that most people turn to first - the acknowledgements. As I am writing this, I am sitting in my hammock in the garden on a Sunday afternoon, with Frida basking in the sun a few meters away. Not much time of this summer has gone to anything but work, but who needs vacation when you have the support of friends and loved ones to power you through even the toughest of times. This is my chance to acknowledge those people, *and acknowledge them I will*[1].

First of all - to the people of Combustion Physics, who all contribute to making it such a welcoming work environment - thank you! The fact that so many students return to the division in one form or another speaks volumes. As is true for everyone, I have gotten to know some of you better than others, and so some special thanks are in order. To Sebastian, Sam, Martin, David Sanned, David Frantz, Adrian, Saga, Zhiyong, Can, Meena and many others of my fellow PhD-students and postdocs - thank you for all the banter and laughs, whether it be at or outside of work. I have also had the pleasure of collaborating with many of you in my research - thank you for all your hard work! To the backbone of the division, meaning the senior staff, I want to say the following: You are the main reason this division continues to do well and attracts so many nice people! PhD and postdocs come and go - there would be no great atmosphere year after year if it was not carried over by you. Thank you for all the hard work you put in for the division and for your students!

To my co-supervisors, Ed and Andreas - thank you for contributing to making this education so enjoyable. Your official roles as co-supervisors have not been used to any significant extent (which I believe is every co-supervisor's dream) - not for a lack of dedication on your part, but because I have had an excellent main supervisor. You have always given input whenever I have asked for it, while also providing an enjoyable work environment[2]. Thank you both! I also want to say a special thank you to Zhongshan Li for keeping your door open to answer my questions on pyrometry.

To past and current members of the FRAME research group - Elias, Vasse, Simon, Yupan, Karolina, Sam and Klara - thank you for all the useful exchanges and input! The helpful atmosphere and focus beyond personal prestige that you all have contributed to has made working within this group incredibly rewarding. Simon - your attention to detail and drive to not only make things work, but make them work *well*, really shows in the quality of your results. I am glad we got to work together a little bit more before you left the division. Yupan - thank you for your constant enthusiasm and the

[1]Please read this last part with an ominous voice for comedic effect.

[2]Ed, your legendary dance moves will forever be imprinted on my retina.

many laughs both when working and when taking a break.

To Steph, Martin, Emma, Hans, Jonas and Zoe - holy diver! Thank you for welcoming me into your role-playing group with such warmth, and for putting up with my overtired shenanigans over the course of several years. I never would have imagined that the invitation from Jonas to join "just to try it out" would become such a steady pillar in my weekly life. Thank you all for the laughs, snacks and memes, and a special thank you to Steph and Martin for opening up their home every week and for being such great hosts. Martin, thank you for the lovely cooking, and Steph, thank you for your hard work in keeping the group sane - it can't be easy.

Anton - thank you for the many evenings of talking over Discord, whether it has been about work, games, or something else entirely. Your support and company means a lot to me, and I am very grateful for the almost daily updates we give each other on what is going on in our lives. Thank you for always listening, and for being such a force of kindness in my life!

Nycklis - one can't help but be inspired when being around you. You always have a project or a build going, whether it is building your own hot tub, making your own wine, or salvaging and repairing a 5.1 surround sound system from the bulky waste of your student housing. You somehow manage to combine a down-to-earth calmness with the energy and drive of a Duracell rabbit, and to top that off you are one of the most thoroughly kind people I know. Thank you for the laughter, hype and thoughtfulness you bring into my life!

Tim - I feel honored that you are my friend. Hanging out with you is effortless, and always results in tonnes of laughter, whether it is meeting up for a beer after work or a whole day in the woods. Your knowledge of nature and your eagerness to share it with others is truly inspiring, and I learn something new every time I see you. I am very grateful for the invitations I have gotten from you to join you and the others for days of hunting in Slättåkra, and I hope to one day be able to return the favor. Thank you for being such a caring friend and for your tutorship, it means a lot to me! I also want to thank the rest of the hunting team at Slättåkra for the warm hospitality you have shown me - a special thank you to Thomas and Måns for your excellent leadership and the entertainment over the radio.

To Glenn, Fia, Cajsa and Wilma - thank you for the support and encouragement you have given not only your daughter and sister, but me as well. Your home has become a place of rest and pleasant company[3] for me and Frida during our years of study, and you never seem burdened by it - quite the opposite! Your warm welcome every time we visit means a great deal to me, and I am deeply grateful for the kindness you always show.

[3]As well as great food!

Oscar - my oldest friend. Even though we only see each other about once a year nowadays, it is still one of the highlights of my year. I still get excited just like I did when I was 7 years old and my mom told me we were going to go see you and your family, and it only takes about 5 minutes of talking until we are back where we were 10 or 20 years ago. It warms my heart to see you embrace fatherhood with such joy, and I hope that I will be able to see you and your family more frequently in the coming years. Thank you for your support and for being such a reliable friend, both in present day and during our upbringing!

To my grandmother Anita, whose skill in the kitchen is only rivaled by her love and support for her children, grandchildren, and great-grandchildren. The trips we took to Denmark to see "gammelmormor" when I was a child are some of my fondest childhood memories. Thank you for your abundance of love and for all the things you do for your family!

To the rest of my extended family, who always show interest in what I do and cheer me on in my pursuits - thank you! Your encouragement has meant a lot to me, especially during these last few months. To always be welcomed with open arms and smiles, despite probably having missed more than 90% of the birthday parties in the last 10 years, is a blessing.

And so for the people I cannot thank enough - some of which have already been mentioned by name, but not to the extent that they deserve. In short, these are the people without whom I would not be here.

First of all, my supervisor Elias. Much of the reason these years have been as comfortable[4] as they have been is because of your leadership. You have many times believed in me more than I have believed in myself, for which I will always be grateful. Always positive, always available, and always coming up with new perspectives - an invaluable skill that I am still a bit in awe over how well you have mastered, despite having known you for more than half a decade by this point. I have sometimes found myself going to your office just to hang out, which I imagine is quite rare for a PhD-student. Your dedication to the people who depend on you - students, colleagues, employees, and most importantly your family - is truly admirable. I think your priorities are clear to anyone who has ever worked with you - your child's football practice trumps work (as it should)! Thank you for being a joy to be around, for always answering my questions and for believing in me!

Vasse - it is an art to simultaneously be both sharp as a tack, empathetic like few others, and at the same time the most unhinged person I know. Your willingness to help others seemingly knows no bounds, and the uncompromised optimism you display

[4]All things are relative. Jokes aside, you have made them as easy as they can be.

every day in every situation is unparalleled. It doesn't matter if it's about walking from the office to the lunch room, or unloading 10 literal metric tonnes of equipment out of a truck at 4 in the morning after having worked 20 hours straight - your answer is always the same: "*LET'S DO THIS!!*". A large part of why I wanted to work at the division was because I knew that I would be working alongside you, and it has been every part as good as I imagined it would be. You have been a mentor to me in many ways, and our trip to Yosemite is a memory I will treasure for the rest of my life, along with many others like it. Thank you for all your help and your friendship - I could not have done this without you.

Jonas - the big brother that I didn't know that I needed, but now can't live without. It is a blessing to have a friend with which you can always be 100% honest when they ask you how you are doing, with no fear of being judged, while at the same time knowing the same trust will be extended to you. To on top of that be able to do a PhD alongside that friend is beyond anything I could have wished for. Our patented blend of serious talks and stupid jokes is a large part of what has gotten me through this, and I would not be where I am without your support. Thank you for choosing to talk to that new student next to you on "Ärtor och Punsch" all those years ago - he got a friend for life.

To PO and Johan, founders of the Physics- and Lasershow - I don't think anyone has ever had such a profound impact on my life with a single action as you did when you asked me to join the show. Little did I know that it would give me friends that I consider family, not to mention introduce me to the love of my life! As if that wasn't enough, you have (among a lot of other things) also given me an appreciation for being on stage, for the value of education, and for hard work (and the rest that follows it). You saw something in me that I did not, and you awoke it and nurtured it - thank you for everything!

To all past and current members of the Physics- and Lasershow that I have had the honor to get to know - it is hard to overestimate the importance of your friendship, and the joy it has brought to my years at the university. Friendships are forged in fire, something that we have all learned on many occasions[5]. I consider you family, and I am certain that I will still be seeing many of you when we are old and gray. A special thank you to Daniel, Tina, Anna-Maria, Alex, Elin, Vasse and Lina for all the memories (so far) at the annual summer-retreats, which are the highlight of the summer. Daniel, thank you for your words of support and for introducing me to the wonderful (and endless) world of board games.

Till min morfar Bengt - jag vill tro att den fallenhet jag har för vetenskap har jag främst fått från två håll - från dig och min pappa. Att ha en morfar som vid 90 års ålder inte

[5]I seem to recall some sort of delay in the delivery of our equipment in Podgorica..?

bara undrar hur “de där neurala nätverken funkar”, utan också redan har läst på lite och därför kan hålla igång en diskussion, är få förunnat. Tack för din inspirerande nyfikenhet, din sprudlande energi och din otroliga uppfinningsrikedom - om jag har fått ärva en tiondel av den så är jag överlycklig!

Till min farmor Kerstin - tack för den bottenlösa kärlek du visade alla i din närhet - den formade min barndom till den saga jag drömmer mig tillbaka till så här i vuxen ålder: Att hämta ved uppe på vedbacken, klippa gräset, spela fotboll på den nyklippta gräsmattan med gungställningen som mål, och bygga Lego på köksgolvet med Allsång på Skansen på den lilla burk-TV:n. Många av mina mest minnesvärda solnedgångar är de över den stora åkern i Önneröd. Tack för allt du gav mig.

Till min bror Simon, den eftertänksamme våghalsen - ditt lugn, din självkännedom och din känslomässiga intelligens är bara några av de saker som jag ser upp till hos dig, och som jag beundrar djupt. Ditt skratt kan lysa upp ett rum, och varje gång vi ses märker jag hur mycket jag har saknat dig. Jag är otroligt tacksam över att ha fått växa upp med en så snäll och tålmodig bror, och för den fina relation vi har idag.

Till min syster Sara, den mest empatiska person jag känner - tillskottet av dig till vår familj är det bästa som kunde ha hänt oss, och jag har svårt att minnas hur det var innan du föddes - det känns som att det alltid saknades en person? Du är en gåva - en glädjespridare utan dess like, stark som få, och det är en ynnest att vara din bror.

Till mina föräldrar, Christina och Sven - inget av detta hade varit möjligt om det inte var för den kärlek, den vägledning och det stöd ni gett mig under hela mitt liv. Ni är delade världsettor i att sätta andras behov framför era egna, och ert hårda arbete för att ge Sara, Simon och mig så många möjligheter i livet som möjligt är något som jag aldrig kommer att kunna uttrycka min tacksamhet över. Tack för er gränslösa kärlek, och för att ni alltid ställer upp.

And finally, Frida - the light of my life, my lighthouse and my harbor. Whatever amount of luck a person gets in life, I must have used up mine the day I met you. You constantly and tirelessly give of yourself, and I am in a perpetual state of awe over how amazingly talented and hard-working you are in all areas of life. Many people have had an impact on this work, but none has been as all-encompassing as yours. Your love and support is the reason this exists, more than any other - it’s not even close. I treasure the fact that we have been able to embark on this journey together, and look forward with excitement to the next chapter of our lives. You are my lobster, and I love you.

Populärvetenskaplig sammanfattning

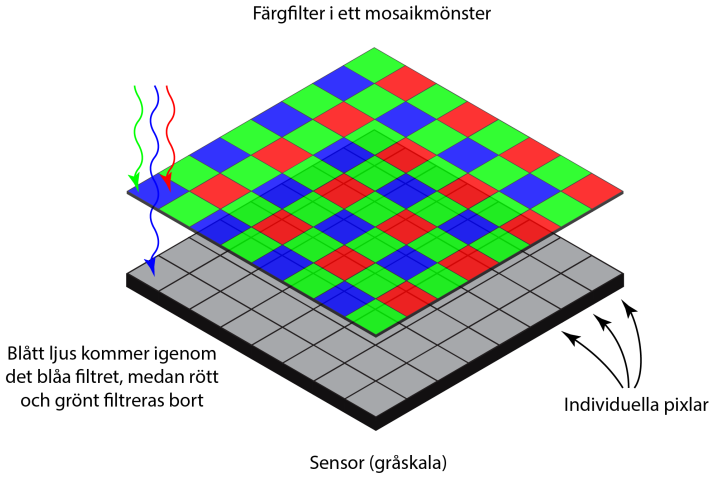
Vad vore livet utan färg? Allt från hur himlen ändrar färg under dygnet, från dagens klarblå till skymningens alla nyanser av rött, oranget och rosa, till alla de tusentals nyanser av grönt som finns i naturen. Men färg är inte bara något som tillför skönhet i våra liv, det innehåller också information. Har du exempelvis någon gång undrat hur det kommer sig att jägare klär sig i orange vid älgjakten? Är inte det det absolut sämsta valet av färg under en jakt, där man vill vara så dold för djuret som möjligt? Faktum är att just orange är ett utmärkt val, eftersom älg och andra klövdjur saknar den sortens *tappar* - sinnesceller som gör att vi kan uppfatta färger - som krävs för att urskilja denna färg. Det gör att för älgen smälter skogens gröna och bruna ihop med det orangea, medan det för oss människor som kan särskilja den orangea färgen är mycket tydligt var de andra i jaktlaget befinner sig. Ett mer utvecklat färgseende ger mer information!

Vi människor har tre sorters tappar i våra ögon, där varje sorts tapp täcker ett visst färgintervall, och kombinerat gör dessa att vi kan se de färger som ingår i regnbågen[6]. Lite förenklat kan vi säga att de tre olika sorters tapparna känner av rött, grönt respektive blått ljus. Faktum är att det är exakt samma princip som ligger bakom en färgbild som vi fångar med till exempel en mobilkamera[7]. När du tar en bild fångas egentligen tre olika bilder (som vi kallar *färgkanaler*) samtidigt - en med bara det röda ljuset, en med bara det gröna, och en med bara det blåa. Detta sker med hjälp av små färgfilter som sitter påklistrade på varje *pixel* (bildelement) på kameran i ett mosaikmönster, så att en specifik pixel bara nås av ljus av en viss färg. Fig. 1 visar hur denna färgfiltrering går till. Vad du sedan ser på skärmen är dessa tre bilder sammanflätade i samma mosaikmönster, men där sammanflätningen är så liten att ögat uppfattar det som att det röda, gröna och blåa ljuset kommer från samma punkt. Detta ger illusionen av den färg som de tillsammans ger.

För en fysiker är färg en väldigt användbar egenskap hos ljus. Ljusets färg bestäms av dess *våglängd*, som är direkt kopplad till hur mycket energi ljuset har. Eftersom det finns många processer inom fysiken som skickar ut ljus så kan ljusets våglängd ofta användas för att förstå hur mycket energi som frigjorts. Problemet med att använda en kamera med påklistrade färgfilter är att olika processer släpper ifrån sig ljus med väldigt olika energier, så det är sällan de färgfilter som sitter påklistrade passar för de färger man vill fånga. Det existerar såklart kameror med olika påklistrade färgfilter, men kameror som används för vetenskapliga mätningar är dyra, och det är inte hållbart

[6] I den synliga delen av regnbågen - den innehåller ju trots allt fler färger än de vi kan se.

[7] Detta är självklart ingen slump - kameran är ju byggd med syftet att skapa färgbilder som stämmer överens med hur vi uppfattar färger, och kamerans "färgseende" är därför uppbyggt på samma sätt som vårt.



Figur 1: En gråskalesensor med färgfilter framför varje individuell pixel gör att en färgbild kan skapas. Detta är hur de flesta färgkameror fungerar, inklusive mobilkameror, men det är oftast inte en praktisk lösning inom forskningen.

att köpa en ny kamera varje gång man ska mäta något nytt. Dessutom är denna sorts kameror inte så effektiva när det gäller att ta vara på allt ljus, eftersom de färgfilter som används filtrerar bort de färger som inte passar. Det är här min forskning kommer in i bilden.

Denna avhandling handlar om mitt arbete med att utveckla en alternativ metod för att ta färgbilder, riktad mot vetenskapliga mätningar där färgfilter behöver kunna bytas ut, men där man samtidigt samlar in alla färger samtidigt. Den metod som jag arbetat med använder sig av en kamera som tar *gråskalebilder*, alltså bilder som inte innehåller färginformation[8]. Genom att märka ljus av olika färger med kända mönster innan det når kameran kan färgerna separeras i efterhand - inte för att deras faktiska färg går att urskilja på bilden, utan för att deras mönster känns igen och kan skiljas åt. Tänk tillbaka på älgen - den hade inte kunnat skilja på orange och grön, men om du hade haft ett schackrutigt oranget mönster på jackan hade den stuckit ut ändå, eftersom detta är ett väldigt onaturligt mönster. Med denna metod fångas alla färger samtidigt, vilket är viktigt vid snabba processer, och vilken kamera som helst kan användas för att fånga vilka färger som helst. Inte bara det, utan metoden med vilken ljuset samlas in och märks sker utan att behöva "slänga bort" ljus på grund av färgfiltret, vilket gör att mer av ljuset kommer till nytta. Detta hjälper oss att titta på processer som är svåra att fånga i färg på grund av att de sker mycket snabbt eller är mycket ljussvaga.

På detta sätt har jag jobbat med att fånga bilder och filmer i färg av snabba processer.

[8]Vad som i folkmun brukar kallas för "svart-vita" bilder. Tänk dig ett gammalt fotografi eller en gammal TV utan färg.

Färgen hjälper oss att förstå den underliggande fysiken, till exempel vilken temperatur ett glödande objekt har. Du har säkert sett att exempelvis järn som hettas upp till tillräckligt hög temperatur börjar lysa, först svagt i mörkrött, och sedan starkare och gulare allt eftersom temperaturen ökar, tills det till slut lyser väldigt starkt och nästan vitt. Eftersom färgen som metallen lyser i är direkt kopplad till dess temperatur kan vi med hjälp av en färgbild ta en bild av ett objekts temperaturfördelning. Detta är en av de saker jag gjort i mitt arbete, men i stället för att titta på en stillastående järnbit har jag undersökt hur temperaturen skiljer sig åt i olika delar av 0,5 mm stora aluminiumpartiklar som brinner och samtidigt rör på sig, och skapat filmer av detta med 50 000 bilder i sekunden(!). På liknande sätt har jag skapat bilder av var olika gaser befinner sig i en plasma-urladdning - samma sorts fysik som orsakar blixtnedslag. Sådana bilder och filmer hjälper oss att förstå vad som pågår i händelseförlopp som annars hade varit omöjliga för oss att se, både för att de sker för snabbt för att ögat ska kunna uppfatta dem och för att våra ögon inte har den färgkänslighet som behövs. I förlängningen för detta forskningen för ett mer hållbart samhälle framåt.

Popular summary

What would life be without color? Everything from how the sky changes its shade throughout the day, from the bright blue of day to the twilight's many hues of red, orange, and pink, to the thousands of shades of green found in nature. But color is not only something that adds beauty to our lives, it also carries information. For instance, have you ever wondered why hunters wear orange during the moose hunt? Isn't that the absolute worst choice of color for a hunt, where the goal is to stay as hidden from the animal as possible? The fact is that orange is an excellent choice, because moose and other hoofed animals lack the kind of *cone cells* - the sensory cells that allow us to perceive color - needed to distinguish this particular color. For the moose, the forest's greens and browns blend together with the orange, while for us humans who can tell orange apart, it becomes very clear where the other members of the hunting party are. A more developed sense of color means more information!

We humans have three types of cones in our eyes, with each type covering a certain range of colors, and combined, these allow us to see the colors that make up the rainbow[9]. Put simply, the three types of cones are sensitive to red, green, and blue light, respectively. In fact, this is the very same principle behind a color image captured with, for example, a mobile phone camera[10]. When you take a picture, what is really captured are three separate images (which we call *color channels*) simultaneously - one with only the red light, one with only the green, and one with only the blue. This is achieved with the help of tiny color filters attached to each *pixel* (picture element) of the camera in a mosaic pattern, so that each pixel only receives light of a specific color. Fig. 2 shows how this color filtering works. What you then see on the screen is these three images interwoven, but blended so finely that the eye perceives the red, green, and blue light as coming from the same point. This creates the illusion of the combined color they produce.

For a physicist, color is a very useful property of light. The color of light is determined by its *wavelength*, which is directly linked to how much energy the light carries. Since many processes in physics emit light, the wavelength of that light can often be used to understand how much energy has been released. The problem with using a camera with attached color filters is that different processes emit light with very different energies, so the built-in filters rarely match the colors one wants to capture. Of course, there are cameras with different sets of filters, but cameras intended for scientific measurements are expensive, and it is not sustainable to buy a new camera every time something new needs to be measured. On top of that, this kind of camera is not very

[9] In the visible part of the rainbow - after all, it contains more colors than the ones we can perceive.

[10] This is of course no coincidence - the camera is designed to create color images that correspond to how we perceive colors, and its "color vision" is therefore built in the same way as ours.

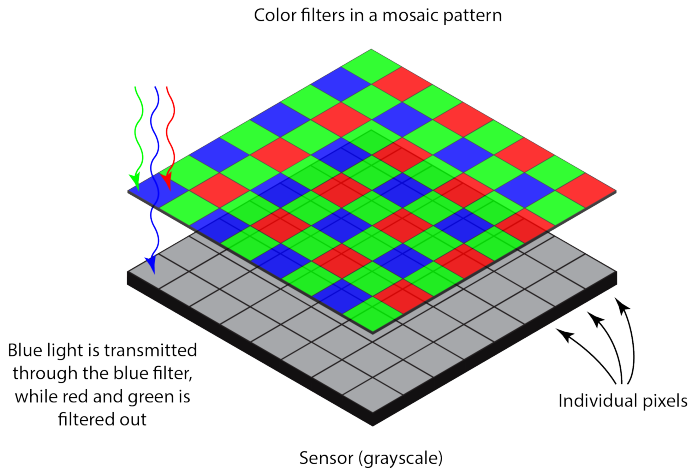


Figure 2: A grayscale sensor with a color filter placed in front of each individual pixel makes it possible to create a color image. This is how most color cameras work, including mobile phone cameras, but it is often not a practical solution in research.

efficient when it comes to capturing all the available light, since the filters block out the colors they are not designed for. This is where my research comes into play.

This thesis is about my work on developing an alternative method for capturing color images, aimed at scientific measurements where color filters need to be exchangeable, but where all colors must still be recorded simultaneously. The method I have been working on uses a camera that records *grayscale images*, that is, images that contain no color information[11]. By tagging light of different colors with known patterns before it reaches the camera, the colors can later be separated - not because their actual color can be distinguished in the image, but because their patterns can be recognized and told apart. Think back to the moose - it would not have been able to tell orange from green, but if your jacket had a checkered orange pattern, it would still have stood out, since this is a very unnatural pattern. With this method, all colors are captured simultaneously, which is important for fast processes, and any camera can be used to capture any set of colors. Not only that, but the method by which the light is collected and tagged does not require "throwing away" light because of the color filter, which means that more of the light is actually used. This allows us to study processes that are difficult to capture in color because they are either very fast or very faint.

In this way, I have worked on capturing images and movies in color of fast processes. Color helps us to understand the underlying physics, for example the temperature of a glowing object. You have surely seen that for instance iron heated to a sufficiently

[11]What is commonly referred to as "black-and-white" images. Think of an old photograph or an old TV without color.

high temperature begins to glow, first faintly in dark red, and then more bright and yellow as the temperature increases, until it eventually shines very brightly and almost white. Since the color that the metal emits is directly linked to its temperature, we can, with the help of a color image, record a picture of an object's temperature distribution. This is one of the things I have done in my work, but instead of looking at a stationary piece of iron, I have investigated how the temperature differs across different parts of 0.5 mm aluminium particles that are burning while simultaneously moving, and created movies of this at $50{,}000$ frames per second(!). In a similar way, I have produced images showing where different gases are located during plasma discharges - the same type of physics that cause lightning strikes. Such images and movies help us to understand what is happening in sequences of events that would otherwise be impossible for us to see - both because they occur too quickly for the human eye to perceive and because our eyes lack the necessary color sensitivity. In the long run, this advances research towards a more sustainable society.

Towards quantitative imaging using coded light

Chapter 1

A first look

"It's a dangerous business, Frodo, going out your door. You step onto the road, and if you don't keep your feet, there's no knowing where you might be swept off to."

- Bilbo Baggins, The Lord of the Rings

1.1 Our vision

When you really think about it, the human sense of vision is truly staggering. We can create a next-to-perfect three-dimensional sense of not only our immediate surroundings, but also the far distant world around us, using only the stereoscopic vision of our two eyes, spaced about 10 cm apart, and the biological processing power of our brain. It is the sense that we as a species unequivocally value the most[1], and plays a part in almost all activities. It is hard to overestimate its importance.

Each one of our eyes produces an image of our field of vision on its inner back, called the retina, using a lens located behind the pupil, the round opening in the center of the iris. The shape of the lens can be altered using muscles, enabling us to focus our vision to different distances [2], and the size of the pupil can also be changed using the iris, to let in more or less light. Not only that, but the sensory cells used to detect the light on the retina also come in a few varieties for extra flexibility - rod cells are very sensitive to light but can only separate brightness from darkness, while cone cells are less sensitive but can distinguish between different wavelengths of light, giving us the ability to see different colors[3]. Using all these components, we can create two images, one with each eye, of the same field of vision. Well, *almost* the same. As the eyes are located at slightly different points in space, their perspective of the object

in focus is also slightly different. This is key to be able to judge the distance to the object[1] [4; 5]. In everyday life, our brain does a fantastic job of not involving our conscious selves in the neurological algorithm that the brain uses to infer distance, but the difference in position of the eyes is the hardware that makes it possible to do this without moving your head. If you would like to test how well you would be able to judge distances without two eyes, try the following: Close one eye, put your arms out in front of you (not completely stretched, that is cheating) and space them about 50 cm apart. Now, point your index fingers towards each other and try to get them to end up point-to-point. Then do it with both eyes open.

1.2 Extending our vision

It might seem strange to start a physics thesis with a lecture on the biology of the eye, but as the subject of this thesis is imaging, I feel it is a natural starting point to go back to the roots. After all, the eye is the technical predecessor to the camera, and almost all modern cameras operate on the same fundamental principles as the eye. Take the modern color camera as an example. Whereas the eye uses a single lens and muscles to change its focusing distance by altering the shape of it, a process known as accommodation [2], the camera instead uses a set of lenses at interchangeable distances to produce a focused image. The aperture of the camera lens is an almost identical mechanical replica of the iris in the eye, being able to open and close, and the sensor is the electronic analogue to our biological sensor, the retina. The sensory cells of the camera sensor - the color-filtered pixels - try to emulate the cone cells of the human eye by filtering light into three different color bands [6]. This is of course no coincidence - the purpose of the everyday consumer color camera is after all to reproduce the scene in colors as we see them, and as the cone cells are sensitive to spectral bands roughly centered around the colors red, green and blue (commonly referred to as RGB), so are the pixels in a color camera. Even the lens lid can be seen as an analogue to the eyelid. Fig. 1.1 demonstrates these analogues between the eye and the camera.

[1] At least when remaining stationary. Many animals utilize for example head movement to in this way get different perspectives. The point is that different perspectives are needed in order to judge distance, and stereoscopic vision allows for two simultaneous perspectives at all times.

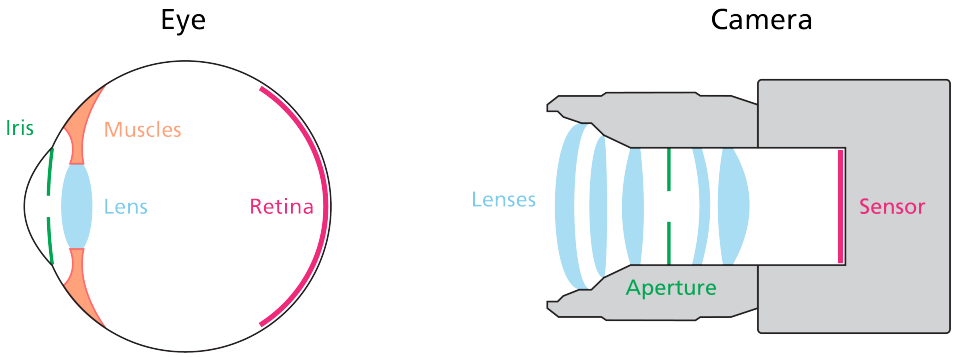


Figure 1.1: The eye and the camera share many components in terms of functionality, such as a lens for focusing, an aperture to vary the level of light entering the system, and a retina/sensor for registering the intensities.

Before the invention of the camera, the only way we had of preserving a visual experience was art, such as painting and sculpting - an inherently subjective method. As the medium of the creation process is a human, art is also limited by the sensory capabilities and skill of the artist, such as the colors with which and the rate at which they can perceive the scene, as well as the quality of and the rate at which they can subsequently reproduce it. This limits depiction to objects within the visible region of the electromagnetic spectrum that also remain relatively stationary over the time it takes to finish the image, or alternatively, to rely on ones visual memory. Humans are incredible in many aspects, but eye-witness testimony is not one of them [7]. Even with a perfect visual memory, some timescales are simply out of the grasp of our senses without technological aid [8]. The invention and development of the camera changed this - it has given us the ability to perceive the world as if frozen in time, to explore areas of vision outside our own, and to accurately share and preserve such perspectives.

At its technological dawn, the camera was a thitherto unrivaled method of objectively and systematically capturing a visual perspective in time, which could also be shared with others in a physical format. As development continued, frame rates increased, allowing images to be captured in an increasingly rapid succession. In a similar fashion, spectral resolution increased. At first, all light was recorded in a single color band, known as monochrome or grayscale imaging, meaning no colors were discernable, only luminescence. This was done by exposing photosensitive chemicals on a film to light, and later developing the photograph using the created negative. With the discovery of chemicals sensitive to different bands of the electromagnetic spectrum [9], different parts of the spectrum could be captured simultaneously by layering the chemicals on the film, and color photography was born. For the first time in human history, a scene could be depicted in color without the filtering of a human artist.

In modern digital cameras, the analog photographic film has been replaced by digital

sensors, consisting of a number of pixels which determine the spatial resolution of the image. Replacing the photosensitive chemicals which gave us spectral division are primarily color filter arrays[2] [10], which filter the light so that each pixel primarily detects photons of a certain wavelength interval. The transition into electronic registration and storage of images has ushered in a new era of information gathering. With the parallel improvement of adjacent technology, such as optics, mechanical engineering and processing algorithms, the images we are able to capture today are truly remarkable.

1.3 The scientific value of images

Perhaps the greatest gift brought by the invention of the camera is this: For the first time in human history, we were not limited to our own visual capabilities and memory. That this is a valuable property for scientific research is hardly shocking, but I want to use this opportunity to break down some of the principal ways in which this helps us, to really emphasize its impact.

The first aspect I want to discuss is time. Change with time is abundant in nature, but it is the only one of the four dimensions that we experience (three spatial and one temporal) that we have no control over. We can to a certain extent control how we move in three-dimensional space, but not in time. Time moves on, unperturbed, regardless of our efforts. We experience time at a sampling rate determined by our neurobiology [11], and as previously mentioned, some timescales are simply out of the grasp of our senses. But, with the ability to record images, we have all the time in the world to get to the bottom of what we are seeing. We do not have to keep up with the process in real time as it unfolds, but can instead analyze images at a leisurely pace (with a cup of coffee) comfortably knowing that it is not going to disappear or expire, and the event be lost in time. A video can be rewinded and watched again and again, and be played slower or faster than we would otherwise be able to experience it. The high frame rates of modern high-end commercially available cameras, being able to film at $\sim$76 kHz at 1 MP resolution [12], also allows us to sample time at a much higher frequency than our own vision, allowing us to explore timescales outside our own.

The second aspect is that of spatial scale. This aspect has less to do with photography and more to do with optics, specifically magnification, but photography plays a part, as we will soon see. Our own vision is limited in that we cannot optically magnify what we see - for this we would need multiple lenses in our eyes, since the distance from the lens to the retina is small compared to the distances at which we can focus.

[2]Although other solutions exist, as we will later discuss.

However, using a second lens (apart from the one in our eye) we can magnify the image, increasing the size of the image before it reaches our eyes. Both astronomy and microscopy are great examples of fields where this is fundamental. Magnification alone has enabled us to understand more of the world around us, but when combined with a camera, it becomes even better, and here we will make use of the first aspect - time. A camera has the ability to dramatically change the exposure time of an image, something that our naked eyes are not capable of. By collecting light for several minutes in a single image, even the dimmest objects in the night sky can be visualized. Similarly, but on the other side of the spectrum, a short exposure can freeze the Brownian motion of the microscopic world [13].

The third and final aspect is color. First of all, I want to be clear with what I mean when I refer to "color". Color is our interpretation of the wavelength of light, as filtered and perceived by human vision, limited to approximately 400-750 nm[14]. We divide this range into smaller parts, with somewhat arbitrary boundaries, and assign names for them such as "green", corresponding to light with an approximate wavelength of 500-570 nm. The wavelength of light is inversely proportional to its energy, and through imaging we can recreate a spatial distribution of said energy in many processes where photons are emitted. This fact alone makes spectral imaging very powerful, as energy is a fundamental quantity in physics. It becomes even more useful when we consider all the parts of the electromagnetic spectrum outside the visible range that we can image - radio, microwave, infrared, X-rays and gamma[3] [15–20]. Once again, the camera helps us discover areas otherwise outside our grasp.

1.4 What is an image?

At its core, an image is a spatial distribution of intensity. There are many types of images, but for this thesis I will limit the discussion to two types - an optical, physical image and a digital, sampled image. Please make extra note of the word *sampled* in the previous sentence, excluding synthetic digital images and the like, and limiting the discussion to images depicting real-life scenes.

An optical image is a redistribution of light from an object, into an (imperfect) spatial light distribution of the object at a different point in space - in other words, a recreation of the light field of the object, seen from a certain perspective. The reason for the addition of the word "imperfect" is that a real optical image[4] is never perfect - various effects, such as aberration and diffraction, prevent a perfect reconstruction. Never-

[3] It should be noted that some of these techniques do not directly and optically image the photons, but rather recreate a spatial distribution of them. Nonetheless, it is imaging.

[4] "Real" in the sense that it is created in practice, not only in theory.

theless, the object can be recognizable, and this is (to my mind) when it becomes an image. A digitally sampled image, on the other hand, is the discretized representation of a physical image, sampled using a digital sensor. This means that intensity has been quantized as a function of space, by converting incident photons into electrons in a number of photodetectors, where each photodetector corresponds to an image element (pixel) in the final representation of the image. The combined charge of the electrons is then measured, quantizing the intensity[5] into "counts". Unlike an optical image, a digital image can be stored, and its dimensionality may vary - one could say that the definition of a digital image is a little bit broader. A tomographic image, an RGB-color image, and an image captured by a streak camera all have different dimensionality, yet are all called images. The most basic form of a digital image is a monochrome image, represented by a two-dimensional matrix. In this matrix, the rows and columns correspond to position in space, and the value of each element corresponds to the intensity in that position[6], integrated over the spectral response of the camera and the exposure time.

1.5 Multispectral images

As my work concerns itself largely with multispectral imaging, I think it fitting to here introduce what these are and how they are represented in an image. The term "multispectral image" refers to an image which contains spectrally resolved image information from a small number of distinct, spectrally separated bands. Multispectral images can cover any part of the electromagnetic spectrum, and typically contain 3-15 spectral channels [21]. If the spectral resolution is further increased, with more and narrower contiguous spectral bands, the image will move towards the "hyperspectral" [22; 23]. I also want to make an (in my mind) important distinction here, which will be relevant in later chapters - an RGB-image, whose purpose is to emulate human color vision, is *not* a good example of a multispectral image, at least not in technical terms. Even though it contains three color channels, the spectral responses of these channels almost always have overlap, just as for our different cone cells. Thus, the resulting intensities in adjacent spectral bands are codependent, which is a problem in many technical applications. As humans, we can actually experience this problem as well - two objects can appear to have the same color, only to look to be different colors in other light conditions. This phenomenon is called metamerism, and is a direct result of the overlapping spectral responses of our three cone cell types [24]. It is also what enables us to perceive a computer screen as for example yellow - even though no light with a wavelength corresponding to yellow is actually emitted, the

[5] Intensity is of course quantized in photons in nature as well.

[6] Technically over an area, albeit incredibly small.

relative response of the cone cell types have the same proportions as if you were to have actually seen yellow light.

As wavelength is a property of light, the resulting colors of an optical image are determined by the continuous spectral profile of the irradiance incident on each point in the image plane. However, when sampled by a digital image sensor, the continuous spectrum is discretized, adding another dimension to the digital image in the form of spectral bands. The size of the added dimension depends on the spectral resolution with which the image is sampled. However, as spectral information such as the spectral response curve of each channel is not stored in the digital version of the image, the interpretation of the spectral channels is an intellectual construct, and without this external information the scene cannot be properly reconstructed. Putting all three dimensions together, along with the spectral meta-information, the spectral image can be described as a data cube, which will be discussed further in Chapter 2.

1.6 Quantitative imaging

Another distinction I want to make early on in this work is what constitutes a quantitative measurement. Some might not agree with all definitions or statements made hereafter, but at least know that this is what I have considered when I choose to call a measurement "quantitative".

The essence of a quantitative measurement, whether it be imaging or something else, is the aim to attach physical quantities to measured values[7]. We want to come as close as possible to an objective truth - the actual value of a physical parameter. In comparison, a qualitative measurement is more subjective, requiring interpretation. Such measurements are also of value, for example in situations where quantitative values do not contribute with very much, or are excessive. A good example of this within imaging is diagnosing bone fractures using X-ray images - the doctor does not need to know the exact intensity values of the X-ray image to understand that the bone is fractured, the relative contrast is enough. Almost all methods start out as qualitative, simply because it is most often the easiest to achieve. Before asking how fast something is moving, you must first establish that it is moving. However, the value of qualitative data is often limited, and sooner or later, most methods develop toward the quantitative. This complicates the process of data collection, as calibration is a requirement for correct quantitative data - independently establishing the relation between the output values of the sensor and physical quantities. At the same time, the potential rewards and insights also become greater.

[7] In other words: Putting numbers and units on things.

The spectrum of quantitative to qualitative is of course also continuous. A good example of this is classification of fluorophores, as we will see later. This process includes collecting spectral data, creating a spectral composition of each fluorophore's emission spectrum, and using these to qualitatively distinguish between them. Even though the end result is often used in a qualitative manner, some quantitative meta-data exist in the form of spectral band intensities, which tells us something about the spectral composition of the fluorophore.

With regards to quantitative imaging, the main hurdles occur, at least in my experience, along the optical axis. No matter which way you spin it, an optical image is a projection of a three-dimensional space onto a two-dimensional surface, and for this reason we can never perfectly reconstruct all three dimensions. As the integration of intensity occurs along the optical axis, the positional information along this dimension is distorted or lost, creating a challenge when it comes to passive imaging methods. Depending on the application and method, this can sometimes be counteracted by for example introducing an active illumination scheme, but this is not always possible. However, accurate and relevant information can still be extracted from this type of data as well, as we will see in later chapters.

1.7 Introducing the work

During the last 5 years, I have been working with developing the imaging technique known as Frequency Recognition Algorithm for Multiple Exposures, or FRAME for short. FRAME is an image multiplexing method, allowing several exposures to be stored in a single camera image, and can be used in a variety of ways to expand the capabilities of a sensor. The main focus of my research has been using this method to create multispectral images, and developing methodology for how the resulting data can be used for quantitative measurements. For this reason, the thesis summary that you are now reading will have a strong focus on multispectral capture, as it constitutes the bulk of my work. That being said, I have also been working with FRAME and other coded light-approaches to obtain other forms of data, which will be covered at the end.

The content of this thesis is presented in 6 chapters, including the introductory one you have already read. Chapter 2, "Methods of spectral imaging", provides an overview of different techniques used to capture spectral images along with their general advantages and drawbacks, adding the context needed to understand the motivation of the research. Chapter 3, "FRAME", dives into the theoretical framework[8] needed to understand the workings of FRAME on paper, while Chapter 4, "Experimental

[8]My west coast origin could not resist.

aspects", provides practical "know-how" needed to actually implement it in a laboratory setting. Chapter 5, "Applications and results", discusses how the techniques have been applied for different measurement scenarios, including application-specific processing and similar, and highlights some of the results achieved. Finally, in Chapter 6, "Outlook and Conclusion", I give my view on the potential future of FRAME and finish with some concluding remarks. I hope that you will find this work both enjoyable and that it proves useful in your own endeavors!

Chapter 2

Methods of spectral imaging

Spectral imaging involves imaging a scene at multiple unique and preferably non-overlapping spectral bands. These are combined into a single multi-channel image, now containing one "layer" for each spectral band, commonly referred to as spectral channels. Such images provide spatially resolved spectral information useful in a wide array of applications and fields. Remote sensing applications such as satellite imaging [25] and agriculture monitoring [26] enables the monitoring of environments and vegetation over large areas, increasing the efficiency with which for example damage to crops can be identified. Within the biomedical field it provides tools for both the study of the microscopic world [27; 28] as well as provides tools for medical diagnosis [29], helping scientists understand the underlying biology of diseases and enabling better and more efficient treatment. Spectral imaging can also be used to non-intrusively study materials [30; 31] as well as provide insight into the composition and temperature of celestial bodies and flames alike [32–35]. The continued development of spectral imaging is one of the tools needed to continue to push our understanding of the world. Fig. 2.1 shows some of these application areas of spectral imaging.

This chapter discusses different techniques of spectral image capture. I have chosen to limit this to single-sensor techniques utilizing Focal Plane Arrays (FPAs), as including multi-sensor techniques is a bit like comparing apples and oranges. Of course, other solutions such as multi-sensor techniques [36] and point detectors [37] exist, each with their own advantages and drawbacks, but the line has to be drawn somewhere, and so this chapter will limit the comparison to single-sensor FPA techniques. I also want to clarify that this is not an all-encompassing review of such techniques, but includes the ones that I have found the most relevant for comparison purposes, either because they are commonly used, or because they are closely related to the technique I have been developing.

Figure 2.1: Three examples of applications where spectral capture plays a crucial role. Left: Spectral microscopy imaging in combination with fluorophore tagging help separate cellular structures. Middle: Multispectral remote sensing applications, such as satellite and aerial imaging, can be used for agricultural monitoring. Right: Telescopes with spectral resolution help us investigate the composition of celestial bodies lightyears away.

As briefly mentioned in the previous chapter, a spectral image can be described as a data cube with two spatial dimensions and one spectral dimension, as seen in Fig. 2.2. The volume of this data cube is comparable to a total "bandwidth" (composed of spatial and spectral bandwidths) of the spectral image. As mentioned in the previous chapter, techniques with a few spectral channels (typically less than 15-20) are referred to as multispectral imaging. When the number of channels increase beyond this point it is instead referred to as hyperspectral imaging, often containing hundreds of contiguous spectral bands [22; 23]. Assuming the same wavelength range is sampled, a higher number of spectral bands increases the spectral resolution, but also leads to worse signal-to-noise ratio (SNR) and/or spatial resolution per band [38]. Most hyperspectral methods also rely on the continuous dispersion of light, rather than optically splitting the light into separate optical channels based on wavelength, meaning the manipulation of individual spectral bands is considerably more difficult.

The discussed techniques for obtaining spectral images will be divided into two main categories - scanning techniques and snapshot techniques. Scanning methods gradually build the data cube in time using multiple camera images, while snapshot methods capture the entire data cube during the exposure time of a single camera image.

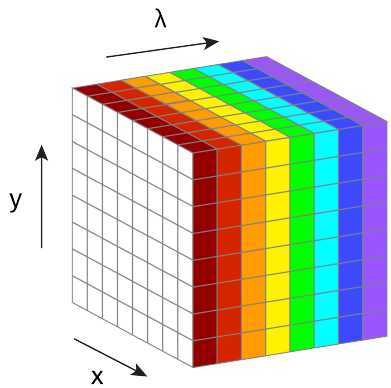


Figure 2.2: A spectral image can be described as a data cube, with two spatial dimensions (x, y) and one spectral dimension λ.

2.1 Scanning techniques

Techniques that use temporal scanning to acquire spectrally resolved images either sample different spectral bands of the entire scene at different times, or sample all spectral bands simultaneously but only for a subsection of the scene at a time. Regardless of which, the resulting image data cube is built up gradually along at least one dimension. The common thread in these techniques are good spatial and spectral resolution at the cost of temporal resolution - using several exposures to build the spectral data cube means more total bandwidth available (total volume of data cube), but also longer acquisition times.

2.1.1 Spectral scanning

As the name eludes, spectral scanning[1] refers to techniques that utilize a sequence of camera images to build the data cube along the spectral axis [39], as shown in Fig. 2.3. The spectral properties of each sub-image are determined by collectively changing the spectral composition of all the light reaching the sensor. A common and economical way of manipulating the spectral composition is to use mechanical filter wheels [40], which cycle through spectral filters in between image captures. Other solutions include tuneable spectral filters such as Acousto-Optic Tunable Filters (AOTFs) [41] and Liquid Crystal Tunable Filters (LCTFs) [42].

[1] Also referred to as "staring" imaging.

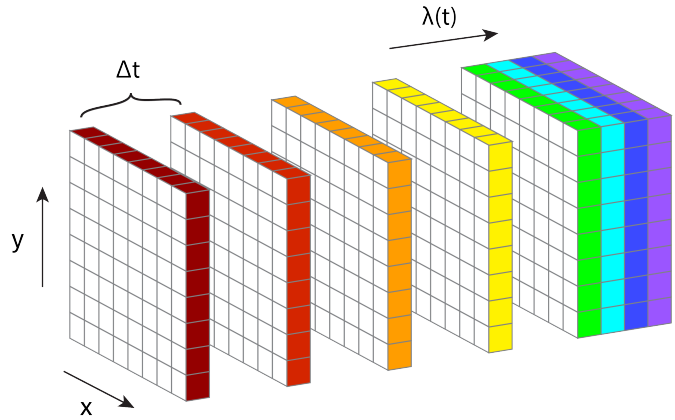


Figure 2.3: Spectral scanning uses multiple camera images taken in sequence to gradually build the data cube in the spectral dimension. Between each image, the spectral band reaching the sensor is altered, so that each camera image samples a unique spectral channel.

A major advantage of this technique is the utilization of the full resolution of the camera for every spectral channel, both in space and intensity. No two spectral bands have to share any aspect of the total bandwidth of the sensor [43]. Sequential capture is also very flexible regarding the spectral composition of each channel - spectral filters are easy to tailor to the specific application, and the process of both changing filters and capturing a sequence of images does not require complicated alignment or reconstruction algorithms - the data cube can be assembled by concatenation of the individual spectral channels, creating the spectral axis.

A fundamental assumption of the spectral image data cube is that each spectral channel depicts the same scene. The spectral resolution of spectral scanning methods is therefore limited by the characteristic timescale of the scene - if the scene changes during the capture of spectral channels, the spatial correlation is lost. This also limits the potential spectral resolution of the method, as spectral resolution scales reciprocally with frame rate. For a characteristic timescale T of a scene, over which no detectable movement occurs, the number of spectral channels N_{ch} of a single spectral image data cube is limited by

$$N_{ch} = \frac{T}{f}, \tag{2.1}$$

where f is the capture rate of the camera in frames per second. The solutions previously mentioned - mechanical filter wheels and LCTF such as Thorlabs Kurios© - typically operate on timescales on the order of 10-100 ms for wide spectral range solutions (wavelength ranges of several hundred nanometers) [44], limiting the spectral

image frame-rate f_{sp} (the number of complete spectral data cubes per second) of such setups to

$$f_{sp} \approx \frac{1}{N_{ch} \cdot 10 \text{ ms}} = \frac{100}{N_{ch}} \text{ [s}^{-1}\text{]}. \tag{2.2}$$

AOTFs are generally faster (µs timescale), but are limited to very narrow spectral bands on the order of 1 nm [45] and can introduce distortions in the image [46].

A common application of spectral scanning is fluorescence microscopy [47]. Combinations of filters are needed to spectrally separate the broad and often overlapping spectral emission bands of fluorophores, requiring spectral specificity and flexibility. At the same time, spatial resolution is of importance in order to spatially resolve microscopic features, as is ease of use. All of these factors combined makes spectral scanning a viable option. However, the long acquisition times associated with it can be problematic when imaging living samples [47], or when photo-bleaching [48] limits the total time that a tagged sample can be exposed.

2.1.2 Spatial scanning

Spatial scanning methods instead build the spectral data cube along one or both of its spatial axes, by spectrally resolving a subregion of the entire scene, as shown in Fig. 2.4. A common example is the push-broom scanner [39; 49], which utilizes spatial scanning with an FPA to image spectrally resolved image slices. For each scan, a narrow slice of the scene, corresponding to a single pixel column, is captured. Using a dispersive element such as a prism or a grating in combination with imaging elements (lenses and/or mirrors), the image column is spectrally resolved and mapped onto the second axis of the two-dimensional FPA, resulting in a single slice of the data cube. By subsequently scanning over the entire scene, moving the camera relative to the scene along the scanning dimension, the spectral data cube is completed.

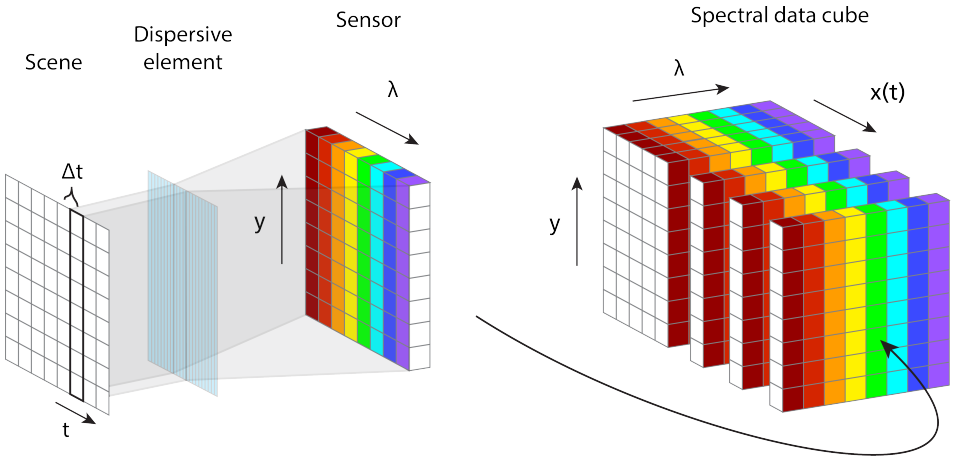


Figure 2.4: Spatial scanning uses a sequence of camera images to gradually build the spectral data cube in one of the spatial dimensions. Scanning a region corresponding to a single pixel column of the scene and dispersing it across one of the spatial axes of the sensor creates spatial slices of the spectral data cube.

Spatial scanning provides excellent potential for spectral resolution, as it utilizes one of the sampling dimensions of the FPA [50]. The maximal number of spectral channels is determined by the number of sampling points along the spectral axis of the sensor, while the total spectral range sampled (and thus the bandwidth of each spectral channel) is determined by the spectral dispersion in the plane of the FPA. Combined, these factors result in a dense and evenly sampled spectrum. The technique is also very stable once aligned, with no need for moving parts. It utilizes the full sensor resolution along the spatial axis of each sub-image, while the spatial resolution along the scanned axis is dependent on the synchronization of acquisition and movement. This puts extra emphasis on calibration and adds a post-process step in the form of "stitching" - the assembling of the second spatial axis. Spatial scanning also relies on the fundamental assumption that no detectable movement occurs over the scanning time. The scanning method of push-broom is akin to a rolling shutter [51] which sweeps over the field-of-view, reading pixels line by line until the entire field-of-view is captured. Any movement during this scanning results in spatial artifacts. This is very similar in nature to the limitations for spectral scanning, where the discrepancy in time will be visible along one of the spatial axes, with all spectral bands having an identically distorted scene.

An application example of this technique is remote sensing, such as satellite imaging [52]. Spectrally resolved satellite images are an indispensable tool in the monitoring of our environment and weather [53; 54]. As satellites move in fast and predictable orbits relative to the surface of Earth, spatial scanning is a clever solution taking advantage of the otherwise problematic relative movement between the sensor and the scene.

2.2 Snapshot techniques

Snapshot spectral imaging techniques captures the entire spectral data cube in a single camera image, circumventing the most prevalent issue of scanning techniques - the long acquisition times. However, this also means a lower total bandwidth (smaller data cube volume) is sampled compared to scanning techniques. The fast acquisition times of snapshot techniques, where all spectral bands are captured concurrently, makes them suitable for capturing transient events.

The technique of combining separate signals into a single one is called multiplexing [55]. It is a technique most often used to increase transmission efficiency when multiple signals need to be transmitted over a single medium or channel, and is achieved by uniquely encoding each of the constituent signals using a parameter of choice. After transmission of the combined signal, the individual constituent signals can be separated by the receiver if the encoding pattern is known. Multiplexing is used in a variety of applications primarily in communication technology, perhaps the most famous example being radio which utilizes both amplitude modulation (AM) and frequency modulation (FM) multiplexing [56]. These are two examples of coding schemes, divided based on what signal parameter is encoded.

Assuming a single monochrome FPA with no spectral sensitivity, snapshot multispectral imaging requires multiplexing - several unique signals (the individual spectral channels) must be captured simultaneously in a single camera image while maintaining separability. As the sensor is not spectrally sensitive, this property has to be encoded using some other parameter, trading another bandwidth for spectral bandwidth. The techniques covered in this chapter all use spatial bandwidth for this trade, albeit in different ways.

2.2.1 Spatial division multiplexing

This type of multiplexed spectral imaging divides spectral channels in the spatial dimension, either into continuous sub-regions or (more commonly) into mosaic patterns covering the entire sensor. As a direct consequence of this, each spectral channel is sampled fewer times in space, as the spatial resolution of the sensor is constant. However, the dynamic range is not shared between different spectral channels, and the post-processing required to produce the finished spectral data cube is most often less complicated than the other multiplexing schemes discussed later in this chapter.

Filter arrays

By mapping spectral channels onto different sub-regions of the FPA, all of them can be imaged onto the sensor and captured concurrently at the cost of a lower spatial sampling frequency. A common technique incorporating spatial division in this way is multispectral filter arrays (MSFAs) [57–59], which consist of a mosaic of pixel-sized spectral filters fixed to the sensor. Each spectral channel is then a sparsely sampled image covering the entire sensor in a periodic mosaic pattern, and the final spectrally resolved image is constructed through a demosaicking algorithm [60]. Fig. 2.5 shows the steps of creating a spectral data cube using an MSFA. The spectral properties of each channel are determined by the spectral filters, which are permanently fixed to the sensor during manufacturing. The most common example of filter arrays is the Bayer pattern used to produce spectral images emulating our own color vision. This pattern samples three color bands - red, green and blue (RGB) - roughly corresponding to the spectral sensitivity of the L-, M- and S-cells in our eyes [61]. However, the large spectral overlap of these three spectral channels make them less useful in many research applications, due to inter-channel dependency [62].

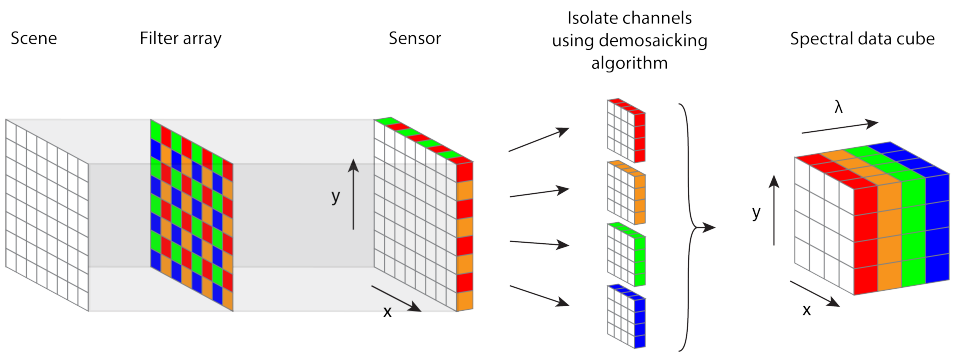


Figure 2.5: Filter array techniques use pixel-size spectral filters, which result in sparsely sampled spectral bands being captured in a mosaic pattern on the sensor. Each spectral channel is isolated in post using a demosaicking algorithm, and the spectral bands are concatenated along the spectral axis to create the data cube.

Filter arrays provide great stability and ease-of-use at the cost of spectral flexibility and discarded intensity. The calibration and demosaicking needed can be implemented at a manufacturing level to provide a ready-to-use solution for the end user. However, changing the spectral properties of the channels is therefore also complicated, if not impossible. As the vast majority of filter arrays are on-chip solutions, they are permanently integrated with the sensor at the manufacturing stage, preventing any spectral division adjustments after this stage. Solutions for modular filter arrays have been demonstrated [63], but these are still much less flexible than a camera-independent filtering solution. Filter arrays also discard inbound light not meeting the wavelength

criteria. Assuming perfect spectral filtering, with 100% transmission in the desired wavelength region and 0% outside, filter arrays still only transmit

$$\frac{I}{I_0} = \frac{1}{N_{ch}} \tag{2.3}$$

of the intensity as a consequence of filtering, where I_0 and I are the collective intensity before and after the spectral filtering respectively. In reality, spectral filters are of course not perfect, further increasing intensity losses. The mosaic patterns also impose geometric restrictions in the form of scalability, as spectral sample points need to be structured in a periodic fashion. This can cause spectral channels to have different sampling frequencies, either intentionally to increase sensitivity and spatial resolution of a certain band, or as a consequence of these restrictions.

Very similar in concept, Fabry-Perot filter arrays (FPFAs) uses Fabry-Perot etalons for spectral filtering [64; 65], incorporating the same mosaicing concept as MSFAs. These offer better optical throughput and spectral selectivity, at a greater cost and more complicated alignment process. However, the drawback in common for all filter array techniques is the spectral filtering loss illustrated in Eq. 2.3, as light is not re-directed based on wavelength.

Other spatial division-solutions

Commercially available image doubler units allows for splitting the sensor into two halves, projecting an individual image on each [66]. This can be combined with spectral filters to simultaneously image two spectral bands using the same sensor. However, optical complexity quickly increases with an increasing number of channels, making this type of solution limited to a low number of channels.

A multi-aperture solution [67], which instead use continuous (non-mosaicked) regions of the sensor to separate spectral channels, offers better spectral flexibility and transmission than filter arrays, but is optically complex to achieve. In such a solution, the light can be split spatially on the criterion of wavelength using dispersive elements or dichroic coatings, meaning losses due to spectral filtering is minimized. However, the optical routing of each channel requires alignment close to the sensor, and in the case of dispersive elements, there is also a spectral variation along at least one spatial axis in each spectral channel.

Lenslet Array Tuneable snapshot Imaging Spectrometer, or LATIS [68], uses a lenslet array to focus the light of image subregions down to lenslet spots, which are imaged on the sensor. As the light of each subregion is focused to a single, smaller spot, void

space is created between spots on the sensor. By introducing a prism into the optical path between the lenslet and the sensor, the lenslet spots are dispersed into the void space, granting spectral sensitivity in snapshot. The trade-off for this achieved spectral sensitivity is the downsampled spatial resolution (lenslet focusing) to achieve the void space needed for spectral dispersion.

In later years, a combination of micro-machined optical surfaces and machine learning have also emerged [69; 70]. These techniques use diffractive effects to separate spectral channels, leveraging the adaptive capabilities of neural networks to interpret the resulting images and reconstruct the individual spectral channels. Techniques such as these feature high transmission but rely on the reconstruction capabilities of the trained network. This reliance also makes changing the spectral properties of the setup cumbersome, as it requires additional training of the neural network.

2.2.2 Compressed sensing

Compressed sensing is a collective term for techniques which sample fewer points than the number of elements of the final data cube [71]. The underlying assumption of these techniques is that the sampled data is sparse in at least one domain, allowing for a solution to be found which uses this sparse representation. Spectral imaging based on compressed sensing therefore rely on more complicated reconstruction algorithms (compared to previously discussed methods) which enforce sparsity to isolate each spectral channel.

The most prevalent method of compressed sensing for spectral imaging is Coded Aperture Snapshot Spectral Imager [72–74], or CASSI for short. It uses a random binary pattern mask in combination with dispersive element(s) to capture a spectrally multiplexed camera image. Fig. 2.6 shows a schematic of the CASSI technique. The image of the scene is first spatially modulated using the binary mask, and the masked image is subsequently dispersed over one of the spatial axes of the sensor, creating an image of the scene on the FPA which is composed of sheared spectral channels. This sampled two-dimensional image $\mathbf{Y}$ can then be described as

$$\mathbf{Y} = \phi\mathbf{X} + \eta \quad , \tag{2.4}$$

where ϕ is the known encoding transformation (intensity modulation and spectral dispersion), $\mathbf{X}$ is the complete spectral data cube, and η is the noise [75]. The spectral data cube $\mathbf{X}$ can then be approximated by searching for the sparse representation which best represents $\mathbf{Y}$. The binary pattern ensures a known, structured pattern which is critical to reconstruct the spectral data cube both spectrally and spatially.

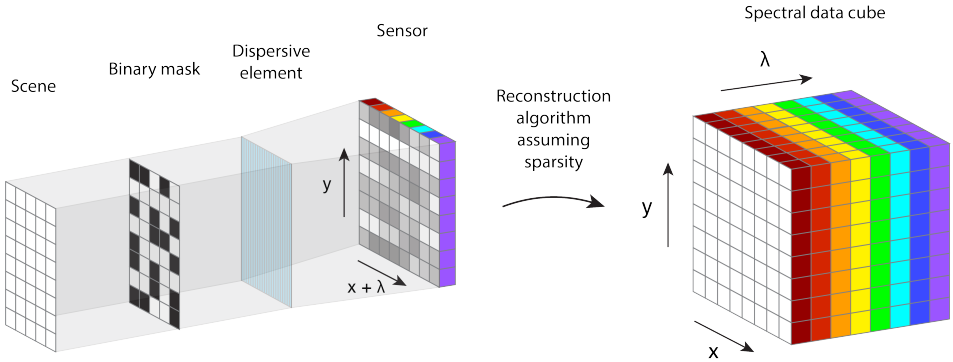


Figure 2.6: CASSI uses a binary mask and a dispersive element to create a camera image mapping both spatial and spectral information along one of its axes. The binary mask acts as a known reference in the reconstruction algorithm, which assumes sparsity in some domain of the image to calculate the solution which best agrees with the sampled image.

Much like spatial scanning, CASSI uses one of the dimensions of the sensor to encode spectral information, with the difference that CASSI also maps spatial information on the same axis. The number of spectral channels and the spectral resolution is determined by a combination of factors, including the number of pixels that the image is spectrally sheared over, the spectral dispersion at the FPA, and the size of the smallest feature in the binary mask [76]. The final spectral data cube has a spatial resolution corresponding to $\mathbf{Y}$ in the direction orthogonal to the spectral dispersion. In the direction of spectral dispersion, the spatial resolution of the reconstructed data cube is reduced by the number of pixels the shearing occurs over [75].

As mentioned, CASSI relies on sparsity in some domain of $\mathbf{X}$, making it compressible. However, the method with which the final data cube is reconstructed also means that CASSI is a product of both measurement and computational processing - some of the intensity values in the resulting spectral data cube are not physically sampled, but instead the result of a mathematical approximation. This results in the total bandwidth of the spectral data cube being higher than those of other snapshot techniques. The approximation is also computationally demanding, and the optical setup requires careful alignment and calibration. Nonetheless, studies have shown that the technique performs well, and a number of reconstruction algorithms have been developed [77].

2.2.3 Frequency division multiplexing

Instead of using the spatial domain to separate spectral signals in an image, the signals can be separated in the frequency domain as demonstrated by Gunturk and Feldman in 2013 [78]. This is known as frequency division multiplexing, and Frequency Recognition Algorithm for Multiple Exposures (FRAME) [79–84] - the main subject of the

research presented in this thesis - is an imaging technique belonging to this category of multiplexing techniques. Albeit few, other techniques utilizing the same multiplexing scheme for spectral capture have also been demonstrated. In 2018, Deng et al. demonstrated an imaging method combining frequency division multiplexing with a Digital Micromirror Device (DMD) and a spectral filter wheel to achieve hyperspectral snapshot imaging of 31 spectral channels [85]. The technique is reminiscent of a combination of a spectral scanning and snapshot - despite the technique being snapshot (capturing all spectral channels in a single camera image), all spectral channels are not synchronously sampled within the camera exposure due to the use of the spectral filter wheel.

As both the theoretical and experimental aspects of using FRAME for multispectral imaging will be dealt with in the coming chapters, this section will only provide a brief overview of the fundamentals of multispectral FRAME and its overarching advantages and drawbacks.

Multispectral FRAME utilizes an optical setup to uniquely modulate image intensities in a number of optical channels, corresponding to spectral bands. The images of each spectral channel are spatially overlapped on the FPA and all captured in the same camera image synchronously. Even though intensities of individual spectral channels are overlapped in space, they are separable in reciprocal space, as the spatial carrier frequencies used to modulate each spectral band isolate them from each other by shifting each to a different, higher frequency band. This phenomenon is illustrated in Fig. 2.7. Each individual spectral channel can then be isolated through sequential spatial frequency lock-in and frequency filtering [86].

The spectral properties of each channel are determined by spectral filters and/or coatings used to create the optical channels in the setup. In practice this means fewer, spectrally wider bands, as channels need to be individually aligned, but with the advantage of good spectral flexibility and easy access to those spectral components. Moreover, the technique is not limited to sampling contiguous spectral bands (a continuous range of wavelengths), but can isolate specific wavelength regions of interest and disregard intermediate spectral bands. As the method uses digital low-pass filtering or truncation to isolate exposures, it trades spatial resolution for spectral resolution just as all other techniques covered in this chapter, albeit during the post-processing instead of directly during the sampling. The post-processing required is computationally quick thanks to optimized algorithms and transforms in combination with an exposure only requiring four operations to be isolated - (1) a Fourier transform, (2) a shift of the Fourier domain (frequency lock-in), (3) digital filtering, and (4) an inverse Fourier transform. In addition, the intensities sampled in the final spectral data cube all originate from actual sampled intensities. As mentioned, the experimental aspects and post-processing will be covered in greater detail in subsequent chapters.

Multispectral FRAME has previously primarily been used for fluorescence imaging [79; 80]. The work in this thesis expands on that work, as well as introduces new applications and imaging setups used for quantitative measurements.

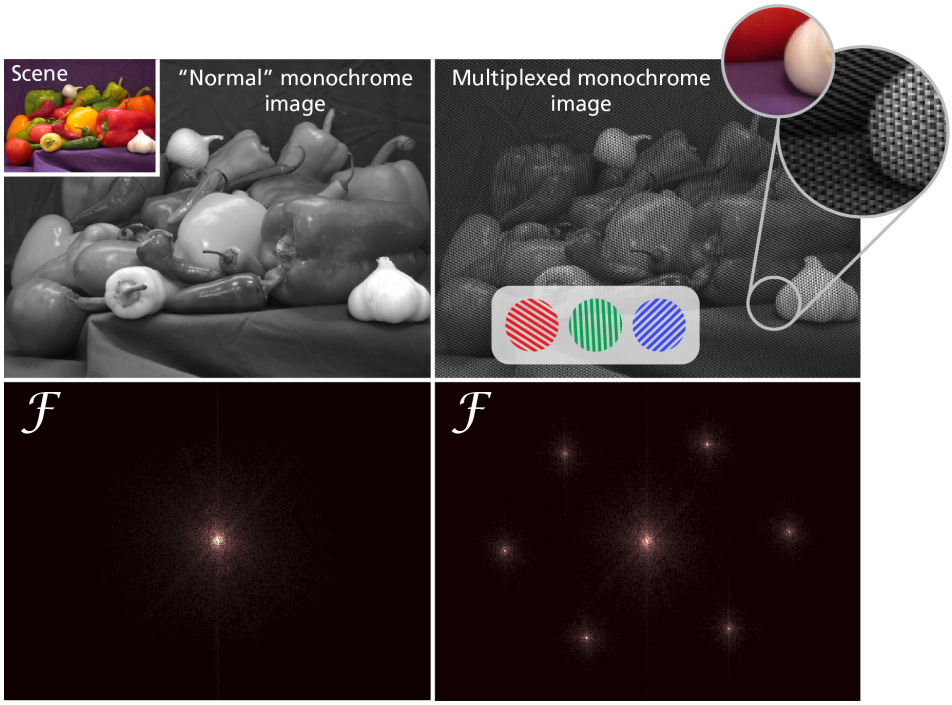


Figure 2.7: The fundamental principle of multispectral FRAME. A colorful scene (top left) is imaged by a monochrome grayscale camera. When no intensity modulation is applied to the scene, the Fourier spectrum of the image (bottom left) exhibits the characteristic shape of a natural image. By imparting a unique spatial carrier frequency onto the image in each spectral channel (top right, gray inset showing the frequency associated with each spectral channel), the frequency information of that image is shifted to a higher frequency band in frequency space (bottom right), isolating it from the other exposures and the central DC-component. This is how FRAME stores each exposure. The theoretical background for this will be discussed in Chapter 3.

Chapter 3

FRAME

"The most beautiful part of every picture is the frame."

- Gilbert K. Chesterton

A small foreword

I am personally not a fan of expressing ideas, whether complex or not, through mathematical notation. That is not to say that I do not think mathematical notation has its place within science - it is the best way we as a species have conceived of to express complex ideas in a compact and efficient way. However, for the same reason, it can also be quite cryptic, and it often takes time to fully understand the underlying message. It is my belief that if an idea can be expressed in words in a pedagogical and fairly efficient way, that is very often preferable, even if it takes up more space on the paper. As I want this thesis to be a rather light read, and not a book where it takes an hour to understand the implications of a single line, I have and will continue to do my best to express my thoughts and discussions in words rather than equations. However, at some point it becomes impossible to continue the discussion without bringing equations into the mix, as the subject of my work is, at least in part, inherently mathematical[1]. That point has now been reached. Brace yourselves[2].

[1]Despite my best intentions.

[2]In all honesty, these words of encouragement are mostly aimed at myself.

3.1 Space and frequency - two interpretations of the same image

This chapter will discuss subjects related to frequency analysis of images. By far the most common way of displaying images are in the spatial domain, mapping intensities to positions. This is how we normally display images for example in our homes, and what most people mean when they think of an image, as it replicates our own sense of vision and allows us to quickly identify and locate objects in the image. However, if the goal is to examine periodic patterns in an image, it is beneficial to examine the image in the frequency domain instead, where the same image information is not expressed in spatial components, but in frequency components instead. While the spatial domain allows us to quickly *locate*, the frequency domain allows us to examine properties such as the proportion of details to structures over the entire image.

When an object is imaged onto a digital sensor, it is sampled in the spatial domain, mapping intensity to position. Therefore, in order to examine the frequency representation of the image, this digital image must first be transformed to its frequency spectrum. This is done using the Fourier transform, which (without a doubt) deserves its own heading.

3.2 The Fourier transform

In the beginning of the 19th century, Joseph Fourier lay the ground works for what would become known as Fourier analysis [87]. The fundamental realization of Fourier was that by varying the amplitude, frequency and phase of an infinite amount of individual sinusoids and summing them, any periodic function could be created. The series of sinusoid functions became known as the Fourier series of the created function, and was proposed as a means of solving the heat equation. Later work expanded on this idea, showing that this expansion was not only possible for periodic functions, but for a wide range of functions that are absolutely integrable and piece-wise continuous [88]. The Fourier transform was born - enabling functions of time or space to be described in terms of their frequency components.

The Fourier transform is essentially a change of basis functions. A digital grayscale image consists of two spatial dimensions which uniquely identifies the position of each image element (pixel) in the image, and associated with each pixel is an intensity value. However, the image can also be described as a sum of unique layers, where each layer is a two-dimensional sinusoid of intensity, with a unique combination of frequency, amplitude and phase. Instead of describing the image as a function of

space, as when considering individual pixels, it is now described as a function of frequency components acting as orthogonal basis functions. Fig. 3.1 illustrates these two representations of a grayscale image. When summing along the intensity dimension to create the complete image, all frequency components contribute to the intensity in a single pixel, in contrast to each pixel being independent as in the spatial interpretation. The proportion of these frequency components are the Fourier spectrum of the spatial image, and the Fourier transform and its inverse are the mathematical transforms that perform the switch between basis functions, from spatial components to frequency components and back.

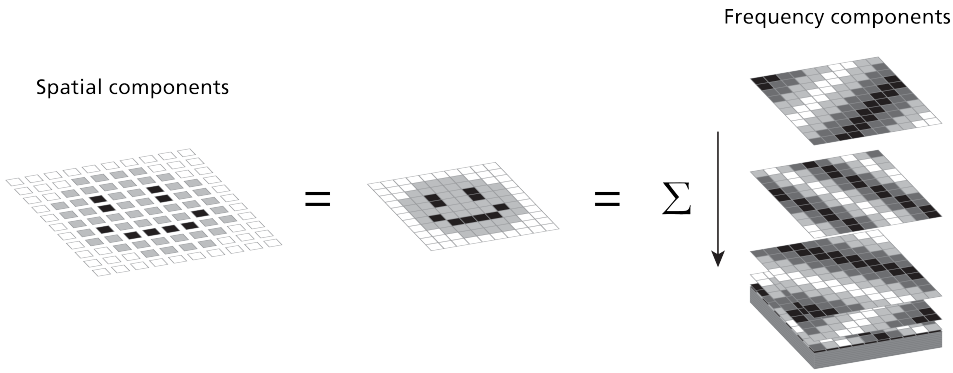


Figure 3.1: A digital grayscale image can be represented in terms of spatial components (pixels) or frequency components (layers of sinusoids). In the case of frequency components, each layer contributes to the intensity in each pixel, meaning a summation of the layers in the intensity dimension is performed.

By this point, it is time to formally introduce the Fourier transform. However, before discussing the discrete version of the Fourier transform, the continuous version will be presented. This first part will also be limited to functions of a single dimension, but the same principles apply in a multi-variable case. The Fourier spectrum $\hat{f}(\xi)$ for a real, continuous function of space $f(x)$ can be expressed as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \cdot e^{-i2\pi\xi x}\, dx, \quad \forall \xi \in \mathbb{R}. \tag{3.1}$$

In other words, for each frequency ξ, the product between our original function $f(x)$ and the complex exponential, which is a function of ξ, is calculated over the entirety of space. Substituting using Euler's formula,

$$e^{-i2\pi\xi x} = \cos(-2\pi\xi x) - i\,\sin(-2\pi\xi x), \tag{3.2}$$

shows that two factors are in fact included, one real and one imaginary, meaning $\hat{f}(\xi)$ is complex. This leads to our final interpretation of the Fourier spectrum of $f(x)$ - a complex function consisting of the relation between $f(x)$ and cosine (real) and sine (imaginary) functions for a range of frequencies. The phase information of each frequency is given by the proportion of cosine and sine coefficients. From the Fourier spectrum, the amplitudes of the individual frequency components can then be calculated by

$$A(\xi) = |\hat{f}(\xi)|. \tag{3.3}$$

The change of basis function, from amplitude of positional coordinates $A(x)$ to amplitude of frequency components $A(\xi)$, is thus completed. In order to revert back to the spatial representation of the image, an inverse Fourier transform can be performed, reversing the change of basis function through a similar integral across the entirety of the frequency domain[3],

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) \cdot e^{i2\pi\xi x} d\xi. \tag{3.4}$$

3.2.1 The Discrete Fourier transform

As this thesis concerns itself with discrete (sampled) functions in the form of digital images, no more time will be spent on the continuous Fourier transform. From this point on, the term "Fourier transform" will instead refer to its discrete version, the DFT. The key difference in this transition is that the original function, that is the signal being sampled, is now sampled over a finite domain at a certain frequency in a total of N sampling points. This means the original function of space is now expressed in terms of sampled points, $f(x_n)$, where $n = 0, 1, 2, 3, ..., \ N - 1$. Replacing the integral from Eq. 3.2 with the discrete sum and expressing the equation in terms of this sampling gives

$$\hat{f}(k) = \sum_{n=0}^{N-1} f(x_n) \cdot e^{-i2\pi \frac{kn}{N}}, \quad k = 0, 1, 2, ..., N - 1 \quad , \tag{3.5}$$

where $\hat{f}$ is now expressed as a function of wave number k. In other words, the original signal $f(x_n)$ is broken down into frequency components corresponding to standing

[3]Omitting any normalization constants.

waves over the sampling space, where each frequency component completes exactly k periods over the sampling domain (illustrated in Fig. 3.3 in the coming section). Expanding this for a two-dimensional signal of space corresponding to a digital grayscale image, with second dimension y sampled M times ($m = 0, 1, 2, 3, ..., M-1$), gives

$$\hat{f}(k_x, k_y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(x_n, y_m) \cdot e^{-i2\pi\left(\frac{k_x n}{N} + \frac{k_y m}{M}\right)} \tag{3.6}$$
$$k_x = 0, 1, 2, \ldots, N-1$$
$$k_y = 0, 1, 2, \ldots, M-1.$$

Fig. 3.2 shows a digital image of a forest and its Fourier spectrum. $k = 0$, known as the DC-component or 0-frequency (corresponding to an offset), is located in the center of the Fourier spectrum[4]. Moving outward from this point, the frequency increases towards the edges, approaching the highest frequency that can be uniquely sampled by the sensor in a certain direction - the Nyquist frequency - which will be the subject of the next segment.

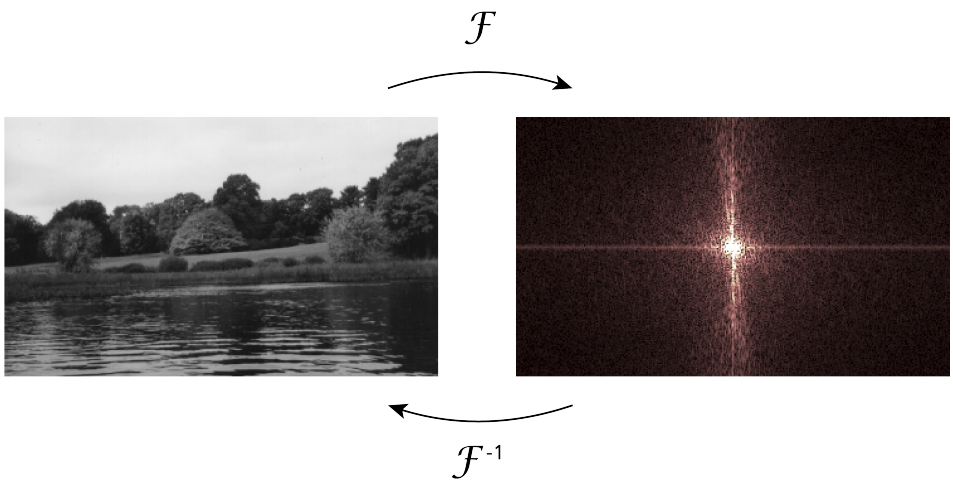


Figure 3.2: A natural grayscale image (left) and the absolute value of its Fourier spectrum (right).

[4]This is due to using a shift-function, making the spectrum easier to interpret. The raw output of the DFT is formatted as described by Eq. 3.6, in rising order of k

3.2.2 Consequences of sampling

The Shannon-Nyquist theorem states that in order for a signal to be perfectly reconstructed, it must be sampled at a frequency twice that of the signal's highest frequency component [89]. Thus, for a given sampling frequency f_s, there exists a highest resolvable frequency called the Nyquist frequency,

$$f_N = \frac{f_s}{2}. \tag{3.7}$$

From an imaging perspective, this means that the highest resolvable frequency in a single image dimension of size N is one which completes $k = N/2$ periods over the sampled space, corresponding to 2 pixels per period. For higher frequency components, such that $k > N/2$, the sampling will be affected by frequency folding, where a high frequency will instead appear as a lower one [90]. Fig. 3.3 demonstrates this effect by showing the sampling at the Nyquist frequency and its two adjacent frequencies, f_N and $f_{N\pm1}$. For f_{N+1}, the resulting sampled signal is the same as for f_{N-1}. This phenomenon, where a higher frequency is masquerading as a lower one, is known as aliasing. Following the same logic for all $k = (N/2)+1, (N/2)+2, ..., N-1$ explains the symmetry of the absolute value of the Fourier spectrum in Fig. 3.2, as all frequency components higher than f_N are folded. Something worthy of notice here is that the Nyquist frequency along the diagonal is higher than in both x and y, as the diagonal utilizes sampling in both x and y.

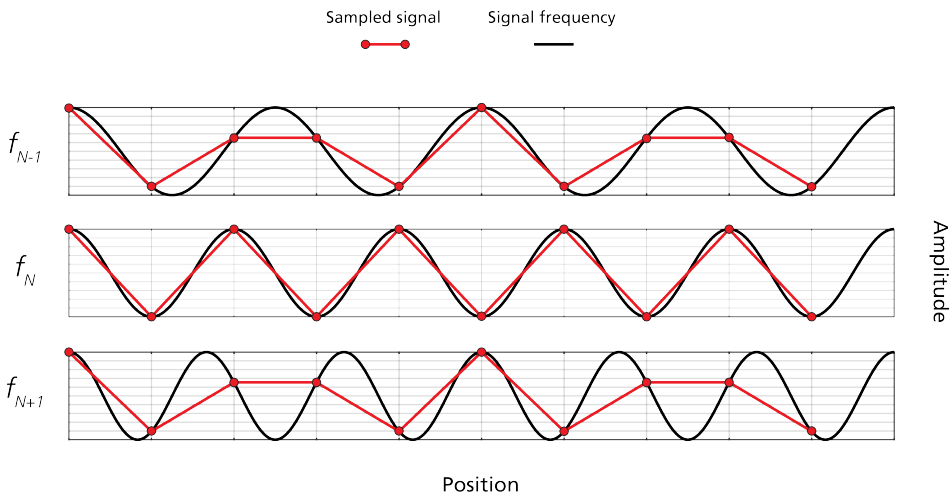


Figure 3.3: Sampling of three spatial frequencies around the Nyquist frequency f_N at a constant sampling rate. Due to frequency folding, the sampled signal of f_{N-1} and f_{N+1} appear identical.

As captured images are rarely band-limited, meaning sampled optical images most often contain higher spatial frequencies than the Nyquist frequency of the sensor, aliasing frequently appears in images. A common example of this is sharp edges, such as those of a square wave, which in theory require an infinite combination of high frequency-components to be perfectly sampled (assuming a perfect edge)[5]. A two-dimensional square wave with fundamental frequency $f_1 = k_1/2\pi$, as well as its Fourier spectrum can be seen in Fig. 3.4. As a square wave is composed of only odd sine terms, the harmonics (integer frequency multiples of the fundamental frequency) are located at $[k_3, k_5, k_7, ...]$. Frequency folding causes components higher than f_N, such as harmonics $\geq k_5$, to fold back into the Fourier domain as indicated by the horizontal dashed line.

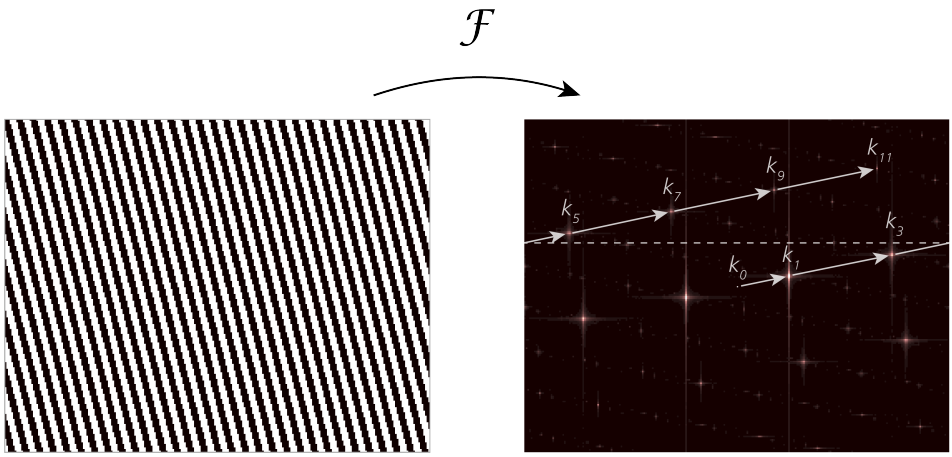


Figure 3.4: Image of a two-dimensional square wave (left) and the absolute value of its Fourier spectrum (right). The frequency folding of the higher-order harmonics have been annotated in the Fourier spectrum.

3.2.3 Natural images

As the name suggests, natural images are a subclass of images depicting real-world scenes. Due to the structure of such scenes, natural images has certain characteristic properties. Intensity levels across the image most often change gradually due to a combination of natural lighting and continuous surfaces, resulting in spatial correlation between many pixels. This results in a much sparser Fourier representation while maintaining most of the information in the image, something which is utilized in image compression [91]. Another property of natural images is scale invariance, and the associated characteristic shape of their Fourier spectrum.

[5]More on this later in the chapter. For a visual example of the Fourier series of a square wave, see Fig. 3.10.

Statistically, natural images exhibit similar proportions of low frequency structure to high frequency detail at different spatial scales [92]. This is the meaning of their scale invariance - nature has no preferred spatial scale. Details and structures exist in roughly the same proportion no matter the scale, as a sort of statistical fractal. A consequence of this is that most of the information in an image is contained in the lower frequencies, with higher frequencies containing less and less of the total energy of the image. This statistical tendency shows up as a characteristic spectrum shape in the form of a power law - the amplitude of the frequency components of natural images scale approximately as $1/f$ [93]. This can be put in perspective by comparing them to other image types. Fig. 3.5 shows the Fourier spectrum for a natural image (a tree in my garden) exhibiting the characteristic $1/f$-shape and an artificial image depicting different shapes of varying intensity. The latter is an example of a scale variant image - its Fourier spectrum changes drastically when zooming in or out of the image, while the Fourier spectrum of the tree maintains the same approximate $1/f$-structure regardless of zoom-level.

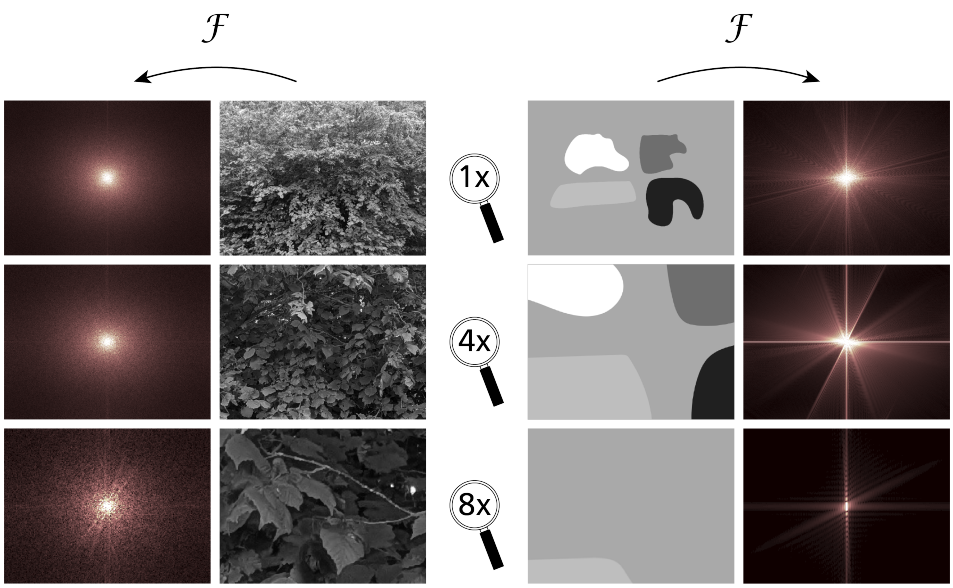


Figure 3.5: Comparison of a natural image and an artificial image at different spatial scales, and their respective Fourier spectra. **Left:** A natural image (sampled intensities of a real scene), depicting the greenery of a tree, at three different spatial scales (levels of zoom), along with the Fourier spectrum at each scale. **Right:** An artificial image (drawn digitally), at the same levels of zoom, along with Fourier spectra for each scale. The images have identical resolution. Digital zooming results in the resolution decreasing for each level of zoom.

3.3 FRAME in a nutshell

By this point, it is time to introduce FRAME - Frequency Recognition Algorithm for Multiple Exposures. As briefly mentioned at the end of Chapter 2, FRAME is an imaging technique built around the concept of spatial frequency multiplexing, and uses coded light to capture multiple individual exposures in the same camera image. The result is an image data cube containing at least one extra dimension apart from the two spatial dimensions, like the spectral data cube example seen in Chapter 2 where the extra dimension was the spectral axis. FRAME allows the encoding of dimensions such as time, wavelength bands and states of polarization [79; 80; 84] in a single camera image by trading the spatial bandwidth of the sensor [43].

As discussed in Section 3.2.3, natural images have sparse high-frequency bands. This is the fundamental fact that FRAME utilizes by shifting images copies into different regions of the empty "frequency real estate". This is done by modulating their intensities using spatial carrier frequencies, a technique known as frequency division multiplexing (covered in the previous chapter). In this way, several exposures can be captured and stored in the same camera image, overlapped in the spatial domain but separated in the frequency domain. Fig. 2.7 in Chapter 2 demonstrated this - it is hard to visually distinguish between them in the raw image captured by the camera, but the information of each exposure is isolated in the frequency domain.

These *sidebands* each contain the information from an individual exposure, and by isolating that information, the individual exposure can be extracted. In order to do so, a technique known as frequency lock-in is used [86]. This removes the carrier frequency and isolates the underlying intensities by shifting the carrier frequency to the center of the Fourier spectrum and filtering around it. Fig. 3.6 shows a schematic of the FRAME process - exposures representing different spectral bands are encoded, leading to them being stored in different frequency bands of the Fourier spectrum of the FRAME-multiplexed image. Through lock-in analysis, they can be extracted in post.

The physical meaning of the extra dimension in the image data cube depends on what parameter is encoded. By exposing a scene using temporally separated pulses, where the light of each pulse has been encoded with a unique spatial frequency, a video can be created from a single camera image, meaning the extra dimension is time. This is an example of "active" FRAME-encoding, where the illumination source is encoded and interacts with a scene. "Passive" FRAME-encoding instead encodes the light emanating from a scene, by splitting it up according to the parameter of encoding. This is how multispectral FRAME works - the light is split into optical channels where the image of each spectral band of interest is encoded using a unique carrier frequency.

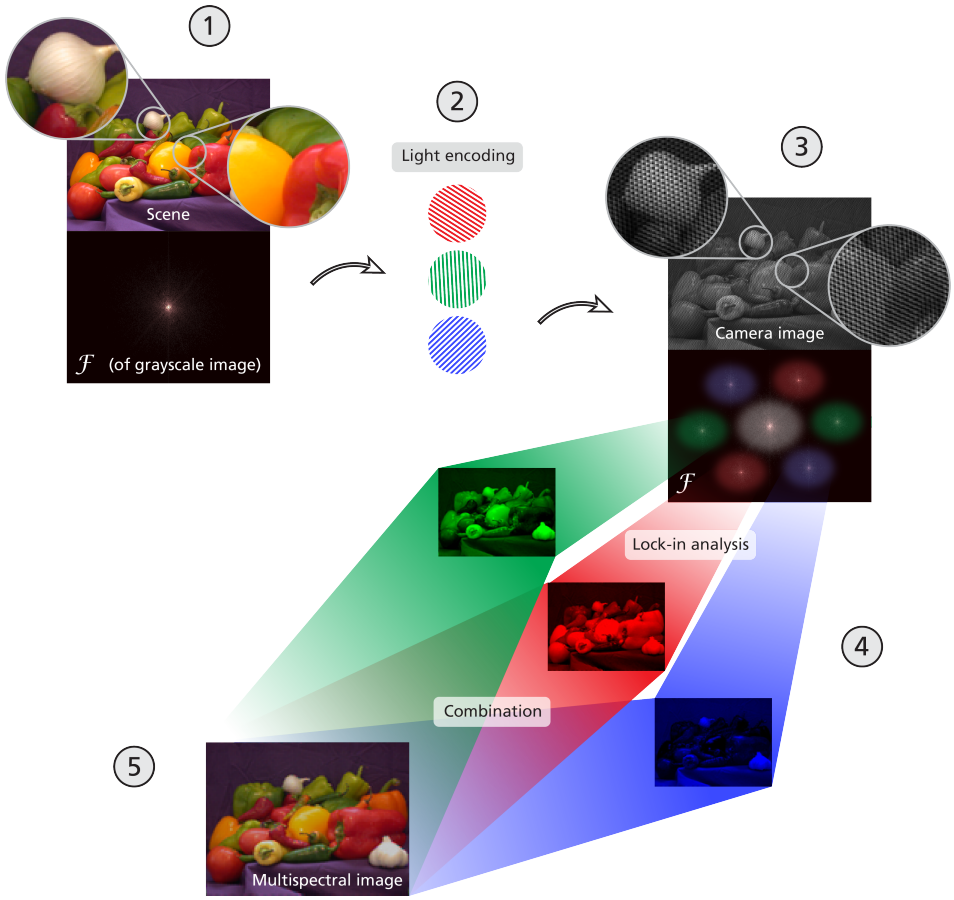


Figure 3.6: The process of multispectral FRAME, divided into five steps. **1.** A natural scene depicting peppers on a table, whose Fourier spectrum is also shown (produced from a monochrome capture), is being imaged by a monochrome camera. **2.** Before the light reaches the camera, it is split into optical channels and encoded with a unique intensity modulation in the form of a spatial frequency. All channels are then overlapped and imaged simultaneously. **3.** The raw camera image is captured, exhibiting the modulation frequencies. Zooming in and comparing with the original scene, the encoded spectral composition can be discerned. The white garlic contains patterns from all three channels, while the red bell pepper almost exclusively contains one pattern. When analyzing the Fourier spectrum of the image, the effect of the spatial modulation becomes evident - the information in each spectral band has been shifted to a separate frequency band, called sidebands. **4.** By removing the carrier frequency through frequency lock-in and filtering around it, the information in each sideband can be isolated. **5.** Through combination, the scene can then be reconstructed in color.

3.4 The mathematics of FRAME

Modulating the intensity of an exposure I using a spatial frequency modulation pattern m (with no phase-shift) can be expressed as

$$I_{mod}(x,y) = I(x,y) \cdot m(x,y) = I(x,y) \cdot \frac{1}{2}(1 + \cos(2\pi(f_x x + f_y y)), \quad (3.8)$$

where f_x and f_y are the spatial frequencies in x and y. This creates copies of the exposure in the frequency band around the carrier frequency, as all exposure frequencies are now interacting with the carrier. In the frequency domain, with the Fourier spectrum $\tilde{I}_{mod} = \mathcal{F}(I_{mod})$, the point-wise multiplication of Eq. 3.8 is equal to

$$\tilde{I}_{mod}(u,v) = \frac{1}{2}\tilde{I}(u,v) + \frac{1}{4}e^{i\phi}\tilde{I}(u - f_x, v - f_y) + \frac{1}{4}e^{-i\phi}\tilde{I}(u + f_x, v + f_y), \quad (3.9)$$

where u and v are the coordinates in the frequency domain. In this expression, the image copies at $\pm(f_x, f_y)$ are clearly visible, as well as their amplitude relative to the original image information. By varying the carrier frequency, the exposures can be "placed" anywhere in the Fourier spectrum of the camera image, ideally evenly spread out in the sparse frequency bands. The final multiplexed camera image I_{mux}, containing N individual exposures indexed as $[n = 1, 2, 3, ...]$, can then be expressed as

$$I_{mux}(x,y) = \sum_{n=1}^{N} I_n \cdot m_n. \quad (3.10)$$

Depending on what quantity is encoded, the exposures can either be incident on the sensor overlapped in time, such as for snapshot applications, or spaced out in time over the exposure time of the sensor, enabling temporal multiplexing. No matter the capturing method, retrieving each individual exposure I_n is performed in the same way - using spatial frequency lock-in. By multiplying the image by the complex exponential of the negative carrier frequency,

$$I'_n(x,y) = I_{mux}(x,y) \cdot e^{-i2\pi(f_{x,n}x + f_{y,n}y)} \quad (3.11)$$

its Fourier spectrum is shifted and centered around the carrier frequency $(f_{x,n}, f_{y,n})$,

$$\tilde{I}'_n(u,v) = \tilde{I}_{mux}(u,v) * \delta(u + f_{x,n}, v + f_{y,n}), \tag{3.12}$$

after which a truncation or digital low-pass filter can be applied to the Fourier domain to isolate the individual exposure,

$$\tilde{I}_n^{\text{filtered}}(u,v) = \tilde{I}'_n(u,v) \cdot \exp\left(-\left[\left(\frac{u}{\sigma_u}\right)^2 + \left(\frac{v}{\sigma_v}\right)^2\right]^N\right), \tag{3.13}$$

where the exponential describes a super-Gaussian filter of order N with standard deviation $[\sigma_u, \sigma_v]$. This process is repeatable for each exposure I_n, starting from Eq. 3.11. Fig. 3.7 visualizes Eqs. 3.9-3.13 in frequency space, for an image containing two multiplexed exposures.

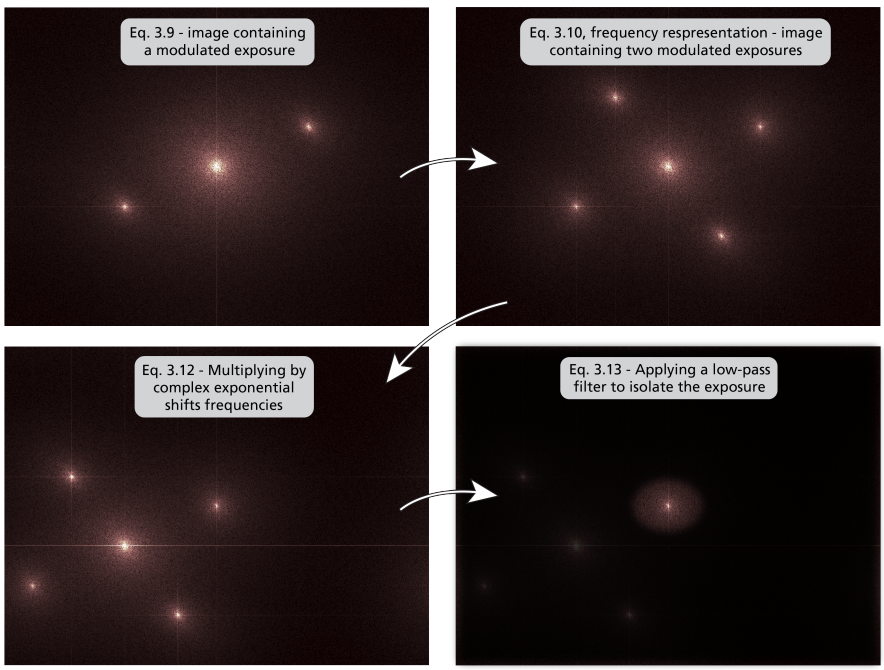


Figure 3.7: An image is multiplexed using spatial carrier frequencies, creating sidebands containing the information of individual exposures. By multiplying with the complex exponential of the negative carrier frequency, one of the carrier frequencies is shifted to the center. This allows for the isolation of the frequency information around the carrier by low-pass filtering, and the recreation of the extracted exposure through the use of an inverse Fourier transform.

Final resolution of isolated exposures

When a digital low-pass filter such as a super-Gaussian with a given radius is used to isolate an exposure, the spatial dimensions of the output will not reflect the actual resolution of the exposure after extraction. As such an operation does not change the dimension sizes, but only sets the amplitude of frequencies higher than the cut-off value (radius) to 0, the resulting image after transforming back to the spatial domain will be oversampled. The actual resolution of the image is given by the diameter of the filter. This leads to artifacts, such as the effect I have come to call "bag of worms", seen in Fig. 3.8, where spatially uncorrelated noise is oversampled, inferring a spatial pattern where there is none. On the other hand, the smoother gradients of the oversampled image can often be beneficial when building a qualitative understanding of the on-going process.

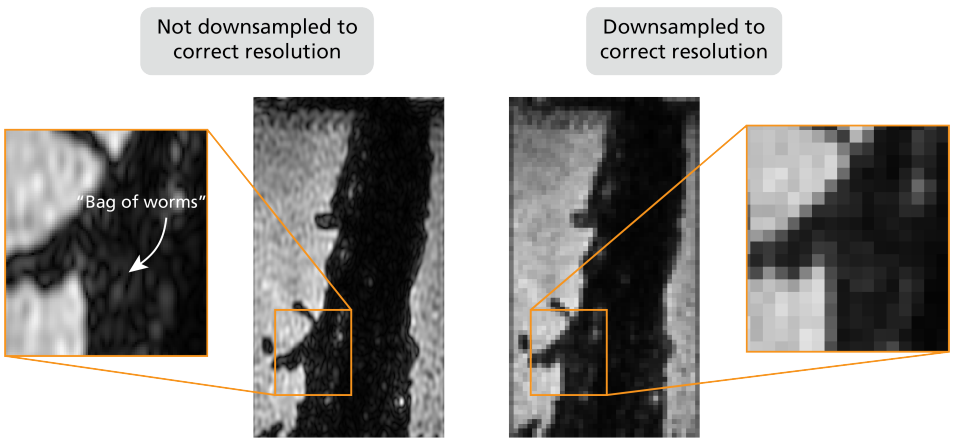


Figure 3.8: FRAME-exposure from shadowgraphy imaging of a spray injection, where the "bag of worms"-effect is visible in the zoom-in of the left, oversampled image. For comparison, a downsampled version of the same image is displayed on the right, which does not feature the artifact.

Instead of low-pass filtering the Fourier domain, it can instead be truncated (cropped) around the new 0-frequency (after lock-in). This leads to the dimension sizes after an inverse Fourier transform directly reflecting the actual spatial resolution. If the truncation is performed in the same aspect ratio as the original image, the resolution of the resulting image is downsampled equally in all dimensions, including the diagonal.

On a related note, filter/truncation shapes affect the final resolution of the extracted exposures. Fig. 3.9 illustrates filter types of different shapes. A rectangular filter will utilize the extra resolution resulting from $\sqrt{k_x^2 + k_y^2}$, while circular filters discard this diagonal resolution. However, a circular filter shape allows for a higher packing density in the Fourier domain, which is relevant if many exposures are to be multiplexed

without interference between exposures. An extension of this is the hexagonal filter shape, which results in the tightest packing ratio. Exposures can also be placed in the Fourier domain to increase resolution in a certain dimension when combined with an elliptical filter [94]. This increases resolution in the direction of the semi-major axis of the ellipse, and was utilized by Kornienko et al. in 2024 to help increase the temporal resolution of multiplexed exposures from a streak camera.

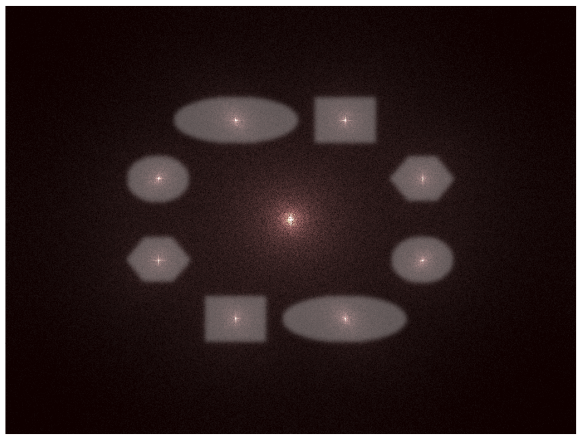


Figure 3.9: Different filter types and their benefits. A rectangular filter shape in the same ratio as the aspect ratio of the image downsamples the resolution equally in all directions, but makes for a somewhat unpractical shape when more than a few exposures are to be multiplexed. Circular and hexagonal filters allow for a larger packing ratio and less overlap between filters, which is beneficial when many exposures are to be multiplexed, but sacrifices resolution in certain directions. An elliptical filter increases the spatial resolution of an extracted exposure in the direction of its semi-major axis, which can be beneficial in special cases.

Cross-talk

Cross-talk in FRAME occurs when frequency bands of exposures overlap, and results in image information from one exposure showing up as stray frequencies in another, and vice versa. Needless to say, this deteriorates the quality of the multiplexed exposures, and is therefore best minimized. Separation of carrier frequencies, both from each other and from the DC-component, is the first step in avoiding this, but the problem is further increased by aliasing. Sharp edges in the image will lead to "streaks" in the Fourier transform for each exposure, complicating the placement in the Fourier domain. Vertical streaking of this type is visible in the Fourier spectrum of Fig. 3.2, caused by the sharp contrast edges in the waves. Aliasing combined with

the harmonics created from square wave modulation further complicates placement, as image copies are folded and can quickly fill up the available frequency bands, as seen in Fig. 3.4.

Modulation depth

As FRAME is a lock-in method, the signal strength of each multiplexed exposure is dependent on modulation depth - the contrast in the modulation pattern, where a perfect modulation depth modulates the signal with a factor varying between 0 and 1. Eq. 3.8 assumes such a perfect modulation depth, resulting in each created copy containing a 16th of the original image power (squaring Eq. 3.9). This corresponds to the theoretical maximum signal strength for an image modulated by a sinusoid. For an imperfect modulation depth $\mu = [0, 1)$, Eq. 3.8 instead becomes

$$I_{\text{mod}}(x, y) = I(x, y) \cdot \left[\frac{1}{2}\left(1 + \mu \cdot \cos\left(2\pi(f_x x + f_y y)\right)\right)\right], \tag{3.14}$$

or in frequency terms,

$$\tilde{I}_{\text{mod}}(u, v) = \frac{1}{2}\tilde{I}(u, v) + \frac{\mu}{4}e^{i\phi}\tilde{I}(u - f_x, v - f_y) + \frac{\mu}{4}e^{-i\phi}\tilde{I}(u + f_x, v + f_y). \tag{3.15}$$

In other words, the signal strength of the image copy scales linearly with modulation depth μ. Accounting for differences in modulation depth, both locally per exposure and globally between exposures, is therefore essential in order to correctly replicate the original image information of the exposure, especially in the context of intensity-based quantitative measurements. The implications of this in an applied context will be discussed in the next chapter.

Modulation shape

As previously stated, the maximum amplitude of each sideband modulated using a sinusoid is $1/4$ (Eq. 3.9). However, the efficiency of such a frequency shift can in fact be increased by using a square wave modulation. Modulating an image using such a waveform gives the expression

$$\tilde{I}_{mod}(u,v) \approx \frac{1}{2}\tilde{I}(u,v) + \frac{1}{\pi}e^{i\phi}\tilde{I}(u-f_x, v-f_y) + \frac{1}{\pi}e^{-i\phi}\tilde{I}(u+f_x, v+f_y) \tag{3.16}$$
$$+ H.O.H.,$$

where H.O.H. are the higher-order harmonics. This is analogous to Eq. 3.9, but using a square wave modulation instead of a sinusoid, which increases the amplitude of the first sideband as $1/\pi > 1/4$. The difference in amplitude of the frequency-shifted image copies can be explained by considering the Fourier series of a square wave, seen in Fig. 3.10. As previously mentioned, the Fourier series of a square wave is a series of odd sine-functions. Of particular note, the fundamental frequency component of the square wave has a larger amplitude than the square wave resulting from the combination of all frequency components. Comparing this to a sinusoid carrier of the same frequency and amplitude as the square wave, this results in the square wave modulation shifting more of the image information into the first sideband around the fundamental carrier frequency, in turn resulting in a potential signal increase of $\sim$27% ($1/\pi$ compared to $1/4$). However, the magnitude of this signal increase is affected by the sampling, meaning that a 27% increase is the ideal case. The signal increase also comes at the price of the aforementioned harmonics potentially filling up the Fourier domain and overlapping with the image frequency band around the fundamental carrier frequencies.

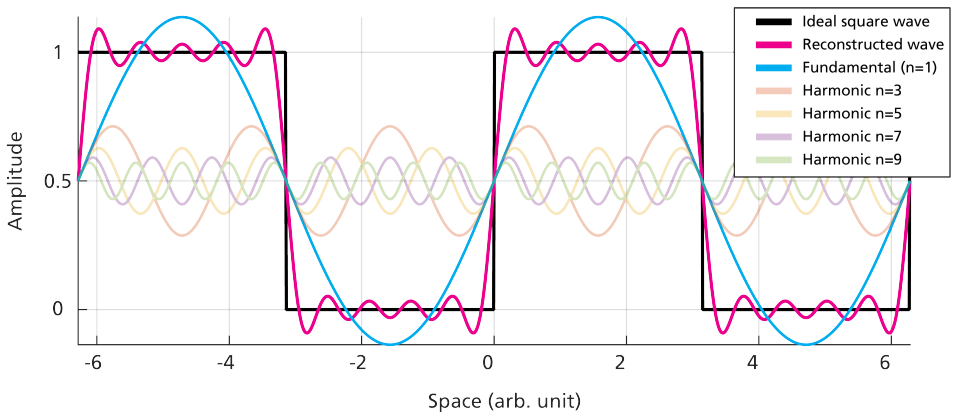


Figure 3.10: The first five terms of the Fourier series of a square wave. The fundamental frequency component (blue) has an amplitude exceeding that of the ideal square wave (black). As more terms are included, the reconstructed wave (magenta) will approach the ideal square wave.

Chapter 4

Experimental aspects of multispectral FRAME

"Theory and practice are the same in theory, but not in practice."

- Scientific adage of unclear origin

This chapter discusses the practical implementation of FRAME - how to optically encode the image information, how to calibrate for effects introduced by the encoding, and some general insights resulting from experience - with a strong focus on the aspects of multispectral encoding.

4.1 Camera considerations

The first and most fundamental thing you are going to need for encoding images is, *surprise*, a camera. FRAME works with practically any camera - there is no limiting factor in sensor technology or similar, other than that the sensor needs to be a focal plane array as FRAME relies on optical encoding of the image and subsequent digital processing. That being said, FRAME benefits more from some sensor properties than others, due to the nature of spatial frequency encoding. Building on the discussions of the previous chapter, the following interactions between sensor parameters and FRAME are worth mentioning:

Pixel resolution

A higher digital resolution (large amount of pixels) of the sensor corresponds to a higher sampling frequency, which in turn means higher spatial frequencies can be resolved. As FRAME trades spatial resolution for the ability to multiplex exposures, it is beneficial for the raw image to be as well-resolved as possible, as that means that the extracted exposures are also more well-resolved.

Intensity resolution and dynamic range

The intensity resolution is also an important factor - how many levels of brightness an individual pixel can resolve, corresponding to bit-depth in the digital representation. Saturating a sensor means information loss, as the relative intensities of the image are no longer comparable. As all exposures in FRAME are overlayed onto each other, covering the entire sensor, they also have to share the available intensity resolution. For this reason, the same argument applies as for pixel resolution - the higher the intensity resolution of the raw image, the higher the resolution for each exposure.

Dynamic range is a more specific formulation of the same idea, taking image noise into account, and is defined as the ratio between the highest and lowest measurable level of signal. This makes it not only a function of the sensor, but also the imaging conditions, as these also affect the noise level in an image. However, different sensors have varying levels of intrinsic noise, making it worthwhile to mention the difference between intensity resolution and dynamic range. After all, if noise absorbs half of your intensity resolution (extremely unlikely case for illustration purposes), you are not going to get anything out of the lower half. Dynamic range reflects this, while intensity resolution does not.

Read-out speed and frame-rate

FRAME multiplexing is not dependent on read-out speed as it is not a temporally scanning method. It puts no "extra strain" on the frame-rate of the camera, which therefore only needs to be adapted to remove motion blur in the scene. In fact, as will be discussed in Chapter 5, FRAME can be used to encode exposures in time by using pulses of structured light, increasing the possible frame-rate of a sequence capture.

Read-out noise

As the electronic noise introduced during read-out is spatially uncorrelated, it spreads out evenly in the Fourier spectrum of the captured image (equally over all frequency bands). Thus, this type of noise cannot be filtered out by shifting information into higher frequency bands, and therefore affect the dynamic range of the extracted exposures. For this reason, reducing sources of particularly high-frequency noise such as read-out noise is a bonus, but is not a primary factor compared to the other more important parameters mentioned. The effect of white noise on the intensity modulation can be seen in Fig. 4.1. As covered in Section 3.4, this reduction in modulation depth will decrease the signal of the extracted exposure.

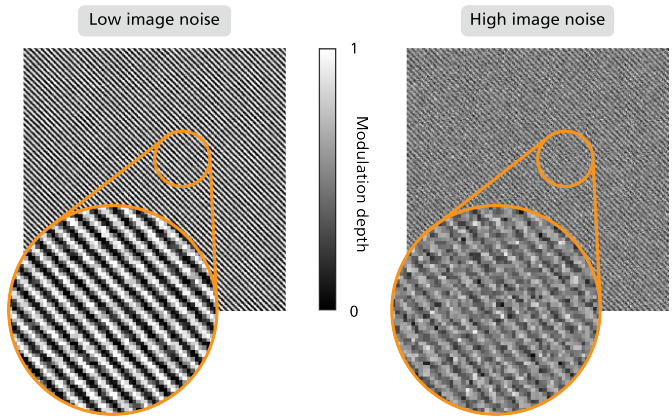


Figure 4.1: Two images featuring different levels of noise. The low noise image maintains good contrast in the modulation pattern, increasing the signal of the carried exposure.

4.2 Optical designs for multispectral encoding

In order to optically encode the light emanating from a scene[1], an optical setup is placed as an intermediate imaging system between the scene and the camera. The purpose of the optical setup is to perform the following actions in sequence (steps 2 and 3 are interchangeable in order):

[1]Known as "passive" FRAME, as discussed in the previous chapter.

1. Split the light into optical channels, creating an image of the scene in each using a relay lens.

2. Manipulate the light in the manner that the encoding should reflect - for example, if the goal is multiplexing of spectral channels, this step would be to spectrally filter the light.

3. Impart a unique spatial frequency onto the image in each channel. This is what enables the separation in the frequency domain in post-processing - a sort of fingerprint which identifies the information from each channel.

4. Recombine (spatially overlap) the channels, creating an image on the sensor of the intermediate image planes in each channel. These images are spatially overlapped on the sensor, but will be separable due to their unique encoding.

Fig. 4.2 shows a schematic of a setup for spectral multiplexing, with the above steps annotated in a single channel to indicate the optical component performing the action. The Ronchi rulings used to impart the intensity modulation are low-frequency transmission gratings - binary square wave masks which extinguish half of the image in a periodically shadowing pattern. This design was used for spectral encoding in Paper V, but has also been used to encode other parameters such as polarization and different focus-depths [79].

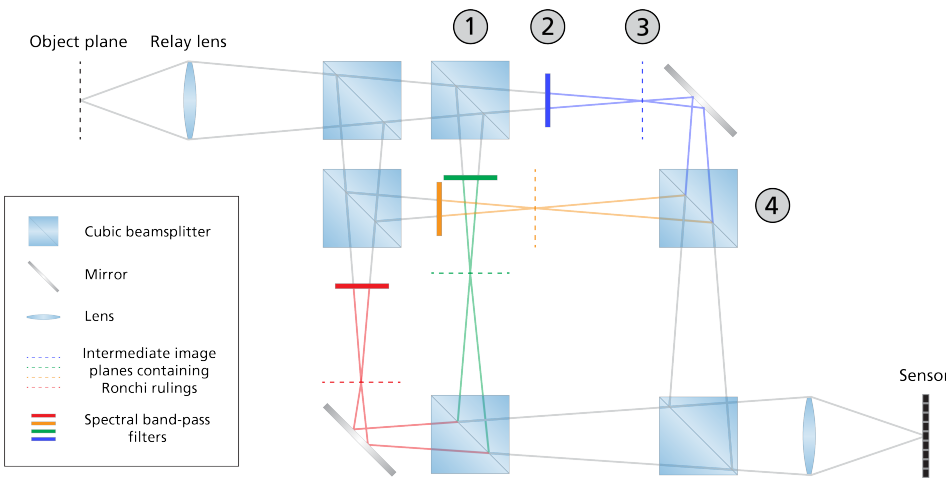


Figure 4.2: Optical setup for spectral multiplexing utilizing cubic beamsplitters. The numbers in gray indicate the component associated with each step in the list of actions needed for spectral encoding. The cubic beamsplitters both split and recombine the optical channels. In between these two actions, the light is spectrally filtered and modulated using a Ronchi ruling.

The setup shown in Fig. 4.2 is versatile with regards to encoding different parameters - replacing the spectral filters with polarizers enables the imaging of different directions of linear polarization, and when performing spectral multiplexing, filters can easily be exchanged to change the spectral properties of a channel without having to realign the system. However, it also discards a large amount of the intensity collected from the scene for two different reasons:

1. The light is split independently of the encoding parameter, meaning a large portion of the light is discarded in the subsequent filtering. Using the example of spectral encoding, the light first entering each channel is broadband, meaning a lot of intensity is discarded by the spectral filtering.

2. The recombination of light is done using cubic 50-50 beamsplitters, which transmit 50% of the incoming intensity in each direction - straight through and reflected. This causes intensity losses for each recombination, as light is transmitted out of the system.

Assuming perfect spectral filters, with 100% transmission in the desired band and 0% outside of it, these two effects each cause 75% of the inbound intensity to be discarded, resulting in only 6.25% of the intensity entering the optical multiplexing setup to be transmitted to the camera[2].

To address this issue, I have been developing a more light-efficient optical setup specifically for the multispectral encoding, which can be seen in Fig. 4.3. This alternative setup increases light-efficiency (the proportion of intensity collected by the optical multiplexing setup that is transmitted to the camera) by using dichroic beamsplitters to split the light into optical channels. These optical components split light based on wavelength, combining step 2 in the list of actions with steps 1 and 4. This avoids both causes of discarded light from the setup in Fig. 4.2, and also allows for the number of channels to be changed freely by adding or removing a set of dichroics. In the cubic beamsplitter setup of Fig. 4.2 this is not as easy, as the number of channels have to be increased or decreased by a factor of two, assuming equal intensity distribution in all channels. However, this more light-efficient setup cannot be modified to multiplex different states of polarization at the same spectral band, and changing spectral bands often involve re-aligning the setup.

[2]This is assuming 4 channels, as in Fig. 4.2.

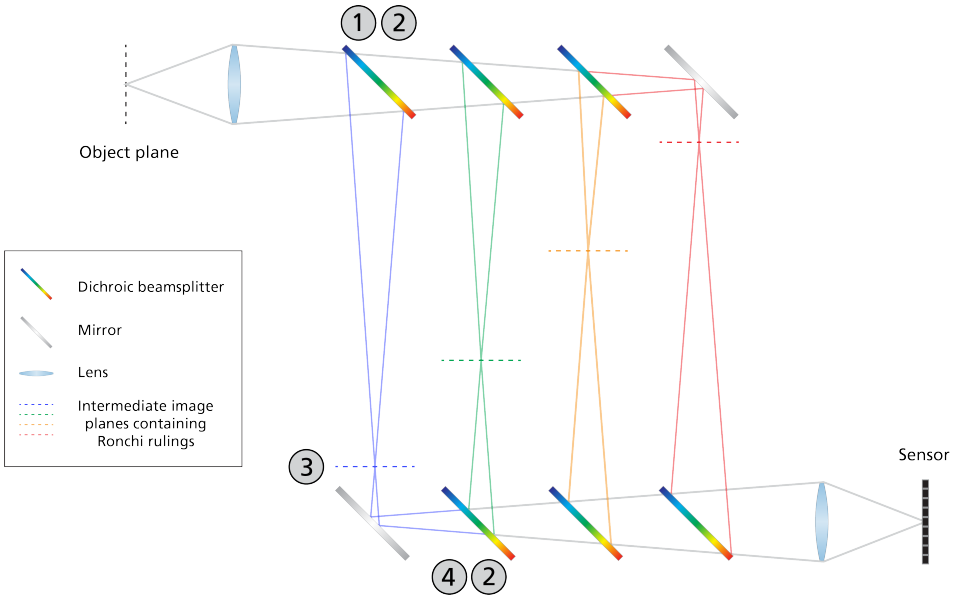


Figure 4.3: Optical setup for spectral multiplexing utilizing dichroic beamsplitters. Just as for Fig. 4.2, the numbers in gray indicate the component associated with each step in the list of actions needed for spectral encoding. The dichroic beamsplitters split the light on the criterion of wavelength, meaning that they combine step 2 with 1 and 4. This significantly decreases the amount of light discarded.

4.3 Alignment of a multi-channel optical setup with multiple intermediate image planes

The optical setup displayed in Fig. 4.3 is the design I have worked with for the majority part of my research, and therefore the one I have come to know the best. This section will discuss the alignment of this optical setup. In order to do so, a definition of "good alignment" must first be established, which will be listed below. The system is aligned when the following criteria are met:

1. Each intermediate image plane is in focus on the camera sensor, creating a focused image of the scene on the sensor (alignment along the optical axis).
2. The intermediate images are spatially overlapped on the sensor (alignment perpendicular to the optical axis).
3. The Ronchi rulings used to impart the intensity modulation in each channel is located in the intermediate image plane.

The complexity of aligning this setup springs from one primary factor - the diagonal symmetry of the setup. In order for the first alignment condition to be fulfilled, the optical path length (OPL) from the camera to each intermediate image must be (approximately[3]) equal. As a consequence of this, the OPL from each intermediate image to the scene must also be equal, in order for all intermediate images to be in focus in the camera. Combining this with the spectral properties of dichroics requires the optical setup to be arranged in the diagonally symmetric way displayed in Fig. 4.3. Each component in such an optical system has many degrees of freedom, and may affect more than just one channel.

4.3.1 Positional alignment

Each optical component in Fig. 4.3 has degrees of freedom in the form of position and angle. The first step in tackling the problem of aligning such a setup is to limit these - fixing some of them in their approximately correct position and/or angle. I solved this problem by first simulating the optical path of each channel in the system, finding the theoretically correct position of each component assuming the optical path reflected at right angles. Using this simulated path, I 3D-printed templates for the layout of the optical components and incorporated precision mounts for the angle alignment of the dichroic beamsplitters and mirrors. In this way, the positions of the optics are fixed in the correct position, as well as roughly the correct angle, before the manual alignment process has even begun. 3D-printing templates for optical setups has been a game-changing tool for my research group, and has enabled very compact multiplexing solutions to be made in a printable format [95].

4.3.2 Alignment along the optical axis

As mentioned, the diagonal symmetry of the setup is a result of the equal OPL-requirements for each channel, which is a requirement for the second alignment condition to be fulfilled. However, due to a couple of effects the intermediate image planes do not fall on a perfect diagonal in practice, meaning the positions of the Ronchi rulings are better kept free along the optical axis and not locked using the 3D-printed template. Firstly, chromatic aberration will affect the distance at which the image in each channel is focused. This can be countered by introducing an iris to the relay lens and increasing the numerical aperture, increasing the depth-of-field at the cost of intensity, so that all channels are in focus despite the difference in optical path length. If the effect is small, it can also be countered by introducing pieces of transparent glass into individual channels, changing the optical path length of each

[3]There is some margin of error here due to the depth-of-field.

individually. This second approach is overall less effective and much more labor intensive, but has the advantage of not affecting the intensity levels of all channels collectively. Secondly, the light passing through channels located further away from the scene will be transmitted through more glass (transmission through other dichroics) before reaching the intermediate image plane. This equates to the physical distance along the optical path from the intermediate image plane to the first lens being shorter, resulting in a "stretch" of the diagonal out towards the dichroics. The magnitude of this effect depends on the number of channels and thickness of the dichroics. However, this second effect is at least partially compensated for in the next imaging stage per the very symmetry of the setup, as channels further from the scene now have less glass to travel through than channels closer to the scene. Given that the magnitude of these two effects are hard to predict, the position of the Ronchi rulings are best kept variable, and not locked by the 3D-printed template.

4.3.3 Alignment perpendicular to the optical axis

The second type of alignment is lateral alignment (perpendicular to the optical axis), which is required for both conditions 1 and 2 to be satisfied. In the case of the first condition, the optical path length of each channel is (of course) affected by the angles at which the rays are reflected, which in turn affects the intermediate image planes position and rotation. Therefore, the ideal alignment is for all channels to be reflected at right angles. When this is achieved for each channel, the second alignment condition, the alignment of the object plane between channels, will also be fulfilled.

To achieve such a right angle alignment, assuming only an angle dependency after having locked the position of the reflecting elements, I have been using a far-/near-field approach. This approach is similar to beam walking when aligning a laser system, adjusting the angles of the two dichroics of each channel, only this involves using two reference points and switching between imaging each of them (using a camera lens as the second lens in Fig. 4.3). The two reference points used are the object plane and an iris adjacent to the relay lens, and when these converge on the sensor, that optical channel is aligned. Fig. 4.4 shows these two reference planes, where spatial alignment is good in the object plane, but not in the iris plane, showing that the three channels have differing optical paths.

In this alignment procedure, one of the channels is established as the reference, to which all other channels are aligned. This also means that any errors in alignment of the reference channel will be inherited by the other channels, meaning that the reference channel must first be aligned. Here, the 3D-printed templates play a crucial role, as they lock the relative positions of optical components, ensuring that both

reference planes align with the first mirror[4] of the reference channel. By following the optical path of the reference channel through the system, the reference channel can be aligned at right angles.

As mentioned, the iterative far-/near-field alignment is done per channel by adjusting the angle of the two dichroics associated with each channel. However, the diagonal symmetry of the setup introduces complexity in the form of a channel-dependency to this process, as the distance from the sensor to the first and second reflecting element (dichroic beamsplitter/mirror, hereafter both referred to as "mirrors") of each channel varies. Both mirrors affect the position in both reference planes, but to varying degrees, which means convergence is a matter of associating the right optical element to the right reference plane.

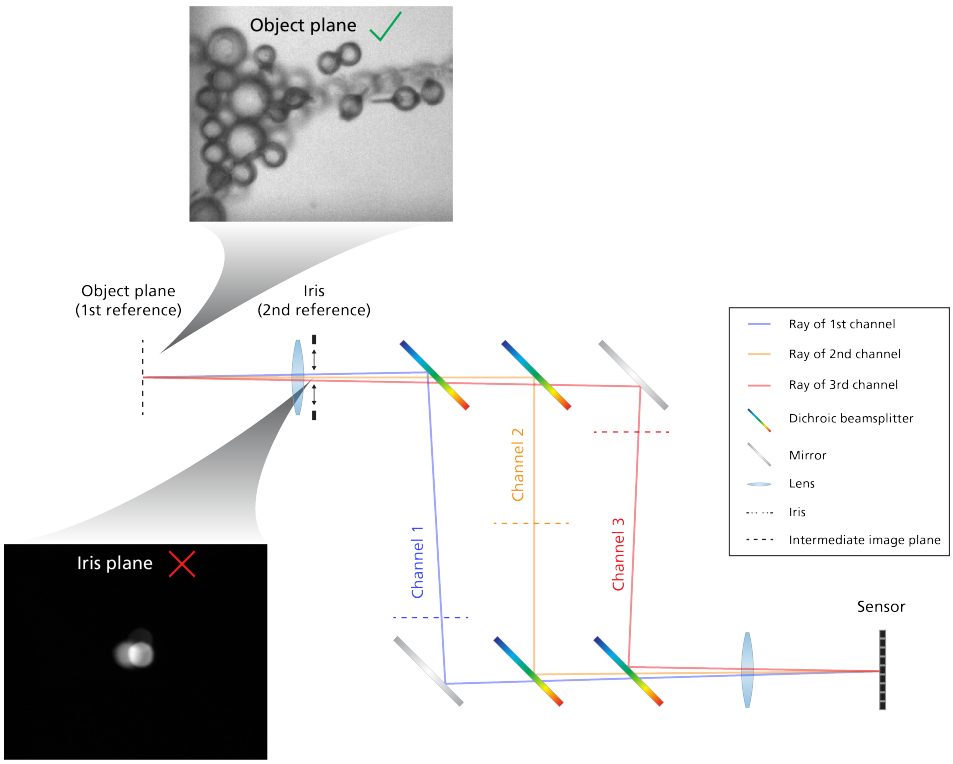


Figure 4.4: Example showing how the optical setup can be misaligned despite image overlap in the object plane. All three channels spatially overlap in the object plane. However, the non-overlapping images in the iris plane show that the optical channels have different beam paths and thus different optical path lengths, as seen in the overview of the optical setup. This can lead to a difference in focus in the object plane.

[4]First and second element are counted in the direction of propagation, meaning the first reflecting element is located closest to the scene, while the second is located closest to the sensor.

To illustrate how this affects the alignment process in practice, Fig. 4.5 provides a step-by-step overview of the alignment of the two extreme channels in a simplified dichroic optical setup (3 channels). In this example, the middle channel is used as reference. Beginning with Channel 1, the intermediate image plane is located close to the second mirror, meaning the resulting image translation from an angle-shift is low compared to the first mirror which is located far from the intermediate image plane. At the same time, the first mirror is closer to the iris, while the second mirror is far from it, resulting in the opposite angle-dependency. The "sum of the cardamom"[5] is that controlling the first mirror when aligning the object plane and the second when aligning the iris will lead to convergence. The blue inset illustrates this, showing the two reference planes for the first channel. At the starting point, both planes are misaligned, but after aligning the image plane using the first mirror, alignment is achieved in the image plane. This action also affects the alignment in the iris plane, but further alignment in this plane is still needed. This is done by adjusting the second mirror, which affect the alignment in the object plane, but with an important caveat - the translation is smaller than before the last step. By iterating between aligning the object plane and the iris in this fashion, simultaneous alignment in both planes is eventually achieved.

Following the same logic for Channel 3, we can deduce that the plane-to-mirror connection is now the opposite. Due to the diagonal symmetry of the setup, the intermediate image plane is now farther from mirror 2 than mirror 1, meaning the resulting translation in the object plane of moving mirror 1 is very small. Both mirrors have a big impact on the translation in the iris plane, but the big difference between angular translation in the image plane dominates. This results in the second mirror now being used to align the image plane, while the first mirror aligns the iris. In this way, convergence and simultaneous alignment in both reference planes is achieved in this channel as well.

[5] This is a Swedish idiom, roughly translating to "the conclusion".

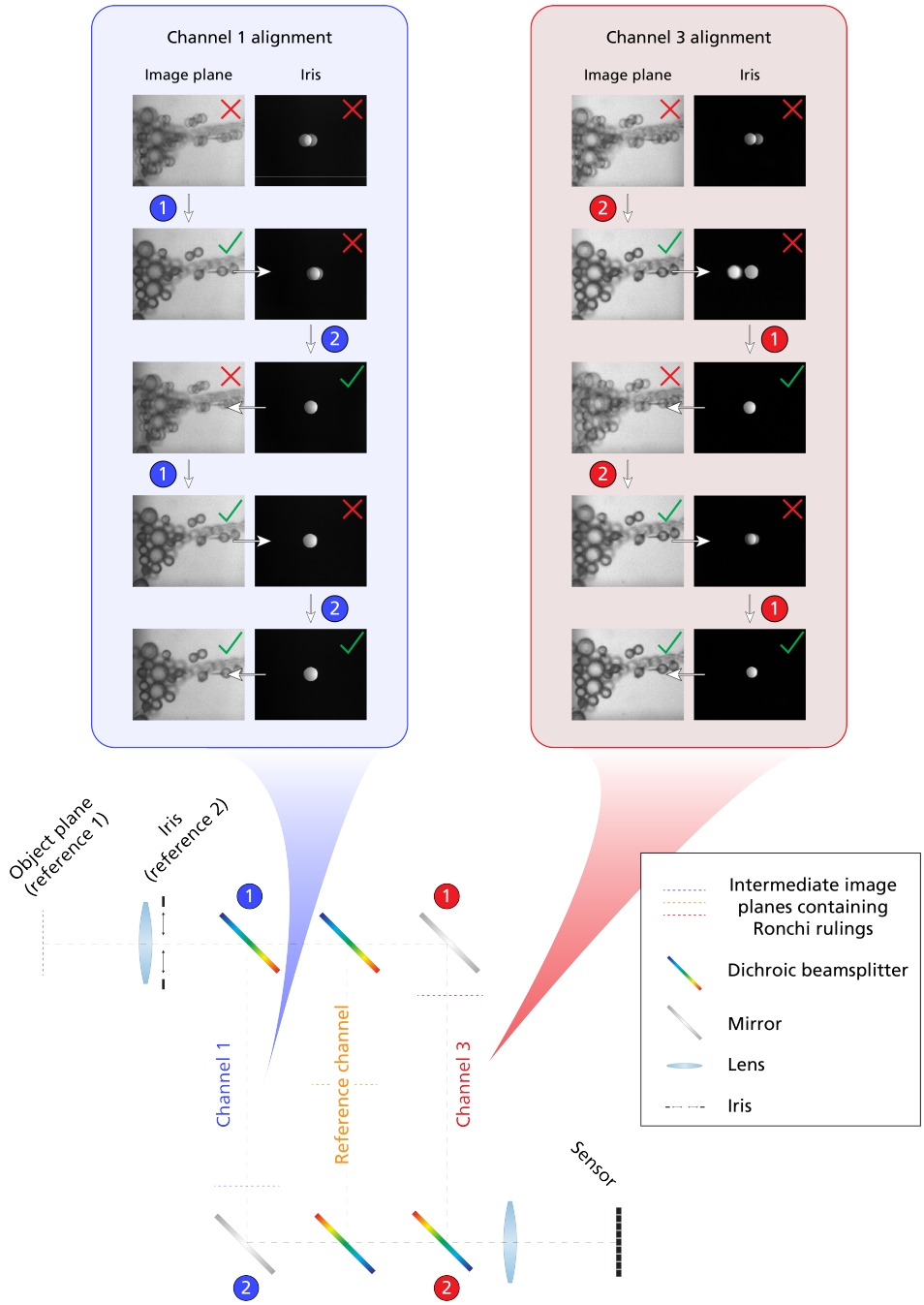


Figure 4.5: Alignment procedure of the dichroic setup. For the first channel (1), with the intermediate image plane located close to the second mirror, the first mirror (dichroic) is used to align in the object plane, and the second at the iris. By iterating between the two, convergence is achieved where both reference planes are simultaneously aligned. For the last channel (3), the mirror-to-plane association is reversed due to the diagonal symmetry of the setup. The middle channel (2) is used as a reference while aligning.

4.3.4 Intensity modulation alignment and flat-field correction

The last alignment condition involves placing a Ronchi ruling in the intermediate image plane of each channel, in order to impart a spatial frequency modulation. In my experience, this is easiest done by creating an image of the intermediate image plane on the sensor and manually move the Ronchi ruling back and forth along the optical axis, maximizing the modulation depth. As shown in Section 3.4, the signal strength of a multiplexed exposure scales linearly with its modulation depth. Different types of misalignment and their effect on the extracted signal can be seen in Fig. 4.6. Four cases are shown, demonstrating the effect the modulation depth has on the extracted signal.

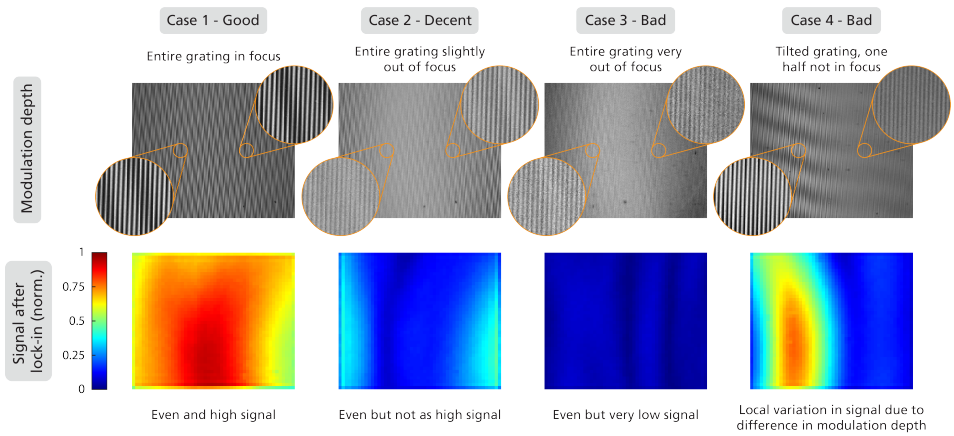


Figure 4.6: Four cases of Ronchi ruling alignment, demonstrating the effect of modulation depth on extracted signal. Case 1 shows a well-aligned grating, leading to uniform and high signal after lock-in. Case 2 and 3 show a grating which is placed out of focus, decreasing the overall modulation depth and therefore the signal after lock-in. Case 4 shows a tilted grating, where half of the grating is in focus, and the other half not, leading to significant local differences in the extracted signal over the field of view.

As modulation depth varies both between exposures (channel-wise, analogous to comparing cases in Fig. 4.6) and locally within an individual exposure (within a single channel, most clearly displayed in Case 4), this variation needs to be calibrated for in order for intensities to be directly comparable. This is done using a flat-field correction - imaging a surface of constant brightness, first through each individual channel with no modulation (one image per channel), and second with a multiplexed exposure through all channels (one image per channel after extracting each multiplexed exposure). Taking the channel-wise ratio of the pairs of images produces a two-dimensional multiplicative calibration-matrix for each channel, compensating for the effects of the modulation depth both locally and between channels.

An experimental aspect to be aware of when designing the setup is to ensure the imag-

ing of the Ronchi ruling is as free from aberration effects as possible. Effects such as field curvature and radial distortion will change the modulation in different ways, affecting the ability to extract an accurate exposure of the scene. Field curvature will cause parts of the modulation pattern to be out of focus, as the focus plane will become warped. Radial distortion causes local changes in the modulation frequency in both direction and magnitude. Thus, the carrier frequency varies within the exposure, spreading the information out in the Fourier spectrum, leading to an inaccurate reconstruction of the scene. For these reasons, using for example a thick lens or a large optical magnification in this imaging stage is best avoided.

The square wave nature of Ronchi rulings

Ronchi rulings are square wave binary patterns which periodically shadow the image in the intermediate image plane, imparting the spatial carrier frequency needed to separate the exposure in post. Drawing from the discussions in the Section 3.2.2, this also means additional image copies are created in the Fourier spectrum of the image, as seen in Fig. 3.4. However, Fig. 3.4 is a perfectly sampled square wave (simulation). In practice, optical resolution, imperfect focus and other sampling effects can actually help to soften the edges of the square wave, leading to this effect not being nearly as pronounced. To avoid the effect entirely, sinusoidal patterns can be used to mask the image instead, but these are harder to manufacture and therefore more expensive. As long as the folded frequencies of the image copies can be avoided, Ronchi rulings work well to modulate exposures. Using square wave modulation can even result in stronger signal of the multiplexed exposure due to the effect described in Section 3.4.

Choosing magnitude and direction of carrier frequencies - Placement in the frequency domain

The separation between carrier frequencies in the frequency domain indirectly affects the final resolution of the extracted exposures, as higher image frequencies can be included without the filtering overlapping for different exposures. Fig. 4.7 illustrates this relationship. Extracting images of as high quality as possible therefore comes down to carrier frequency placement in the frequency domain, which is determined by the rotational directions and magnitudes of the Ronchi rulings.

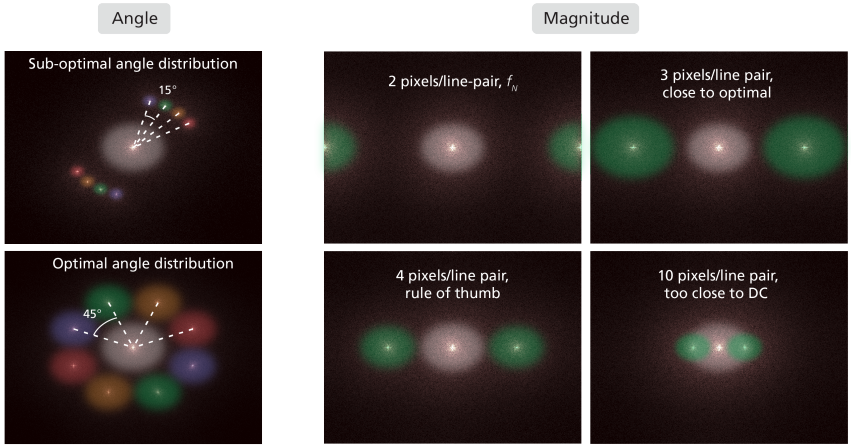


Figure 4.7: Carrier frequency placement in the Fourier spectrum of a natural image. **Left:** The optimal distribution of angles in the case of few exposures ($\lesssim$10) is equally over a 180° span. This maximizes the filter radius possible. **Right:** Four cases of frequency magnitude are shown. At the Nyquist frequency (top left), the image frequencies surrounding the carrier frequency are affected by frequency folding. On the other hand, if carrier frequencies are placed too close to the DC-component, cross-talk will occur between the two (bottom right). The optimal placement is somewhere in between (in this case around 3 pixels/line-pair, top right), where the radius of the filter can be maximized without being affected by cross-talk or frequency folding (this is assuming the optical resolution is not limiting the system). As a rule of thumb, 4 pixels/line-pair is a good starting point (bottom left).

Maximizing the separation between carrier frequencies is a product of both direction and magnitude of each frequency. In the case of multispectral FRAME, which uses Ronchi rulings that are manufactured in set densities such as 5, 10, or 20 line-pairs/mm, most of the freedom when it comes to the "placement" in the frequency domain comes in the form of frequency direction. This works well, as multispectral FRAME does not multiplex enough exposures to make the use of different frequency magnitudes necessary [96]. The placement of the carrier frequencies in relation to each other is therefore fairly straight-forward from an experimental perspective - using Ronchi rulings of the same magnitude, and evenly distributing their rotations over a 180° span, as illustrated in Fig. 4.7. In order to control this distribution of angles in the Fourier spectrum of the image, I have designed and 3D-printed custom holders for the rulings, with unique rotation angles for each holder. Examples of these can be seen in Fig. 4.8, implemented in the optical setup. These ensure even separation of the carrier frequencies in the Fourier spectrum by means of direction.

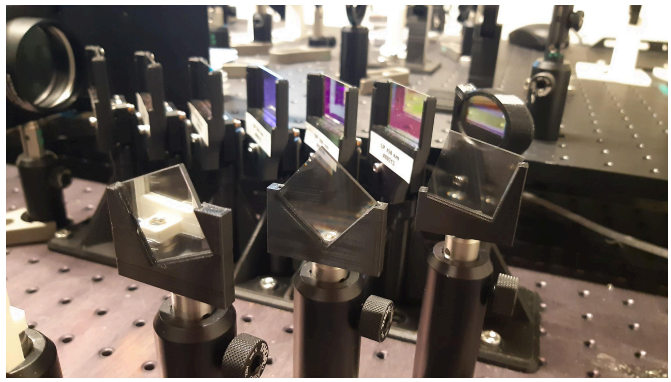


Figure 4.8: 3D-printed holders allow for accurate control of the rotation angles of the imparted carrier frequencies.

As mentioned, the magnitude of Ronchi rulings come in set intervals, meaning fine-tuning of this parameter is not possible. Therefore, in experimental practice it is a matter of choosing the best option available. This decision is weighed by balancing frequency band isolation from the image frequencies surrounding the DC-component (non-modulated image information) with distance from the Nyquist frequency or, in the case of a diffraction-limited image, the optical resolution-limit. The "optimal" placement of the carrier frequency in terms of Fourier spectrum coordinates is therefore right between the (somewhat arbitrary) frequency band edge of the DC-image[6] and the highest resolvable frequency (limited either by the sampling or the optical resolution). As a rule of thumb in the lab, I aim at a ruling density which is sampled at a frequency close to 4 pixels/line-pair. This isolates the multiplexed exposure from most of the frequency information around the DC-component, while still being resolvable in the majority of cases. Fig. 4.7 shows these two cases of carrier frequency placement in the Fourier spectrum of an image, as well as a pour placement too close to the DC-component and placement at the Nyquist frequency.

[6]Where this edge lies is a matter of the acceptable level of cross-talk between the sidebands and the mainband. For an excellent discussion on this balance between specificity (the isolation of signals) and spatial resolution of the extracted exposures, I recommend the thesis of Vassily Kornienko [97].

Chapter 5

Applications and results

"It is better to be vaguely right than exactly wrong."

- Carveth Read

5.1 Temperature imaging

Temperature imaging, where temperatures are mapped to a two-dimensional spatial distribution, has applications within military, medicine, agriculture and industry [98]. Depending on the temperature range of interest, different parts of the spectrum are targeted. For almost all temperatures with a practical application, the bulk of the thermal radiation is emitted in the infra-red (IR) region of the electromagnetic spectrum, but for very high temperatures it becomes increasingly beneficial to instead move towards imaging in the visible region. The imaging discussed in this section focuses on such high temperature applications, where the lower end of measurable temperatures are $\sim$2000 K, making imaging in the visible region a viable solution.

5.1.1 Pyrometry

According to Planck's radiation law, the spectral radiance $B_\lambda(\lambda, T)$ of a black-body can be expressed as

$$B_\lambda(\lambda, T) = \epsilon \cdot \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/(\lambda k_B T)} - 1} \quad , \tag{5.1}$$

where ϵ is the emissivity, λ is the wavelength of light, T is the temperature of the blackbody, h is Planck's constant, c is the speed of causality, and k_B is the Boltzmann constant. A consequence of this expression is that the shape of the spectrum, disregarding absolute intensity values, is uniquely determined by the temperature T of the object, with the peak wavelength of the spectrum being shifted according to Wien's displacement law,

$$\lambda_{peak} = \frac{b}{T} \quad , \tag{5.2}$$

where $b \approx 2.9 \cdot 10^{-3}$ [m · K] is Wien's displacement constant. As such, the temperature of a black-body radiator can be determined by fitting measured intensities to Eq. 5.1. This non-intrusive measurement technique is called pyrometry. A perfect black-body radiator has an emissivity of $\epsilon = 1$, while for an imperfect black-body radiator, called a gray-body, the emissivity takes on a constant value $\epsilon = [0, 1)$, affecting only the absolute radiance of the object. Emissivity may also vary as a function of both wavelength and temperature, which affects the spectrum shape, making it an important factor in any calibration procedure.

As pyrometry uses intensities at different wavelengths, spectral imaging can be used to create a two-dimensional spatial mapping of the temperature of an object. In theory, measuring the absolute radiance at a single, well-defined wavelength (provided enough signal) is enough for a unique solution for the temperature. However, this approach requires rigorous calibration, including emissivity and transmissions of optics, as it relies on absolute values. This makes it less suitable for imaging applications, which involve many components, each with their own spectral transmissions and responses, resulting in large uncertainties. Two-color pyrometry instead uses the proportional radiance at any two wavelengths (provided enough signal), and allows temperature estimation based on relative values. This is commonly done by defining a ratio, hence the alternative name "ratio pyrometry". Two-color pyrometry is better suited for imaging applications as it does not use absolute values, but requires a combination of good signal and that the ratio as a function of temperature is steep enough over the entire temperature range of interest. An extension of two-color pyrometry is multi-wavelength pyrometry (sampling more than two wavelengths), increasing the number of spectral sample points. In theory, this means that temperature can be estimated over a larger range, and also allows for an estimation of the emissivity function if its behavior is known.

5.1.2 Temperature imaging using multispectral FRAME

Papers I and II utilize a multi-wavelength analysis approach for temperature imaging that I have developed for the contiguous and wide spectral bands imaged using the type of dichroic multispectral FRAME setup described in Chapter 4. Each pixel in the resulting multispectral image is a spectral signature, consisting of the integrated intensity transmitted through the system for each spectral band (hereafter referred to as a spectral response vector). If a black-body radiator is imaged, these relative intensities represent distorted black-body spectra, modified by the intermediate optical system and spectral response of the camera.

An explanation of the calibration process used to compensate for these distortions can be seen in Fig. 5.1. By simulating the spectral response of each channel through the imaging system (including the spectral response of the camera) for a range of temperatures and integrating the intensity, the "integrated spectral response vectors" are calculated. The elements of these vectors can be seen plotted as a function of temperature, creating spectral response curves for the three channels. Using a series of calibration images taken of a blackbody emulator through the imaging system, the simulated curves are then corrected for any effects not accounted for in the simulation.

In order to map the relative intensities of the multispectral image to a temperature, a pixel-wise comparison is done. To quantify the similarities between the spectral response vector of a pixel $\mathbf{s}$ and each temperature reference vector $\mathbf{r_T}$, the cosine similarity

$$S_C(\mathbf{s}, \mathbf{r_T}) = \frac{\mathbf{s} \cdot \mathbf{r_T}}{||\mathbf{s}||\,||\mathbf{r_T}||} \quad , \tag{5.3}$$

is used. This metric compares the direction of the two vectors in N-dimensional space (N being the number of channels in the multispectral image), and ranges between 0 and 1, where 1 is a perfect match in direction, and 0 is an orthogonal vector. This is analogous to comparing the relative contributions of each element, disregarding the length of the spectral vector (absolute values of the intensities).

The relatively wide spectral bands used in this approach yields a better signal-to-noise ratio compared to narrow-band pyrometry. However, this comes at the cost of spectral specificity, requiring more rigorous calibration and analysis. It also puts more emphasis on the choice of cut-off wavelengths between channels (spectral band division) as the wider spectral bands can make the technique less sensitive to temperature (as in, spectral response vectors differ less between temperatures).

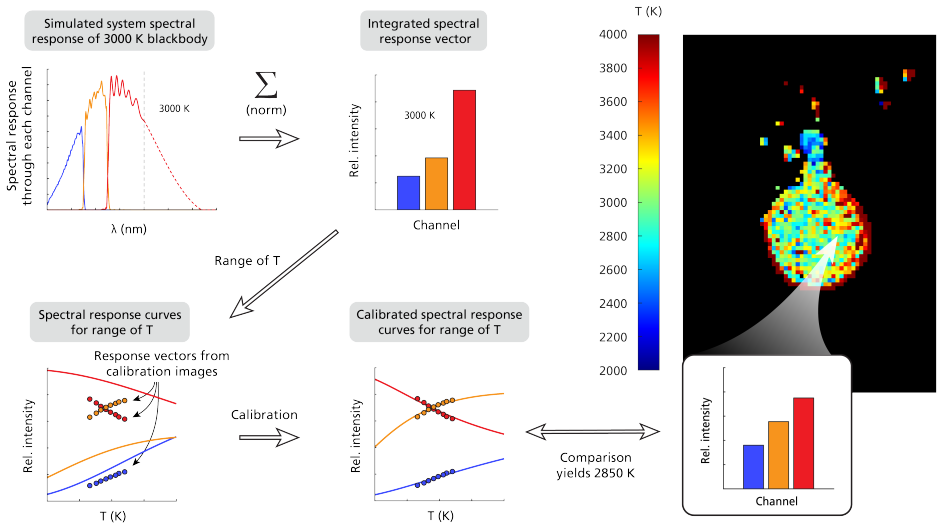


Figure 5.1: Schematic describing the calibration required for pyrometry using the spectral multiplexing setup. Starting at the top-left, the spectral response of the system is simulated for a tungsten blackbody at 3000 K, and the response in each spectral channel is integrated to create an integrated spectral response vector (normalized to a length of 1). Repeating this for a range of temperatures and plotting the values of the response vectors as a function of temperature yields simulated spectral response curves for blackbodies of different temperatures. To compensate for any factors not included in the simulation, the spectral response curves are corrected by being fitted to 8 measured responses using an affine transformation (while also compensating for the emissivity of tungsten). This gives the final spectral response curves used to calculate the temperature image. The temperature of a pixel in an image is then calculated by comparing the relative angle between the integrated spectral vector of the pixel and each reference vector. Figure adapted from Paper II.

5.1.3 Result: Temperature imaging of propane flame using multispectral FRAME (Paper I)

Paper I features the first proof-of-concept for this type of analysis with multispectral FRAME, demonstrating that accurate temperature imaging is possible using multispectral FRAME, given the right circumstances. The setup consists of four spectral channels in the visible range of the electromagnetic spectrum, with cut-off wavelengths at approximately 490 nm, 550 nm and 650 nm. Fig. 5.2 shows a figure taken from said article demonstrating the pyrometry approach, where the temperature of the soot in a soot-rich diffusion flame is estimated. Inset (b) in the figure also shows a plot of estimated temperatures of the same black-body emulator used to for calibration, excluding the temperature at which the setup is calibrated.

The main take-away from this result is, in my opinion, the right-hand side plot. It shows that given an appropriate calibration source, the experimental methodology and analysis developed for temperature imaging using multispectral FRAME works well, with errors as small as 50 K more than 500 K from the calibration point. As for the choice of imaging target, it was not optimal for a first try - imaging the temperature

of a sooty diffusion flame produced using a Bunsen burner is a challenge, albeit a relevant object of study. First of all, the transparency of the flame means line-of-sight affects the measured spectral intensities, as intensities are integrated along the optical axis. I attempted to compensate for this by performing an Abel transform [99], but due to the asymmetry of the diffusion flame this did not yield any convincing results. Secondly, other sources of emission, such as chemiluminescence, are present in the flame. In order to account for this, a minimum threshold value was set for the cosine similarity to exclude points in the flame not dominated by thermal radiation, but a better alternative to this would have been to try to filter out the effected bands during image capture. That being said, the average temperature over the entire flame is 2170 K, which is good considering the sources of error. For reference, the approximate temperature of a stoichiometric propane-air flame using a Bunsen burner is 2220 K [100], while the temperature of a flame of the type imaged in Paper I is likely a few hundred K lower.

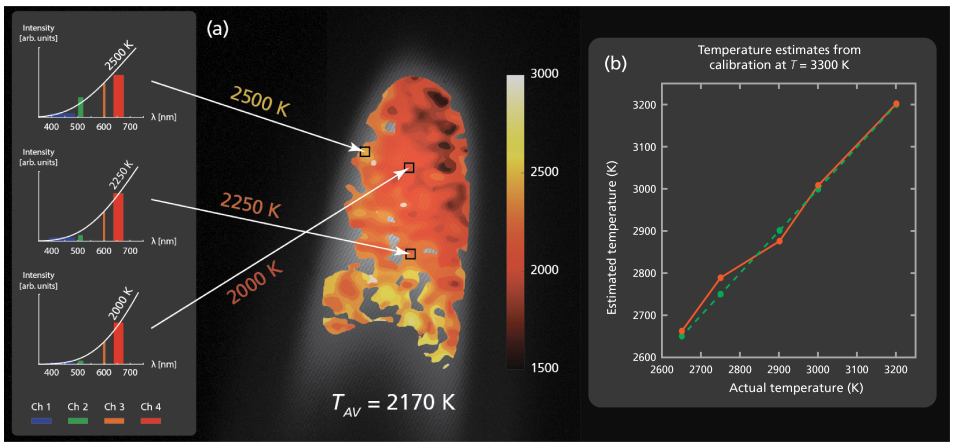


Figure 5.2: Temperature imaging of soot-rich diffusion propane flame, using spectral multiplexing. **(a)** Each pixel of the extracted multispectral image is compared to a reference library of spectral response curves. By calculating the smallest angle between the measured and reference spectral vectors, the closest match for a temperature is established. In effect, this corresponds to fitting the value of the integrated spectral response in a pixel to simulated blackbody curves (gray inset, left-hand side). The holes in the temperature mapping are due to insufficient confidence. **(b)** The method was validated by measuring the temperature of a blackbody emulator lamp capable of a range of temperatures. The system was calibrated at 3300 K and subsequently used to measure at lower temperatures. Despite the large difference (550 K) between calibration measurement temperature, the method gives accurate results (at most $\sim$ 50 K difference). Figure adapted from Paper I.

5.1.4 Result: Temperature imaging of Al-particle combustion using multispectral FRAME (Paper II)

For my second attempt at temperature imaging, the same type of multispectral setup is used, but this time using only three channels (cut-off wavelengths at 550 nm and 650 nm). The motivation behind this paper is to monitor the temperature distribution of burning aluminium (Al) particles as a function of time. As aluminium is a promising fossil-free fuel alternative [101], understanding its combustion process is essential to fully utilize its potential in such applications. The multispectral setup was used in combination with a high capture-rate camera capable of capturing at 50 kHz with a digital image resolution of 1280x896. A field-of-view of approximately 6x4 mm^2 was imaged in the object plane, magnified in two stages (using both the relay lens and the second lens in Fig. 4.2).

Fig. 5.3, taken from Paper II, shows the first 1-10 ms of the combustion process (from time of ignition) of an Al-droplet with a diameter of roughly 550 µm. The droplet is created by placing an Al-wire into a flame, which subsequently burns off (around 3 ms), temporarily suspending the droplet in the flow of the flame until it is eventually transported out of view (typically after 20+ ms). A droplet of this size typically reaches a somewhat stable temperature after around 19 ms [102], meaning the temperature is steadily increasing for the duration of Fig. 5.3. A histogram of the temperatures visible are also shown above each snapshot, showing how the temperature distribution gradually shifts towards higher temperatures with time. The temperatures were calculated using the same analysis pipeline as for Paper I, with the difference that a set of 8 calibration images spanning a temperature range between 2660 K and 3200 K were used to correct the simulated response curves, instead of the previous 1.

Analyzing the resulting temperature mapping, a few things stand out. Firstly, there seem to be a thin region of very high temperature around the droplet, even when surrounding temperatures are very low. This may not accurately reflect temperature, and is most likely caused by de-focusing in one of the spectral channels due to chromatic effects, which throws off the spectral signature in these regions. Similarly to Paper I, a minimum threshold value for the cosine similarity is used to exclude spectral signatures not resembling thermal radiation, but some non-physical values can still persist if they maintain spectral resemblance to thermal radiation. Secondly, as the main droplet burns it pushes more material out into the surrounding reaction zone, creating two zones with very different temperatures. The pushed out nanometer-sized particles have a much higher temperature, and are partially obscuring the droplet. Their emissivity also varies greatly with both temperature and wavelength [102], introducing uncertainties in temperature values as the image captures the combined spectral signature of both different temperatures and emissivities along its line-of-sight. This

is also likely the reason for the very high temperatures in the later part of the sequence, as the nano-particles are accumulating in the reaction zone.

The first thing that needs establishing before looking at the validity of the results is that very few comparable studies have been done. The setup used to generate the particles is, as far as I have understood it, one-of-a-kind when it comes to enabling repeatable experimental scenarios, and the many degrees-of-freedom in such an experiment makes comparisons with other similar studies difficult[1]. For this reason, we have to use either known temperature references or previous studies done under the same conditions in order to evaluate the results. With this in mind, as the outer most layer of the Al-wire is a coating of Al_2O_3, its melting point of 2327 K [103] can be used a reference point for the boundary region between Al starting to undergo combustion (1 ms in Fig. 5.3, light blue) and the rest of the wire (dark blue). This indicates that the temperatures measured in this temperature range are accurate, with calculated temperatures in the boundary region ranging between 2300-2500 K.

Secondly, the whole sequence can be compared to previous studies done using the same Al-combustion setup in combination with a filter array multispectral camera [102], showing good agreement in general, but in particular regarding the temperature of the main particle. To summarize, the results of Fig. 5.3 seem to agree well with the few comparable references we have, indicating that the method has managed to overcome the difficult conditions of the measurement (550 µm particle, very high temperatures, and µs dynamics). These results demonstrate the potential of multispectral FRAME to be used in conjunction with highly specialized sensors, such as intensified and high capture-rate cameras, for quantitative measurements that utilize intensity. Moreover, the independency of the spectral setup from the sensor enables spectral flexibility, allowing the user to tailor the spectral bands monitored to the specific temperature range of interest.

[1] I do not have a literature reference for this, but completely trust the word of my expert colleagues working within this field.

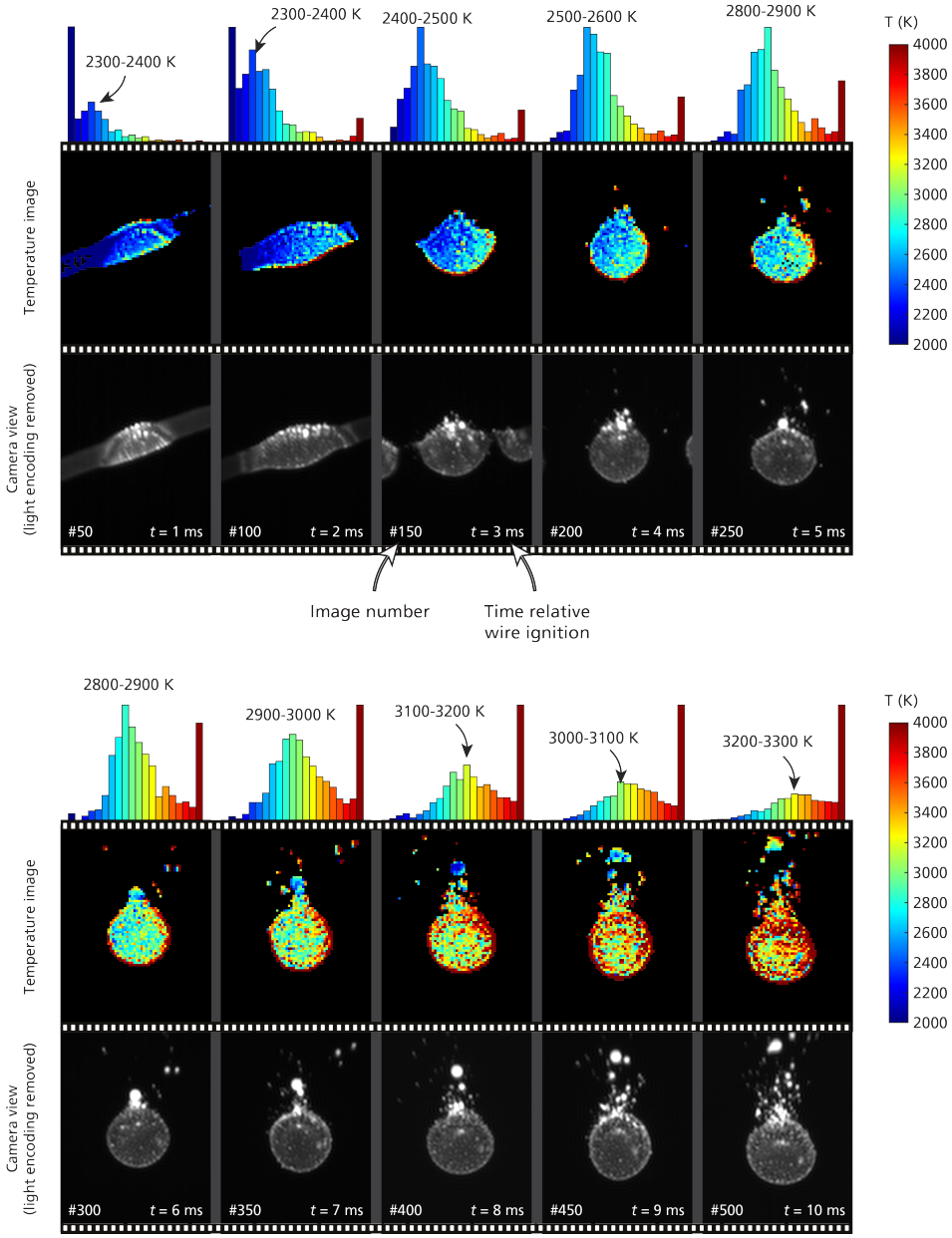


Figure 5.3: Sequence of temperature images, depicting the combustion process of a $\sim$ 550 µm aluminium droplet, captured at 50 kHz. Over the entire sequence, the overall temperature increases from between 2000-3000 K at 1 ms to 3000-4000 K at 10 ms. As the combustion progresses, the radius of the reaction zone surrounding the droplet increases, while the droplet diameter decreases due to mass loss. The different temperatures of the center particles compared to the surrounding particles in the reaction zone can be seen from 6 ms. Figure taken from Paper II.

5.2 Snapshot velocimetry and accelerometry imaging

Velocimetry imaging is a quantitative imaging method where by measuring the physical displacement of features between two images and knowing the time interval between them, velocities in the image plane are extracted. Extending the same logic, three images of known time intervals can be used to calculate acceleration, hereafter referred to as accelerometry imaging, or just accelerometry. The velocities and accelerations measured are then the average over the duration of the image doublet/triplet used to calculate them. Velocimetry imaging is a common diagnostics method in fluid dynamics, most often implemented for particles in the form of Particle Image Velocimetry (PIV) [104].

5.2.1 Temporal FRAME

Just as FRAME can be used to multiplex spectral bands, it can also be used as a videography technique by multiplexing exposures in time, increasing the temporal resolution of a camera. This is done using pulses of spatially modulated illumination (usually coherent) which are all captured within the time-gate of a single camera image exposure. Sequentially, each uniquely modulated pulse interacts with the scene and exposes the sensor, adding encoded temporal information to the resulting image. Through the use of the lock-in procedure described in Chapter 3, the multiplexed exposures can then be isolated in post, identically to how the spectral bands are isolated in multispectral FRAME.

Pulsed illumination circumvents the electronic temporal limitations of the camera, allowing for a higher capture-rate burst to be stored within a single camera image independently of camera capture-rate. The multiplexed exposure capture rate and temporal resolution have been discussed at length [95; 97], but in essence depends on the temporal separation and temporal pulse length of the light pulses. Temporally encoded FRAME has been used extensively and for many applications, among other things to image the propagation of a light pulse with femtosecond resolution [105].

5.2.2 Imaging sprays using temporal FRAME to increase capture rate

As FRAME multiplexing can provide temporally resolved exposures within the gate of a single camera image, three such multiplexed exposures enables snapshot velocimetry and accelerometry. In Paper III, we combined three high repetition-rate pulsed lasers with a high capture-rate camera, tracking velocities and accelerations of features in the spray at 21 kHz camera capture-rate.

Optical setup

The optical setup used for temporal multiplexing of sprays can be seen in Fig. 5.4 (a). It consists of three high-repetition pulsed lasers whose beam paths are overlapped spatially using a dichroic mirror and a polarizing beamsplitter (PBS). The lasers and camera are triggered so that all three pulses are incident on the camera during its gate time, with the pulses arranged in a pulse train. The beam profile of each pulse is uniquely modulated using Ronchi rulings and focusing in combination with a spatial filter, which encodes the temporal image information of the pulse interacting with the spray in the object plane. The object plane is then imaged using a 4x microscopic lens.

An example of the raw multiplexed images resulting from this optical setup are shown in the top-left of Fig. 5.4 (b). As each spatially modulated pulse interacts with the spray, it is partially shadowed, creating the profiles seen in the image. Zooming in reveals the modulation patterns, which vary throughout the image depending on where the spray was at the time of interacting with the light pulse. Located underneath in the same subfigure is the Fourier spectrum of the raw image, showing how the modulation pattern of the pulses shifts the frequency information of the exposures into higher frequency bands. Following the lock-in procedure of FRAME described in Chapter 3, the individual exposures are extracted, showing the propagation of the initial spray injection at three points in time, captured in the same camera image.

An added benefit of this type of imaging is the freedom to space the pulses differently in the gate of the camera. This was demonstrated in the paper and can be used to examine different time-scales within a single exposure. Fig. 5.5 shows examples of this in action.

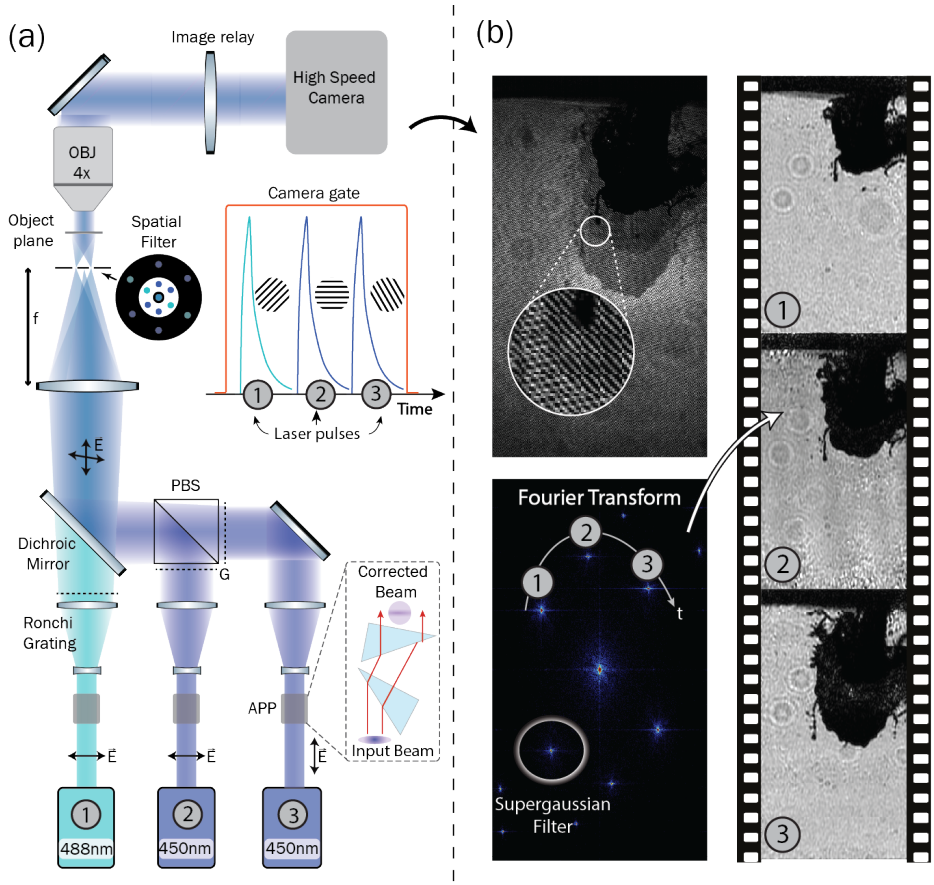


Figure 5.4: The optical setup used to encode temporal exposures in Paper III, and the process with which the FRAME triplets are extracted. **(a)** The optical setup used to temporally multiplex exposures, using three pulsed lasers to create a pulse train that illuminates the spray at three distinct times. All three light pulses are captured in the same camera image. **(b)** Thanks to the unique structure of each light pulse, each exposure is isolated in the frequency domain. Through lock-in and filtering, each exposure can be extracted in post. Figure adapted from Paper III.

Figure 5.5: Three high-speed sequences of sprays, using different pulse train spacings. The manipulation of pulse spacing allows bypassing the capture-rate of the camera for a triplet burst, now instead limited by the pulse width and temporal spacing between pulses. The temporal spacing can also be varied within a single pulse train, allowing for the probing of different timescales in a single exposure. Figure adapted from Paper III.

5.2.3 Extracting velocities and accelerations

I started my doctoral studies in September 2020, and had only been at it for a few months before new restrictions prevented me from doing work in the optics lab. However, one of the measurements that I had been a part of before lockdown was the images seen in Fig. 5.4. I therefore spent some time during the lockdown of winter 2020-2021 writing my own in-house algorithm for extracting velocities and accelerations out of the extracted FRAME-triplets of each raw image. The algorithm was built to work with the type of shadow imaging produced using this optical setup, and consists of the following steps:

1. **Segmentation of the spray.** This step involves different types of thresholding and morphological operations. As holes in the spray are more difficult to track than the edges due to the more turbulent flows (holes are often very temporary and change shape dramatically between exposures), filling operations are used to avoid calculating velocities in central regions of the spray. Conversely, just as holes are hard to track, droplets formed during break-up events and separated from the main body of the spray are relatively easy to track, due to their spatial isolation and predictable movement. For this reason, holes in the main body of the spray are filled, but isolated droplets are not eroded. An example segmentation can be seen overlayed onto an extracted FRAME-exposure in Fig. 5.6.

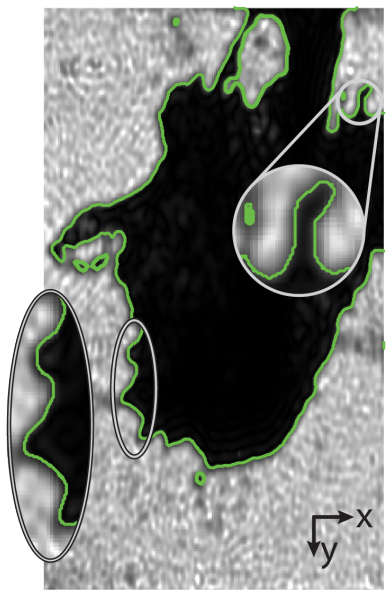


Figure 5.6: The segmentation of a spray, using an extracted exposure. Figure adapted from Paper III.

2. **Template matching along edges of segmentation.** After a satisfactory segmentation of the FRAME-triplet is achieved, a template (square window) is traced along the edge of the spray profile. In even intervals, a template matching is performed to the segmentation of the subsequent exposure in the FRAME-triplet, using normalized two-dimensional cross-correlation. In order to both speed up processing and avoid erroneous matches, a user input in the form of a search radius is also entered.

3. **Calculating a velocity.** Once the best match for the template has been found in the search window, a pixel distance is calculated corresponding to the displacement

between images. Knowing what physical distance this pixel displacement corresponds to, and the time interval between the exposures, the velocity for that region of the spray (the template) is calculated. The template is then moved along the edge of the first segmentation, repeating step 2 and 3 along the edges of the spray and droplets.

Since a FRAME-triplet consists of three exposures, this process is performed twice (excluding step 1) - once between exposure 1 and 2, and once between 2 and 3 - resulting in two velocity vector fields. By matching these to each-other, accelerations can be calculated using a final step

4. **Calculating accelerations from velocity vector fields.** By matching end-points of velocity vectors from the first velocity vector field to starting points in the second, a feature can be tracked over the entire FRAME-triplet. The two correlated velocities can then be used to calculate individual accelerations of features $a_{x,y}$ using

$$a_{x,y} = \frac{\Delta v_{x,y}}{2\Delta t} \quad , \tag{5.4}$$

where $\Delta v_{x,y}$ is the difference in velocity between the second and first matched velocity vectors in both image directions, and Δt is the time interval between each FRAME triplet (assuming the same time interval between exposures). The calculated accelerations can then be used to create an acceleration vector field for the FRAME-triplet.

5.2.4 Result: Validating the algorithm

In order to validate the algorithm used to calculate the velocities and accelerations, a separate measurement was performed with a known force acting on a rigid body - dropping a ball bearing, and letting gravity do its thing. The result of the measurement can be seen in Fig. 5.7 (a). The resulting distribution of vertical accelerations has a mean of $\mu_{ay} = 9.74$ m/s^2, with a standard deviation of $\sigma_{ay} = 13.4$ m/s^2. This high variance is caused by the uncertainty in the position of the edge established in the segmentation process, and corresponds to a spatial standard deviation of $\sigma_x = 2.07$ pixels, demonstrating the large impact that a small error in position has on the final calculated acceleration[2]. The implication of this analysis is that longer inter-pulse times leads to more consistent and therefore more accurate acceleration data, as the displacement becomes larger. However, this also makes the template matching harder.

[2]For a full derivation of this, see Appendix B of Paper III.

To verify that the acceleration variance indeed stems from the uncertainty in position, the algorithm was also tested on synthetic data with perfect segmentation (Fig. 5.7 (b)). In this simulated case it performs very well, with all acceleration vectors showing a value of 9.82 m/s^2. This measurement and analysis confirms that the accelerations and velocities calculated using the algorithm are reasonable.

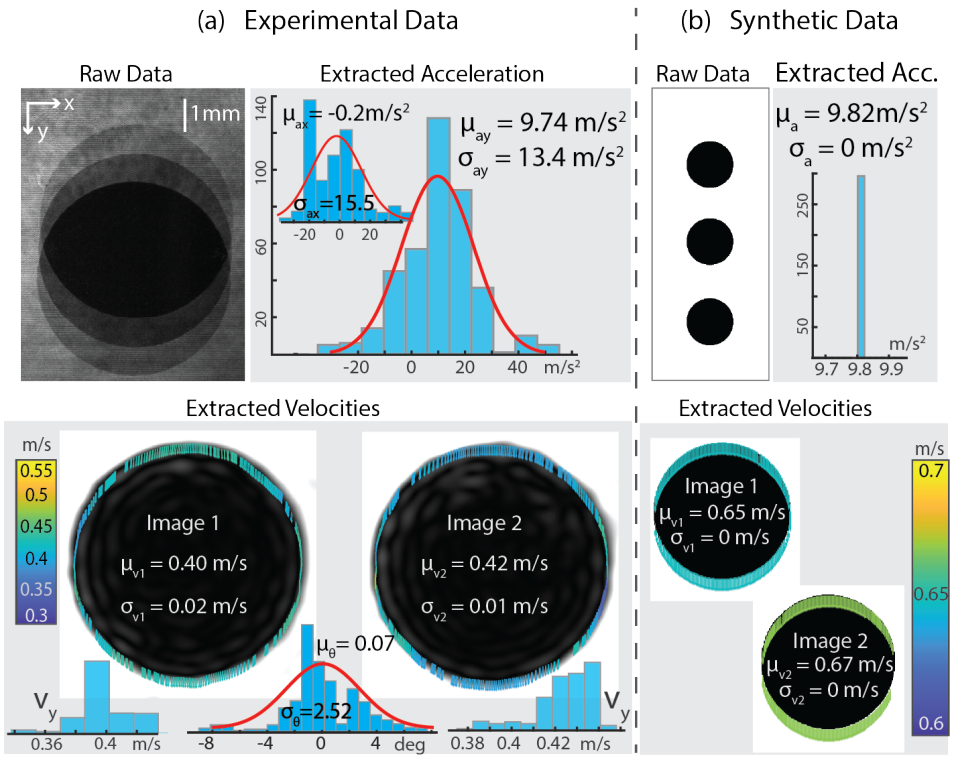


Figure 5.7: Calibration measurement for validating the acceleration extraction algorithm. **(a)** A FRAME-triplet of a falling ball bearing was processed using the algorithm to measure the gravitational acceleration. **(b)** A simulation of the same ball bearing was also performed. The low variance in the calculated accelerations of the simulated data show that the variance of extracted accelerations in the experimental data is caused by an uncertainty in position. Figure taken from Paper III.

5.2.5 Result: Measuring velocities and accelerations in a spray using temporal FRAME

The velocity- and acceleration vector fields resulting from multiplexed FRAME-triplets of a spray injection are shown in Fig. 5.8, overlayed onto the extracted exposures. Also shown are histograms of velocity- and acceleration distributions, indicating the overall dominating movement of the spray. At this point of the injection, accelerations at the leading edge of the spray are relatively low, indicating some sort of equilibrium

has set in between the force of the injection pressure and the air resistance. High accelerations can be observed on the left-hand side of the image, where air resistance tears apart the spray, starting a breakup event.

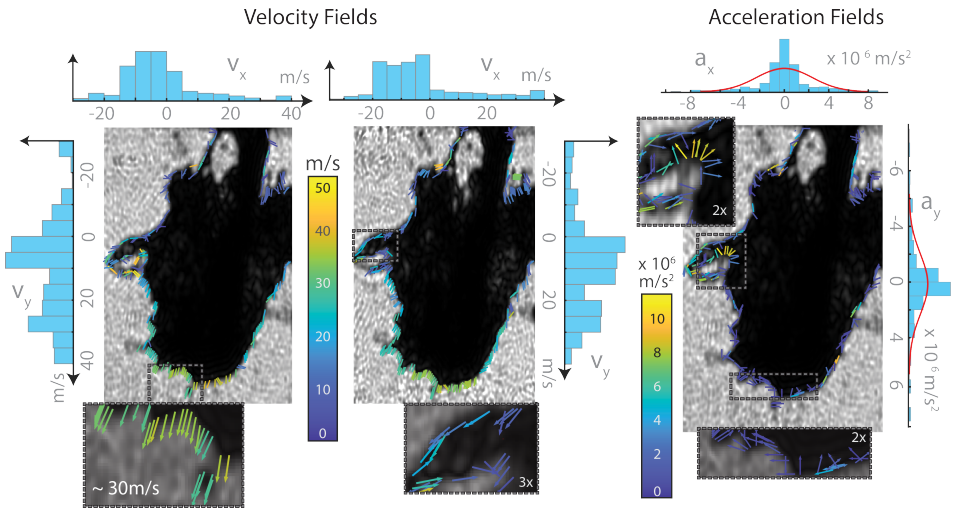


Figure 5.8: Extracted velocity and acceleration fields. The histograms displayed show the overall movement of the spray - downwards and slightly to the left in the image. The zoom-in of the acceleration field highlights the accelerations in a breakup event. Figure adapted from Paper III.

Fig. 5.9 zooms in on a breakup event (not from the same image as Fig. 5.8) generating a few droplets. As the highlighted droplet breaks free from the main body of the spray it can be seen to quickly decelerate due to air resistance. This deceleration is shown in the two velocity vector fields (left, red vectors) and the acceleration vector field (right, yellow vectors). The deceleration is on the scale of million meters per second squared (!), a number that we had to double-check both once and twice[3].

5.2.6 Movement along the optical axis

Perhaps the largest uncertainty factor in this type of analysis is the unknown movement along the optical axis. Here we once again run into "ye olde" limitation that imaging is a two-dimensional projection of a three-dimensional space. While there are methods of countering this issue, such as imaging at different focus depths, they would require that additional complex steps be added into an already complex imaging setup. For this reason, it is hard to motivate their addition, especially since the measurement method shown has already provided snapshot accelerometry.

[3]Although the acceleration is large, the mass of the droplet is quite small, leading to a force that is reasonable, albeit still large.

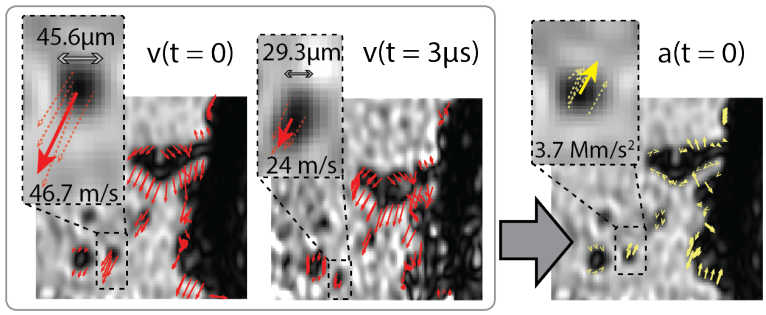


Figure 5.9: Zoom-in on individual droplets that have broken of from the main body of the spray. The dashed vectors show individually calculated vectors, while the larger, non-dashed vector display the average over the droplet. Over the duration of the FRAME-triplet, the droplet can be seen to decelerate due to air resistance. Figure adapted from Paper III.

The unknown movement along the optical axis likely causes artifacts in the velocity- and acceleration vector fields, as they provide an un-mappable[4] degree of freedom that lead to changes in the image projection of the spray. However, a lot of valuable information can still be extracted despite this, as the results show.

5.3 Quantifying the impact of data quality on neural network training

Machine learning-assisted processing has rapidly become the golden standard for image processing in a number of fields, including but not limited to medical imaging, remote sensing, microscopy and bioimaging, and astronomy [106–109]. Convolutional Neural Networks (CNNs) utilize trainable convolutional kernels, apt at identifying image features such as edges and textures, making them especially useful in these scenarios. While a lot of focus in put on how to improve the networks from an architectural point-of-view, they are also highly dependent on the quality of the data used to train them. This section discusses work I have done to highlight the impact of such an improvement in quality by training a well-established network architecture on two different datasets and comparing their relative performance.

[4] I do not know if this is a real word, but I will use it anyway as I think it is understandable.

5.3.1 Convolutional Neural Networks

Artificial Neural Networks (ANNs) are computational models that are trained on sets of data, with the goal to detect underlying patterns that can be used to extrapolate and make predictions on previously unseen but comparable data. They consist of computational units called nodes or neurons (hence the "neural" in ANN) which are organized into layers. The network learns by repeatedly adjusting trainable parameters called weights and biases, which are located in each node. A network consists of an input layer, into which the data is entered, a number of intermediate "hidden" layers where most of the data processing is done, and an output layer, which outputs the prediction of the network. Layers come in many different types, some of which will be discussed in this chapter. The size and type of layers and how they are organized is referred to as the network architecture, the most basic form of which is a sequential model, in which data is fed linearly from one layer to the next.

Convolutional neural networks are a subclass of ANNs which contain layers of trainable convolutional kernels, called convolutional layers[5]. Convolutional layers are particularly useful in image processing applications such as segmentation and object detection, as the convolutional kernels can be trained to find edges, textures, and even shapes. These features are very useful for tasks such as object detection (locating an object of a certain type in an image, such as "find the bike in this image of a street"), as objects in natural images are often separable by such features (change in texture, sharp edge, etc.). CNNs most often use a combination of layer types that complement each other well, helping to optimize the size of the network while simultaneously increasing its efficiency.

Mechanisms in training

A neural network can be trained by using data with or without an associated "correct answer" provided by a human supervisor. These types of training are referred to as supervised and unsupervised training respectively. Unsupervised training is an excellent tool to for example discover new dependencies in large datasets, but as this thesis concerns itself with images and image processing, the focus for this chapter will be on supervised learning.

Training a neural network for object detection requires correctly annotated data which can be used to provide feedback to the network during training and assessment. Such annotation is very often an arduous and tedious task. In order for the network to be able to make accurate predictions, the variance of the data used to train the network

[5]Who'd a thunk it?

need to accurately represent the data on which it will be applied. Therefore, the more (relevant) data the network has access to during training, the better its final performance. However, a low variance in the data also means less data is needed compared to a high variance.

For training, the corrected dataset is divided into parts. Most of the data (commonly around 80%) is put into a training dataset, while the remaining part is divided equally into validation- and test datasets. During the training process, the training dataset is given to the network in batches. After the network has made predictions for a batch, it is provided the correct answers and updates the trainable parameters of the network to better adjust to the new information provided. This process is called back propagation, and how radically the parameters are changed depends on learning rate - a parameter set by the user. Training continues in this manner, until the network has been fed the entire training dataset. This marks the end of what is known as an epoch, at which point the performance of the network is evaluated using the validation dataset not previously seen by the network. This quantifies the prediction capabilities of the trained network, but is not used for back propagation. Once training is done (can be done over many epochs), the final test of the network's extrapolative strength is performed, using the hitherto unseen test dataset.

The feedback to the network is provided using a customizable loss function, whose purpose is to quantify how wrong the network was during the last batch of samples. The choice of loss function is important and varies depending on the ultimate goal of the network, as it affects how weights are adjusted.

Extrapolation vs memorization, and when to end training

If the network is big enough compared to the complexity of the problem it is trying to solve and the size of the dataset, it can achieve 100% accuracy on the training dataset by memorizing the ground-truth of all samples in it. This is known as overfitting, and leads to perfect accuracy for the training dataset, but very poor performance on the two other datasets, as the network has learned the correct answers instead of the underlying patterns. Therefore, the stopping point for training is generally determined by monitoring the loss of the validation set instead, as this gives a good indication of when the extrapolating capabilities of the network are the best. Fig. 5.10 (a) shows the behavior of the training- and validation set when overfitting occurs. A way to discourage the network from memorizing samples while maintaining network complexity is to use dropout during training, which randomly excludes a certain portion of nodes on a batch-by-batch basis. This makes memorizing samples much harder, as the network never has access to all its nodes during training, and also has the effect of discouraging redundant nodes.

Common types of layers in a CNN trained for image processing

All layers contain a number of nodes, which may vary from layer to layer. The function of the layer determines the operation performed in its nodes. Simpler types of layers such as "fully connected" require less user inputs, while more complicated layers such as convolutional layers require more user-determined inputs.

Convolutional layers contain a number of trainable convolutional kernels, which are convolved with the input to produce the outputs of the layer, called feature maps. They are often located early in the network, to interpret the raw image inputs before they are distorted by any other layers. Typical user inputs for a convolutional layer are the number of nodes (kernels to be trained), size of kernels, stride, and padding options.

Max-pooling layers are downsampling layers which help to reduce the size of the samples. These are commonly placed between sets of convolutional layers, gradually reducing the complexity of the data and "boiling the problem down". This helps to decrease the computational load of the network. User input defines the size of the max-pooling kernel, padding, and stride.

Dense/fully connected layers consist of a number of neurons calculating a weighted sum from the inputs of all nodes of the previous layer. The trainable parameters of these nodes is the weights used in the summation, as well as the bias which is added to it. Dense layers are often located in the end of sequential CNNs as deducing layers, interpreting the information extracted from the convolutional layers before them.

Activation function (not a layer, but strongly related) Associated with each layer is an activation function, which is applied as a final step before output. Its function is to enable the layer to model non-linear relationships.

Fig. 5.10 (b) shows a schematic for the architecture of a sequential CNN. In the context of image analysis, this type of architecture can be used for image classification (2 classes, for example if there is a dog or a cat in the image). It utilizes a fully connected input layer, which is fed into a combination of convolutional- and max-pooling layers. The output of the last convolutional layer is fed to a segment of fully connected layers, resulting in two outputs of the network - the probability that the image is a dog and a cat respectively. Choosing the maximum argument of the output generates the prediction of the network.

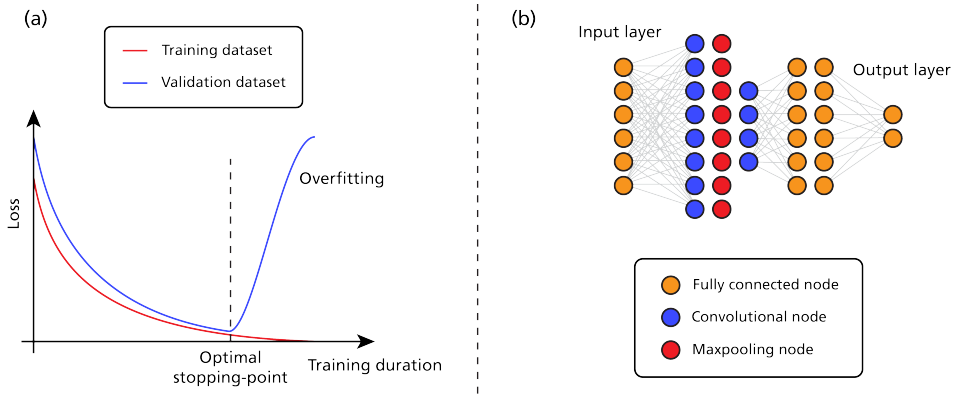


Figure 5.10: **(a)** Illustration displaying how the losses of the training- and validation datasets behave when a network is overfitted. The network has memorized the training dataset, meaning its loss goes to zero, but the loss when tested on the validation set sky-rockets due to the network's very poor extrapolation capabilities. **(b)** Schematic showing an example architecture of a Convolutional Neural Network used for classification (2 classes).

Hyperparameter tuning

Hyperparameters are the parameters of a network defined by the user, such as the number of layers and nodes in those layers, choice of activation functions and loss function, learning rate and batch size, to name a few. Optimizing these parameters gives the network the optimal conditions to succeed in its task, and is commonly done using methods like grid-searches, running through different combinations of parameters.

5.3.2 SILMAS

Structured Illumination Light-sheet Microscopy with Axial Sweeping, or SILMAS [110; 111], is a tomographic microscopy technique utilizing structured light-sheets to improve contrast and optical sectioning [112]. By shaping the excitation source into a homogeneous light-sheet and temporally scanning over the sample, techniques such as Axially Swept Light-sheet Microscopy (ASLM) [113] can produce three-dimensional images of fluorescently tagged structures. However, as these optical slices are imaged through intermediate material, the quality of imaged slices deteriorates as a function of depth along the optical axis due to scattering. SILMAS counters this effect by sinusoidally modulating the light-sheet and capturing three images of each slice with a 120° phase-shift between each modulation. As scattered light does not conserve this spatial structure, lock-in on the frequency of the modulation isolates the signal emanating from the optical section illuminated, suppressing scattered light, and in extension im-

proving contrast and optical sectioning. Each volumetric image slice I resulting from SILMAS can be calculated from the three phase-shifted images $I_{0,120,240}$ using [111]

$$I = \frac{\sqrt{2}}{3}\sqrt{(I_0 - I_{120})^2 + (I_0 - I_{240})^2 + (I_{120} - I_{240})^2}. \tag{5.5}$$

Experimental setup

The experimental setup used for SILMAS is shown in Fig. 5.11. First, the sinusoidal modulation is created through interference of the first orders produced by the diffractive optical element, and the light-sheet is then formed using the final cylindrical lens. By scanning the sheet over different heights, capturing three images of varying modulation phase at each, the three-dimensional image is built. The ASLM data was captured using the same setup, omitting the modulating of the light-sheet. Here I also want to clarify that this is not a setup that I have actually used myself, but as I have worked with the data produced with it, I think it is relevant to show how that data was collected.

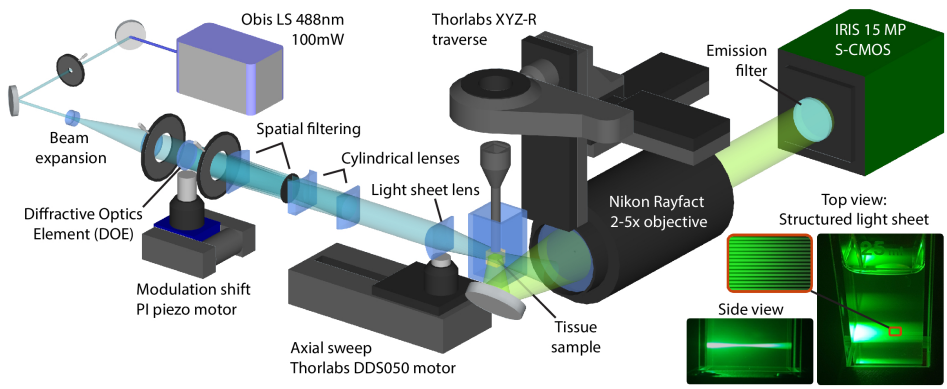


Figure 5.11: The experimental SILMAS/ASLM-setup used to capture the dataset used in the training of the two networks. By engaging or disengaging the modulating unit (dichroic and spatial filter), the optical setup can be used to capture either SILMAS or ASLM data. Image reused with permission from Frantz *et al.* [111].

5.3.3 Result: Comparison of network performances using SILMAS and ASLM datasets (Paper IV)

The motivation behind Paper IV was to concretize the impact that better raw data has on neural network-based processing. The study was performed using SILMAS and ASLM data of the same biological sample - chemically cleared mouse brain tissue

containing fluorescently tagged alpha-synuclein (aSyn) aggregations, a pre-cursor for Parkinson's disease [114]. As machine-learning approaches are extensively used for image processing in biomedical microscopy, it is important to show the impact that data quality has on the final result. In the current research climate, a lot of focus is put on neural networks as an almighty solution able to make use of any data, while the truth is that their success is very much dependent on the quality of the data they are trained on. To demonstrate this effect, two datasets consisting of three-dimensional image subsets (25^3 voxels) were sampled from the same biological tissue sample using different acquisition methods (SILMAS and ASLM). Annotating the same ground-truth to each subset pair (based on the inspection of the same subsets of the samples simultaneously) ensures a fair and direct comparison between them.

In order to ensure relevance within the field, the comparison was made using a U-net architecture [115] - a well-documented and successful Fully Convolutional Network[6] (FCN) developed specifically for biomedical image segmentation. The architecture was adapted for three-dimensional image subsets, and can be seen in Fig. 5.12.

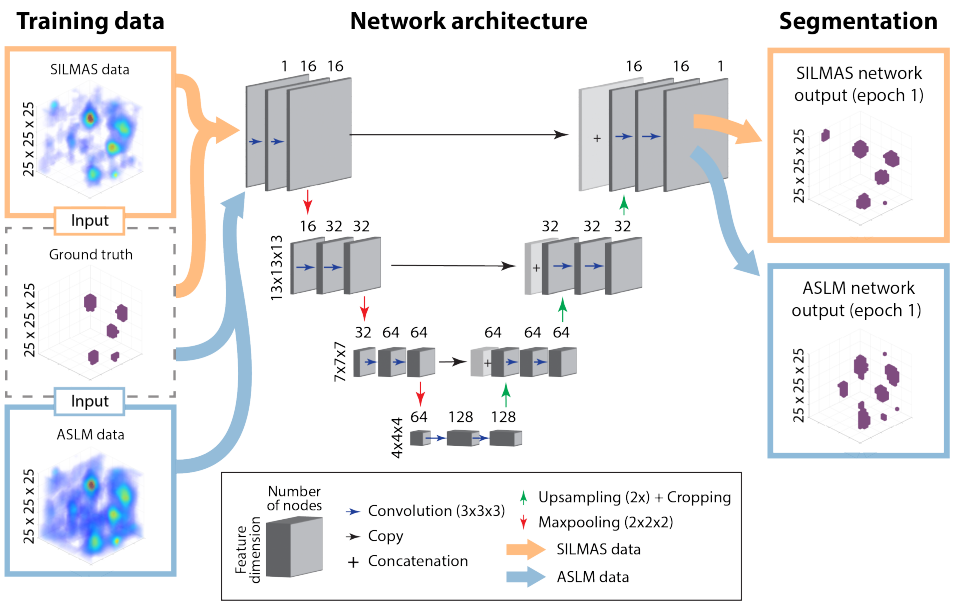


Figure 5.12: Network architecture and input/output of the network. The U-net architecture was adapted for 3D-image subsets. The input to the network is displayed on the left. Each subset (sample in the dataset) consists of 25^3 voxel data cubes, where intensity values correspond to concentrations of the tagged aSyn. A ground-truth segmentation shared between the capturing methods is also inputted. The output of the network is a predicted segmentation with the same 25^3 volume, directly comparable to the ground-truth value.

[6] A FCN is similar to a CNN, but does not use any fully connected layers.

The task of the network was to segment the globular aggregations of tagged aSyn (foreground) from the non-fluorescing cleared tissue (background). As this is a case of two-class segmentation, a binary cross-entropy loss function was used during training. In this loss function, the two classes were also weighted as to compensate for the prevalence of background in relation to foreground. The performance of the two networks[7] was evaluated using the F_1-score, defined as

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad , \tag{5.6}$$

where TP, FP and FN are true positives, false positives and false negatives respectively[8]. The F_1-score is a popular metric within medical research as it weighs in aspects of false diagnosis, which is of high ethical importance.

In order to more accurately reflect a real-world application - where perfect segmentation of a fluorescent cluster is not as important as correctly identifying points of aggregation - the prediction capabilities of the network was evaluated using the F_1-score for correctly identified clusters, instead of on an individual voxel level. In this context, a single object (point of aggregation) is defined as any group of 26-connected voxels, and a true positive is defined as the predicted segmentation having any overlap with the ground-truth segmentation. The precision (what fraction of predicted positives are correctly identified), recall (what fraction of the actual positives are correctly identified) and F_1-score for this type of object detection can be seen in Fig. 5.13 averaged over a number of runs. These plots were produced by extracting the network weights after each epoch, to independently evaluate the prediction capabilities as a function of training time. The scores for the first 40 epochs are shown, after which the training reaches a plateau.

The network trained on SILMAS data can be seen to a have a noticeable edge in all metrics over the entire training process. Of particular notice is the difference in precision, meaning the network trained on SILMAS data has fewer false positives. The results in Fig. 5.13 show that the improved contrast and optical slicing of SILMAS has a tangible impact on the prediction performances, even when golden-standard neural network architectures are used. In a research climate where a lot of focus is put on improving the machine learning-aspects of analysis, this highlights the importance of parallel improvement of measurement techniques to also improve the data quality on which the network is relying.

[7] Referring to the two instances of U-net trained on ASLM and SILMAS-data respectively.

[8] On an individual voxel level.

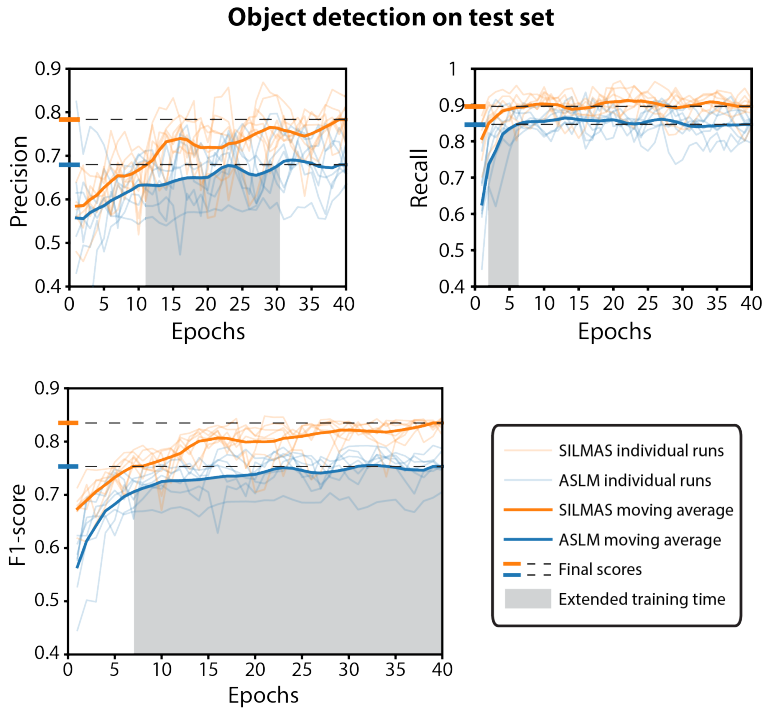


Figure 5.13: Compared performance of object detection of the networks trained on ASLM and SILMAS data respectively. The network trained on SILMAS data performs better in terms of both precision and recall. Combining the two using the F_1-score demonstrates that SILMAS data provides a "head start" which is maintained over the duration of the training. Figure taken from Paper IV.

5.4 Spectral imaging using intensified cameras

Image intensifiers artificially increase the signal of images by converting the incident photons to photo-electrons. These photo-electrons are passed through a microchannel plate (MCP) which multiplies the number of photo-electrons in what is known as an "electron avalanche", thereby increasing the signal[9] [116]. The electrons are then converted back to photons using a phosphor plate, which are registered by the sensor. Image intensifiers are useful in imaging scenarios where the photon-yield is too low for any signal to be registered using only the native sensitivity of the sensor.

The mechanics of an image intensifier prevents the effective use of integrated multispectral solutions, such as MSFAs. Filter array solutions need to be located near the sensor to avoid cross-talk (often also utilizing a microlens array to focus the light on the photosite). Thus, the image intensifier needs to be located before this stage,

[9]It should also be said that just as intensifiers amplify the signal, they also amplify any photon-related noise, as well as introduce their own sources of noise.

which means the spectral information of the incoming photons is lost in the photon to photo-electron conversion, and spectral filtering will not work. Consequentially, multispectral capture with intensifiers requires other solutions. A potential solution is to use a multi-camera setup, but this is very cost-ineffective, especially when increasing the number of spectral channels. Spectral multiplexing using FRAME provides a cost-effective single-sensor solution to this problem.

5.4.1 Multispectral FRAME and intensified cameras

The optical setup for spectral multiplexing presented in Chapter 4 provides a camera-independent, light-conserving and spectrally flexible method of capturing wide spectral bands using a monochrome sensor. All of these factors makes it suitable for applications that require intensified imaging. The optical multiplexing setup is simply placed as an intermediate imaging step between the object and the intensified camera, and the spectral channels are separated in post, as described in Chapters 3 and 4. This type of application was demonstrated in Paper I by imaging a pulsed plasma jet, which occur on the timescale of nanoseconds, requiring a time-gated intensified camera.

5.4.2 Result: Species imaging in a pulsed plasma jet using multispectral FRAME and an intensified camera

Fig. 5.14 (a) shows the experimental setup used to create the pulsed plasma jet. A steady flow of helium (He) was injected into air at ambient air pressure. A pulsed electric field was created between the grounding plate and the He-injection nozzle by supplying a high voltage to the electrode needle located inside the nozzle. The electric field dissociates electrons from atoms, ionizing the gas-mixture. The subsequent recombination results in line emissions corresponding to specific species generating the corona discharge seen in Fig. 5.14 (b) and (c) (false-color), which occurs on a timescale of nanoseconds.

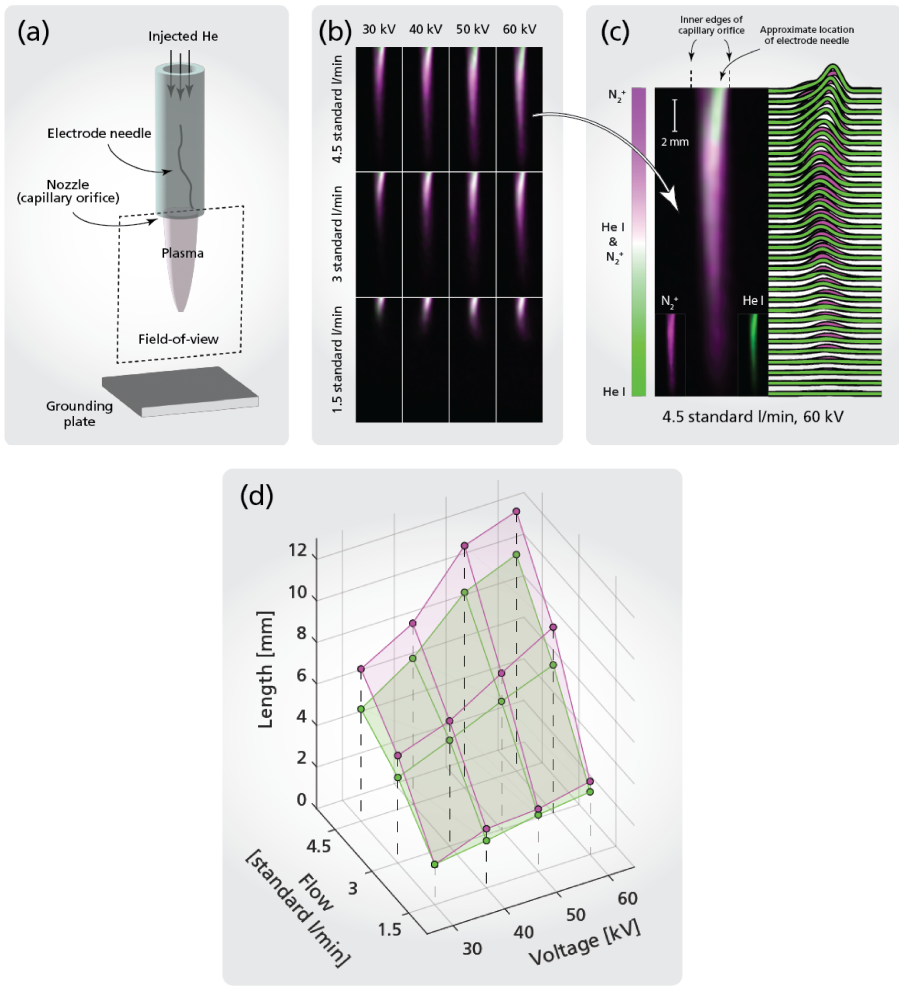


Figure 5.14: **(a)** The experimental setup used to generate the pulsed jet plasma. **(b)** False-colored images of the spatial distributions of He and N in the parameter study, using different gas flows and electrode voltages. **(c)** Highlight of the [4.5 standard l/min, 60 kV] case, also showing the intensity profile of the two species. **(d)** Surface plot showing the length of the jet as a function of both gas flow and voltage. Figure adapted from Paper I.

For this measurement, the emission of neutral helium (He I) and ionized nitrogen (N_2^+) were spectrally separated using multispectral FRAME in combination with an intensified camera. This results in spatial distributions of the two species, seen in Fig. 5.14 (b) and (c) captured at varying gas flows and applied voltages. The length of the plasma jet is defined as the distance from the nozzle to when its intensity has decayed to $e^{-1} \cdot I_{max}$, where I_{max} is the maximum value. Using this, the surface plot of Fig. 5.14 (d) can be created, showing how the length of the plasma jet varies for the two species as a function of both applied voltage and gas flow.

From the perspective of somebody working with method development, the main take-away from this measurement is, according to me, the proof-of-concept aspect: There are no intrinsic properties of multispectral FRAME preventing it from being used with intensified cameras. The one potential downside of frequency filtering in the case of intensified imaging is that the hexagonal honeycomb pattern of the MCP used to intensify the image can sometimes[10] contribute with high frequency components in the image. This leads to potential frequency band overlap between the MCP-pattern and the multiplexed exposures in the Fourier spectra. This is primarily an issue in darker regions of the image, and can be countered by reducing the gain of the intensifier or at least partially be calibrated for using a flat-field measurement. That being said, in this particular measurement it did not prove to be a major hindrance.

5.4.3 Unpublished work: Multispectral FRAME in the ultra-violet region, using an intensified camera

As a brief side-note, I also want to mention an unpublished work that I have worked with in this section, as it also used an intensifier in combination with multispectral FRAME. In this project, we aimed to spectrally separate and image the ultra-violet (UV) fluorescence of toluene and formaldehyde in a jet flame to characterize the combustion. The optical multiplexing setup was reminiscent of that in Fig. 4.2, only using two channels instead of four, and a dichroic mirror for recombination of the optical channels[11]. Unfortunately this work did not culminate in any publication, as the signal strength in this particular imaging scenario was not high enough. However, it was still the first time multispectral FRAME was applied in UV, showing that there are no intrinsic mechanisms of multispectral FRAME preventing it from being applied in this region of the electromagnetic spectrum.

5.5 Multispectral microscopy

At the time of writing this thesis, I am also working on combining the dichroic multispectral setup with microscopy, the optical setup for which is shown in Fig. 5.15. It uses a combination of the previously discussed dichroic design and an interchangeable microscopic objective, and has been used with both broadband bright-field and laser illumination. This new setup has been developed with the particular application of fluorescence microscopy in mind. For this reason, the number of spectral channels has also been increased to 7, increasing the maximum number of unique

[10] At low signal and/or high gain.

[11] Of course, all optical components were also appropriate for UV.

fluorophores differentiable through linear unmixing to the same. Fluorescence microscopy imaging is often performed using a spectral scanning approach, sequentially switching between spectral filters and excitation sources [47]. Multispectral FRAME offers spectrally flexible snapshot imaging at the cost of reduced spatial resolution, reducing the time needed to capture the scene. Moreover, combined with modulated illumination sources, it can be used to differentiate between fluorophores exhibiting near identical emission spectra, as demonstrated in Paper V.

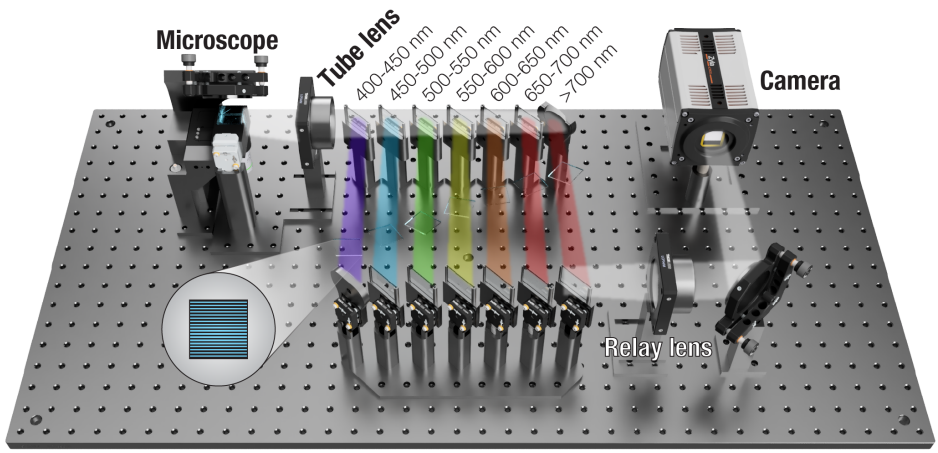


Figure 5.15: The 7 channel optical multiplexing setup used for multispectral microscopy. The setup is very similar to Fig. 4.3, with the addition of a microscopic lens. The setup can be used with brightfield illumination or laser excitation of the sample. Figure taken from Paper VI.

5.5.1 Linear unmixing

Linear unmixing is a mathematical method used to separate overlapping spectral signatures [117; 118]. As fluorophores have broad and often overlapping emission spectra, linear unmixing is frequently used within fluorescence microscopy to visualize the spatial distribution and intensity of fluorophores when their spectra can not be separated into different spectral bands.

Mathematically, linear unmixing is a change of spectral basis vectors, from the spectral basis vector determined by the capturing method to the spectral signatures of each fluorophore. For the same reason, the method cannot be used to differentiate between more spectral signatures than spectral bands captured. To perform the basis change and find the relative intensities $\mathbf{f}$ of each fluorophore in a pixel, given the intensity values $\mathbf{i}$ in each spectral channel and a calibration matrix $\mathbf{M}$, the equation

$$\mathbf{f} = \mathbf{M}^{-1}\mathbf{i} \quad , \quad \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \dots \end{bmatrix} \quad , \quad \mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \dots \\ m_{21} & m_{22} & \dots \\ \dots & \dots & \dots \end{bmatrix} \quad , \quad \mathbf{i} = \begin{bmatrix} i_1 \\ i_2 \\ \dots \end{bmatrix} \tag{5.7}$$

is solved, where each column in $\mathbf{M}$ is the spectral signature of a unique fluorophore.

5.5.2 Result: 7-channel multispectral snapshot microscopy

The microscopy setup has been used to perform both fluorescence microscopy (laser-induced fluorescence, fluorescent beads), bright-field microscopy (broadband backlight) and microscopy of a light-emitting display. The imaging has been performed at two different magnifications (4x and 20x) using the interchangeable microscopy objectives.

Laser-induced fluorescence microscopy

Two types of fluorescent beads, differentiable both by size (27 µm and 48 µm respectively) and fluorophore emission spectra, were imaged at 20x magnification using the microscopic setup. Fig. 5.16 shows the raw image and its Fourier spectrum, as well as each extracted color channel after lock-in (false-colored according to wavelength). Examining the response of each fluorophore in each spectral channel reveals rough emission spectra which peak in green and blue respectively. Using these signatures as the matrix $\mathbf{M}$ in the linear unmixing Eq. 5.7 reveals the relative contribution of each fluorophore in each pixel.

Light-emitting display

Fig. 5.17 shows the spectral decomposition of a light-emitting display, imaged at 20x magnification. The display consists of a mosaic of RGB-emitting diodes, whose spectral signatures through the system can be seen to the right in the figure. The first and last channel were omitted due to low signal-to-noise ratio. Linear unmixing using the three spectral signatures reveals the three types of diodes, false-colored according to wavelength.

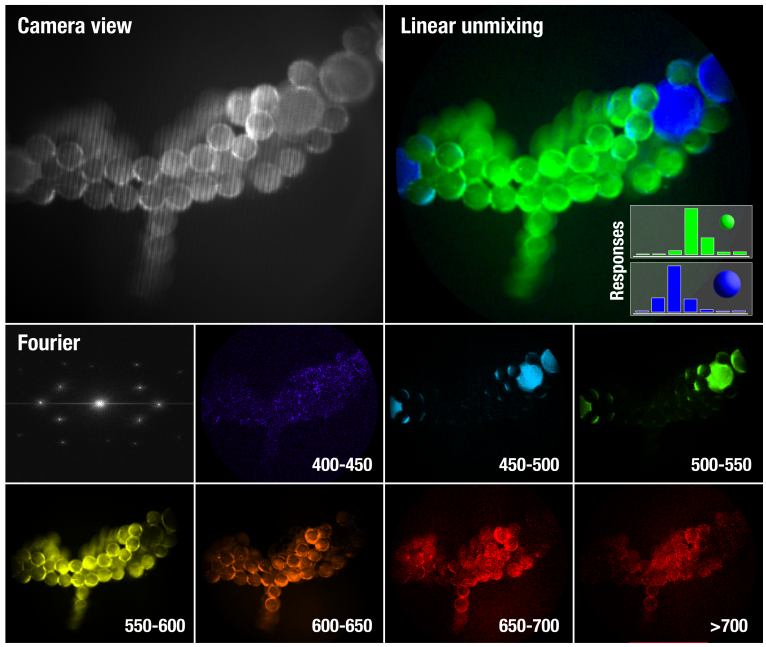


Figure 5.16: Multispectral image of laser-induced fluorescence in beads, captured using the multispectral microscopy setup. Using linear unmixing, the two types of beads can be separated, visualized through false coloring (top right). Also shown in this subimage is the spectral response in each channel for both fluorophores. Figure adapted from Paper VI.

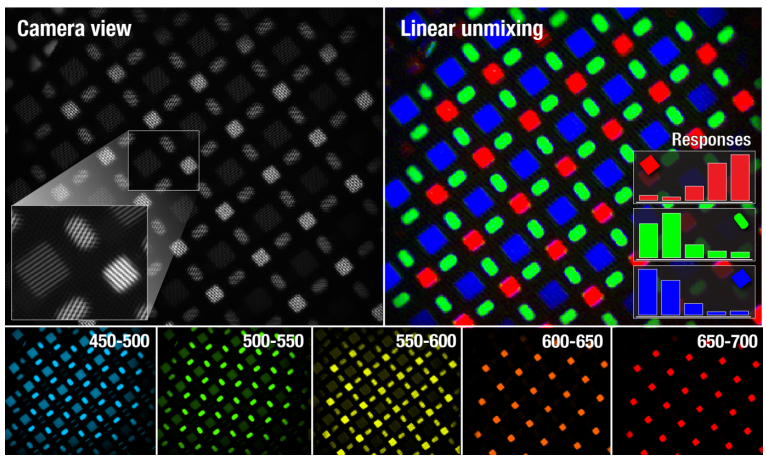


Figure 5.17: Multispectral image of a light-emitting display, where the individual diodes are visible. Using linear unmixing, the three types of diodes (red, green and blue) have been separated and false colored for visualization (top right). The spectral response of each diode is also shown. Figure taken from Paper VI.

Bright-field microscopy

The same beads were also imaged at two different magnifications (4x and 20x) using a bright-field scheme. A broadband light-source located behind the object plane illuminated the beads in Fig. 5.18, enabling measurements of relative absorption at different spectral bands. The reddest channel was omitted in the analysis due to its low signal-to-noise ratio, resulting from the weak output of the illumination source at these wavelengths.

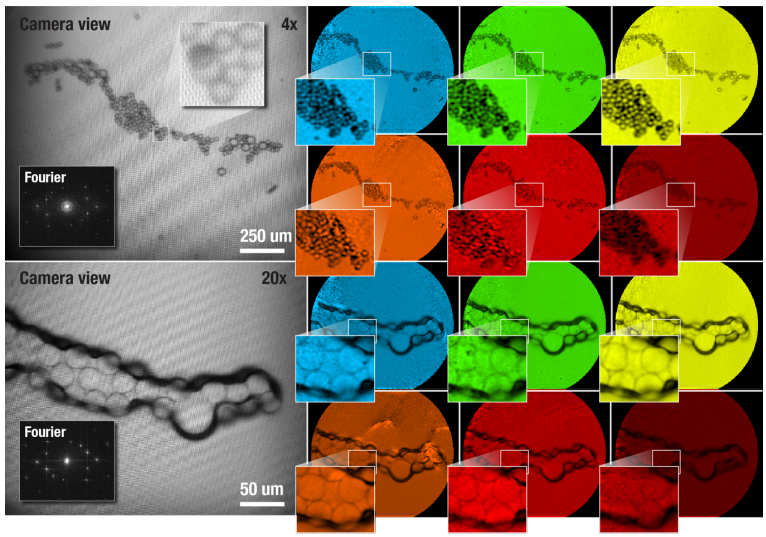


Figure 5.18: Bright-field multispectral image of beads at two different magnification, 4x and 20x, demonstrating the capabilities of the setup to both image at different magnifications and measure absorption in different bands. Figure taken from Paper VI.

Blood cells (typically ∼8 µm [119]) were also imaged using the bright-field scheme to test the system's ability to resolve fine structures close to the optical resolution limit. Two spectral channels were used, allowing for a larger filter radius to be used to, to test if the blood cells could be resolved with minimal resolution impact from the FRAME algorithm. The raw camera image as well as the two extracted exposures (500-550 nm and 600-650 nm) can be seen in Fig. 5.19. The zoom-ins clearly show the system's ability to resolve these structures, demonstrating its potential for future application within the biomedical field. The eccentricity of a few of the visible blood cells are also measured - a precursor to pathological conditions such as infectious diseases and blood disorders [120].

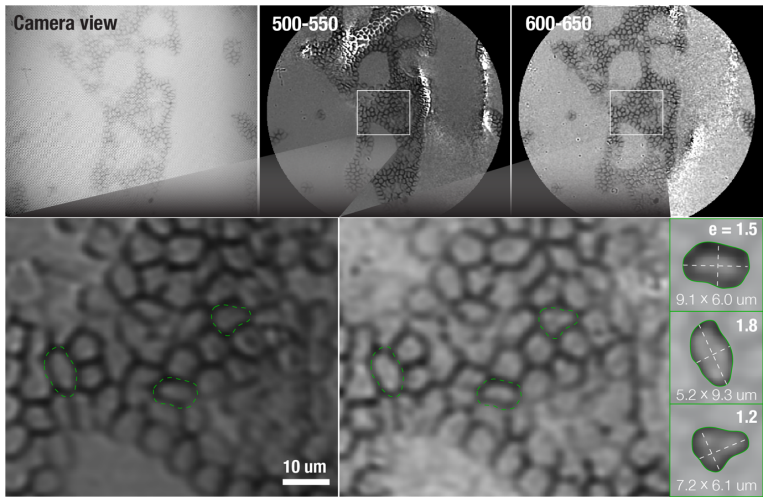


Figure 5.19: Red blood cells imaged in two spectral bands, demonstrating the system's ability to resolve fine structures. The shape of three red blood cells have been isolated to measure their eccentricity. Figure taken from Paper VI.

Chapter 6

Outlook and Conclusion

6.1 Outlook

Looking ahead in relation to my own work with multispectral FRAME, I see a few paths which are yet to be explored. Some are well under way, like the work currently being done with microscopy, and others are lying in wait to be tried by somebody else. I would like to raise three of them here, to peak the interest of readers from other fields and inspire those that are interested in working with FRAME's continued development.

First, the double modulation approach to FRAME presented in Paper V, where two parameters are simultaneously multiplexed in a single image, is not something I have explored further, but which has a lot of potential. In the case of Paper V, it was used to encode the response from fluorophores both as a function of excitation and emission spectra, which helped distinguish between fluorophores with very similar emission characteristics. The power of this technique lies in the scaling of additional information bands. When using double modulation, the amount of exposures containing unique information scales quadratically, in contrast to the linear scaling of single modulation. The potential combinations of parameters to be encoded is also great. Imagine combining the powers of active temporal FRAME with passive multispectral FRAME, to create a spectrally resolved high-speed video sequence in a single camera image. I believe venturing further down this path would result in technically impressive measurements and results, and I highly encourage the continued exploration of this area of FRAME.

My second point is concerning pyrometry. As I wrote in the introductory part of Section 5.1, the bulk of the blackbody radiation is emitted in the infra-red. In my

own work, I have used FRAME in the visible and ultra-violet region, and so maybe the next step is to try to apply the pyrometry approach described in this work in the infrared? The light-efficiency of the dichroic setup I have developed would come to good use when measuring in lower temperature ranges, as the total radiance of a blackbody decreases with temperature. Infra-red imaging of course comes with its own set of challenges such as thermal dark noise [121], but it would be exciting to see if it is possible in practice, as I see no theoretical limitations with regards to the FRAME technique.

Third and last is the continuation of what I am currently working on, namely fluorescence microscopy. Passive FRAME has already been shown to be an effective snapshot method for spectrally classifying and separating fluorophores [79; 80], and we have now also shown that it is possible to combine this powerful tool with microscopy imaging. We hope that this will open up for new applications within the field of biomedicine. One such potential application is cytometry [122], the analysis and characterization of cells. Imaging flow cytometry [37; 123] characterizes cells in a flow based on both spectral properties and cell morphology (shape, size, textures, etc.) - maybe FRAME could be used to make such an imaging technique more spectrally flexible? The world of frequency multiplexed imaging is still in its infancy, and personally, I can not wait to see what is next to come!

6.2 Conclusion

And so we arrive at the beginning of the end. My goal over the last 6 chapters has been to give my own view of the work I have been doing for the last 5 years - in part because that is what you as a reader expect and (hopefully) want from a text like this, but also because that is the only thing I can do. As I stated at the beginning of chapter 3, before venturing into the deep abyss that is the mathematics surrounding FRAME, I tend to prefer qualitative and practical discussions to equations. I think this trait shines through in a lot of ways in this work, and of this I am proud. I believe the most important task in science is making it understandable to a wider audience - that is where its value lies. In the case of a doctoral thesis, which I guess has to be called "cutting edge knowledge" even though it feels a little bit weird to call it that myself, the term "wide audience" is of course relative, but the point still stands. I hope that my ambition to make the subjects of this thesis accessible for as many as possible has succeeded, or at the very least that the effort has been noticed.

By this point, my hope is that the title has been sufficiently motivated by the content, but in case there is still confusion surrounding it I thought it fitting to explain it now.

After all, this is the last chance I have[1]. Beginning with the final words of the title, the common thread in all of the work I have done is coded light. Both FRAME and SILMAS, the experimental techniques I have worked with, use coded light to either expand on the dimensions captured in, or improve the quality of, camera images. As for the first two words, "Towards quantitative", almost all measurements I have participated in have resulted in the quantification of some physical quantity or quantitative comparison. To add to this, both the spatial dimensions and, more importantly, the *relative intensities* of exposures have been used to produce such results. To some this might be commonplace, but considering that FRAME is (1) a technique that requires manipulation of intensities even before light reaches the sensor in order to work, and (2) is still very much in a development phase, this progress is significant and something to be lifted. There is still some way to go and wrinkles to be ironed out, which is where the "towards" comes in, but the results presented demonstrate the viability of FRAME as a quantitative method. The pyrometry measurements demonstrate that multispectral FRAME can be used in conjunction with techniques utilizing relative intensities for quantitative measurements, and the spectral flexibility of FRAME and its ability turn any camera multispectral has been key to producing the temporally well-resolved temperature images in Paper II. This step forward in the development and application of the FRAME technique is something I am very proud of.

With this final statement, I would like to thank you for making it all the way to the end. This work has been a journey in many aspects, and I am grateful that I have gotten to share some of what I have learned with you. After all, what would be the use of all this work if I did not share it?

Snipp snapp snut, så var sagan om färgavbildning slut.

[1]...in writing. I can give no other guarantees.

Hard work is not hard when done with friends.

References

[1] Jamie Enoch, Leanne McDonald, Lee Jones, Pete R Jones, and David P Crabb. Evaluating whether sight is the most valued sense. *JAMA ophthalmology*, 137(11):1317–1320, 2019.

[2] George K. Hung, Kenneth J. Ciuffreda, Madjid Khosroyani, and Bai-Chuan Jiang. *Models of Accommodation*, page 287–339. Springer US, 2002.

[3] Robert S Molday and Orson L Moritz. Photoreceptors at a glance. *Journal of cell science*, 128(22):4039–4045, 2015.

[4] Dennis M. Levi. Learning to see in depth. *Vision Research*, 200:108082, November 2022.

[5] Christopher P Heesy. Seeing in stereo: The ecology and evolution of primate binocular vision and stereopsis. *Evolutionary Anthropology: Issues, News, and Reviews*, 18(1):21–35, 2009.

[6] Leo Anzagira and Eric R. Fossum. Color filter array patterns for small-pixel image sensors with substantial cross talk. *Journal of the Optical Society of America A*, 32(1):28, December 2014.

[7] Lauren O'Neill Shermer, Karen C. Rose, and Ashley Hoffman. Perceptions and credibility: Understanding the nuances of eyewitness testimony. *Journal of Contemporary Criminal Justice*, 27(2):183–203, April 2011.

[8] Auria Eisen-Enosh, Nairouz Farah, Zvia Burgansky-Eliash, Uri Polat, and Yossi Mandel. Evaluation of critical flicker-fusion frequency measurement methods for the investigation of visual temporal resolution. *Scientific Reports*, 7(1), November 2017.

[9] Howard G. Rogers. Beginnings of instant color photography. *Research Technology Management*, 31(5):42–45, 1988.

[10] R. Lukac and K.N. Plataniotis. Color filter arrays: design and performance analysis. *IEEE Transactions on Consumer Electronics*, 51(4):1260–1267, November 2005.

[11] Marc Wittmann and Virginie Van Wassenhove. The experience of time: neural mechanisms and the interplay of emotion, cognition and embodiment, 2009.

[12] Fabien Halter, Valentin Glasziou, Marco Di Lorenzo, Stany Gallier, and Christian Chauveau. Peculiarities of aluminum particle combustion in steam. *Proceedings of the Combustion Institute*, 39(3):3605–3614, 2023.

[13] Xin Bian, Changho Kim, and George Em Karniadakis. 111 years of brownian motion. *Soft Matter*, 12(30):6331–6346, 2016.

[14] Brian M. Tissue. Ultraviolet and visible absorption spectroscopy. *Characterization of Materials*, pages 1–13, 2012.

[15] Dale E Gary. New insights from imaging spectroscopy of solar radio emission. *Annual Review of Astronomy and Astrophysics*, 61(1):427–472, 2023.

[16] Elise C Fear. Microwave imaging of the breast. *Technology in cancer research & treatment*, 4(1):69–82, 2005.

[17] Fujin Hou, Yan Zhang, Yong Zhou, Mei Zhang, Bin Lv, and Jianqing Wu. Review on infrared imaging technology. *Sustainability*, 14(18):11161, 2022.

[18] Sevgi Türker-Kaya and Christian W Huck. A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. *Molecules*, 22(1):168, 2017.

[19] Xiangyu Ou, Xue Chen, Xianning Xu, Lili Xie, Xiaofeng Chen, Zhongzhu Hong, Hua Bai, Xiaowang Liu, Qiushui Chen, Lin Li, et al. Recent development in x-ray imaging technology: Future and challenges. *Research*, 2021.

[20] Marco Pinto. Prompt-gamma imaging in particle therapy. *The European Physical Journal Plus*, 139(10):884, 2024.

[21] Nils Genser, Jürgen Seiler, and André Kaup. Camera array for multi-spectral imaging. *IEEE Transactions on Image Processing*, 29:9234–9249, 2020.

[22] Megandhren Govender, Kershani Chetty, and Hartley Bulcock. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2):145–151, 2007.

[23] Gary A Shaw and Hsiaohua K Burke. Spectral imaging for remote sensing. *Lincoln laboratory journal*, 14(1):3–28, 2003.

[24] David H Foster, Kinjiro Amano, Sérgio MC Nascimento, and Michael J Foster. Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23(10):2359–2372, 2006.

[25] Stuart Phinn, Chris Roelfsema, Arnold Dekker, Vittoro Brando, and Janet Anstee. Mapping seagrass species, cover and biomass in shallow waters: An assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in moreton bay (australia). *Remote sensing of Environment*, 112(8):3413–3425, 2008.

[26] Bing Lu, Phuong D Dao, Jiangui Liu, Yuhong He, and Jiali Shang. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, 12(16):2659, 2020.

[27] Ivan A Vorobjev, Aigul Kussanova, and Natasha S Barteneva. Development of spectral imaging cytometry. *Spectral and Imaging Cytometry: Methods and Protocols*, pages 3–22, 2023.

[28] Tokuko Haraguchi, Takeshi Shimi, Takako Koujin, Noriyo Hashiguchi, and Yasushi Hiraoka. Spectral imaging fluorescence microscopy. *Genes to Cells*, 7(9):881–887, 2002.

[29] Minh H Tran and Baowei Fei. Compact and ultracompact spectral imagers: technology and applications in biomedical imaging. *Journal of biomedical optics*, 28(4):040901–040901, 2023.

[30] Anna Pelagotti, Andrea Del Mastio, Alessia De Rosa, and Alessandro Piva. Multispectral imaging of paintings. *IEEE Signal Processing Magazine*, 25(4):27–36, 2008.

[31] Petter Weibring, Thomas Johansson, Hans Edner, Sune Svanberg, Barbro Sundner, Valentina Raimondi, Giovanna Cecchi, and Luca Pantani. Fluorescence lidar imaging of historical monuments. *Applied Optics*, 40(33):6111–6120, 2001.

[32] JP Maillard, L Drissen, F Grandmont, and S Thibault. Integral wide-field spectroscopy in astronomy: the imaging fts solution. *Experimental Astronomy*, 35(3):527–559, 2013.

[33] Paul R Medwell, Qing N Chan, Peter AM Kalt, Zeyad T Alwahabi, Bassam B Dally, and Graham J Nathan. Instantaneous temperature imaging of diffusion flames using two-line atomic fluorescence. *Applied Spectroscopy*, 64(2):173–176, 2010.

[34] Paul R Medwell, Qing N Chan, Peter AM Kalt, Zeyad T Alwahabi, Bassam B Dally, and Graham J Nathan. Development of temperature imaging using two-line atomic fluorescence. *Applied Optics*, 48(6):1237–1248, 2009.

[35] Zheming Li, Jesper Borggren, Edouard Berrocal, Andreas Ehn, Marcus Alden, Mattias Richter, and Elias Kristensson. Simultaneous multispectral imaging of flame species using frequency recognition algorithm for multiple exposures (frame). *Combustion and Flame*, 192:160–169, 2018.

[36] Johan Sjöholm, Joakim Rosell, Bo Li, Mattias Richter, Zhongshan Li, Xue-Song Bai, and Marcus Aldén. Simultaneous visualization of oh, ch, ch20 and toluene plif in a methane jet flame with varying degrees of turbulence. *Proceedings of the Combustion Institute*, 34(1):1475–1482, 2013.

[37] Daniel Schraivogel, Terra M Kuhn, Benedikt Rauscher, Marta Rodríguez-Martínez, Malte Paulsen, Keegan Owsley, Aaron Middlebrook, Christian Tischer, Beáta Ramasz, Diana Ordoñez-Rueda, et al. High-speed fluorescence image–enabled cell sorting. *Science*, 375(6578):315–320, 2022.

[38] Xiaoxiao Feng, Luxiao He, Qimin Cheng, Xiaoyi Long, and Yuxin Yuan. Hyperspectral and multispectral remote sensing image fusion based on endmember spatial information. *Remote Sensing*, 12(6):1009, 2020.

[39] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of biomedical optics*, 19(1):010901–010901, 2014.

[40] Johannes Brauers, Nils Schulte, and Til Aach. Multispectral filter-wheel cameras: Geometric distortion model and compensation algorithms. *IEEE transactions on image processing*, 17(12):2368–2380, 2008.

[41] Ling Bei, Glenn I Dennis, Heather M Miller, Thomas W Spaine, and Jon W Carnahan. Acousto-optic tunable filters: fundamentals and applications as applied to chemical analysis techniques. *Progress in Quantum Electronics*, 28(2):67–87, 2004.

[42] Jon Y Hardeberg, Francis Schmitt, and Hans Brettel. Multispectral color image capture using a liquid crystal tunable filter. *Optical engineering*, 41(10):2532–2548, 2002.

[43] I. J. Cox and C. J. R. Sheppard. Information capacity and resolution in an optical system. *J. Opt. Soc. Am. A*, 3(8):1152–1158, Aug 1986.

[44] Fatima A. Merchant and Kenneth R. Castleman. 10.10 - computer-assisted microscopy. In AL BOVIK, editor, *Handbook of Image and Video Processing*

(Second Edition), Communications, Networking and Multimedia, pages 1311–XLIV. Academic Press, Burlington, second edition edition, 2005.

[45] Andre Perennou, Veronique Quintard, and Adan Omar Arellanes. Advanced control of an acousto-optic filter performance through arbitrary rf signals. *Applied Optics*, 62(21):5636, July 2023.

[46] Alexander Machikhin, Vladislav Batshev, and Vitold Pozhar. Aberration analysis of aotf-based spectral imaging systems. *Journal of the Optical Society of America A*, 34(7):1109–1113, 2017.

[47] Kun Chen, Rui Yan, Limin Xiang, and Ke Xu. Excitation spectral microscopy for highly multiplexed fluorescence imaging and quantitative biosensing. *Light: Science & Applications*, 10(1), May 2021.

[48] Alexander P Demchenko. Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection. *Methods and applications in fluorescence*, 8(2):022001, 2020.

[49] Kévin Barbieux. Pushbroom hyperspectral data orientation by combining feature-based and area-based co-registration techniques. *Remote Sensing*, 10(4):645, 2018.

[50] Igor Neri, Silvia Caponi, Francesco Bonacci, Giacomo Clementi, Francesco Cottone, Luca Gammaitoni, Simone Figorilli, Luciano Ortenzi, Simone Aisa, Federico Pallottino, et al. Real-time ai-assisted push-broom hyperspectral system for precision agriculture. *Sensors*, 24(2):344, 2024.

[51] Luc Oth, Paul Furgale, Laurent Kneip, and Roland Siegwart. Rolling shutter camera calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013.

[52] Li Zhang and Armin Gruen. Multi-image matching for dsm generation from ikonos imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(3):195–211, 2006. Extraction of Topographic Information from High-Resolution Satellite Imagery.

[53] D.P. Roy, M.A. Wulder, T.R. Loveland, Woodcock C.E., R.G. Allen, M.C. Anderson, D. Helder, J.R. Irons, D.M. Johnson, R. Kennedy, T.A. Scambos, C.B. Schaaf, J.R. Schott, Y. Sheng, E.F. Vermote, A.S. Belward, R. Bindschadler, W.B. Cohen, F. Gao, J.D. Hipple, P. Hostert, J. Huntington, C.O. Justice, A. Kilic, V. Kovalskyy, Z.P. Lee, L. Lymburner, J.G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R.H. Wynne, and Z. Zhu. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145:154–172, April 2014.

[54] Changyong Cao, Jack Xiong, Slawomir Blonski, Quanhua Liu, Sirish Uprety, Xi Shao, Yan Bai, and Fuzhong Weng. Suomi npp viirs sensor data record verification, validation, and long□term performance monitoring. *Journal of Geophysical Research: Atmospheres*, 118(20), October 2013.

[55] Andrew Butterfield, Gerard Ekembe Ngondi, and Anne Kerr, editors. *A Dictionary of Computer Science (7 ed.)*. Oxford University Press, 2016.

[56] Ghafour Amouzad Mahdiraji and Ahmad Fauzi Abas. *Advanced modulation formats and multiplexing techniques for optical telecommunication systems*. IntechOpen, 2010.

[57] Pierre-Jean Lapray, Xingbo Wang, Jean-Baptiste Thomas, and Pierre Gouton. Multispectral filter arrays: Recent advances and practical implementation. *Sensors*, 14(11):21626–21659, 2014.

[58] Johannes Brauers and Til Aach. A color filter array based multispectral camera. In *12. Workshop Farbbildverarbeitung*, pages 5–6. Ilmenau, 2006.

[59] Pilar Gonzalez, Klaas Tack, Bert Geelen, Bart Masschelein, Wouter Charle, Bart Vereecke, and Andy Lambrechts. A novel cmos-compatible, monolithically integrated line-scan hyperspectral imager covering the vis-nir range. In *Next-Generation Spectroscopic Technologies IX*, volume 9855, pages 129–137. SPIE, 2016.

[60] Bahadir K Gunturk, John Glotzbach, Yucel Altunbasak, Ronald W Schafer, and Russel M Mersereau. Demosaicking: color filter array interpolation. *IEEE Signal processing magazine*, 22(1):44–54, 2005.

[61] Andrew Stockman, Donald IA MacLeod, and Nancy E Johnson. Spectral sensitivities of the human cones. *Journal of the Optical Society of America A*, 10(12):2491–2521, 1993.

[62] Zoltán Kolláth, Andrew Cool, Andreas Jechow, Kornél Kolláth, Dénes Száz, and Kai Pong Tong. Introducing the dark sky unit for multi-spectral measurement of the night sky quality with commercial digital cameras. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 253:107162, 2020.

[63] Xiaochang Yu, Jia Hao, Junzhuo Zhou, Yang Su, Shahid Karim, and Yiting Yu. Modular snapshot multispectral-panchromatic imager (mspi) with customized filter arrays. *Optics Express*, 31(2):1475–1485, 2023.

[64] Motoki Yako, Yoshikazu Yamaoka, Takayuki Kiyohara, Chikai Hosokawa, Akihiro Noda, Klaas Tack, Nick Spooren, Taku Hirasawa, and Atsushi Ishikawa.

Video-rate hyperspectral camera based on a cmos-compatible random array of fabry–pérot filters. *Nature Photonics*, 17(3):218–223, 2023.

[65] HH Mai, A Albrecht, C Woidt, X Wang, V Daneker, O Setyawati, T Woit, K Schultz, M Bartels, and H Hillmer. 3d nanoimprinted fabry–pérot filter arrays and methodologies for optical characterization. *Applied Physics B*, 107(3):755–764, 2012.

[66] S Shi, L Wang, J Yin, and X Sun. Measurement of liquid-phase velocity field around taylor bubbles in slug flows. In *Japan-US Seminar on Two-phase Flow Dynamics, Sapporo, Japan*, 2017.

[67] Martin Hubold, Elisabeth Montag, René Berlich, Robert Brunner, and Robert Brüning. Multi-aperture system approach for snapshot multispectral imaging applications. *Opt. Express*, 29(5):7361–7378, Mar 2021.

[68] Jason G Dwight and Tomasz S Tkaczyk. Lenslet array tunable snapshot imaging spectrometer (latis) for hyperspectral fluorescence microscopy. *Biomedical optics express*, 8(3):1950–1964, 2017.

[69] Maksim Makarenko, Arturo Burguete-Lopez, Qizhou Wang, Fedor Getman, Silvio Giancola, Bernard Ghanem, and Andrea Fratalocchi. Real-time hyperspectral imaging in hardware via trained metasurface encoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12692–12702, June 2022.

[70] Daniel S. Jeon, Seung-Hwan Baek, Shinyoung Yi, Qiang Fu, Xiong Dun, Wolfgang Heidrich, and Min H. Kim. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Transactions on Graphics*, 38(4):1–13, July 2019.

[71] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[72] Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008.

[73] Ashwin A Wagadarikar, Nikos P Pitsianis, Xiaobai Sun, and David J Brady. Video rate spectral imaging using a coded aperture snapshot spectral imager. *Optics express*, 17(8):6368–6388, 2009.

[74] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2013.

[75] Xian-Hua Han, Jian Wang, and Huiyan Jiang. Recent advancements in hyperspectral image reconstruction from a compressive measurement. *Sensors*, 25(11), 2025.

[76] Ruixuan Zhao, Chengshuai Yang, R Theodore Smith, and Liang Gao. Coded aperture snapshot spectral imaging fundus camera. *Scientific Reports*, 13(1):12007, 2023.

[77] Xin Yuan, David J Brady, and Aggelos K Katsaggelos. Snapshot compressive imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 38(2):65–88, 2021.

[78] Bahadir K Gunturk and Martin Feldman. Frequency division multiplexed imaging. In *Digital Photography IX*, volume 8660, pages 174–180. SPIE, 2013.

[79] Karolina Dorozynska, Vassily Kornienko, Marcus Aldén, and Elias Kristensson. A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Opt. Express*, 28(7):9572–9586, Mar 2020.

[80] Karolina Dorozynska and Elias Kristensson. Implementation of a multiplexed structured illumination method to achieve snapshot multispectral imaging. *Opt. Express*, 25(15):17211–17226, Jul 2017.

[81] Elias Kristensson, Zheming Li, Edouard Berrocal, Mattias Richter, and Marcus Aldén. Instantaneous 3d imaging of flame species using coded laser illumination. *Proceedings of the Combustion Institute*, 36(3):4585–4591, 2017.

[82] Vassily Kornienko, Elias Kristensson, Andreas Ehn, Antoine Fourriere, and Edouard Berrocal. Beyond mhz image recordings using leds and the frame concept. *Scientific reports*, 10(1):16650, 2020.

[83] Simon Ek, Vassily Kornienko, Adrian Roth, Edouard Berrocal, and Elias Kristensson. High-speed videography of transparent media using illumination-based multiplexed schlieren. *Scientific Reports*, 12(1):19018, 2022.

[84] Vassily Kornienko, Elias Kristensson, Andreas Ehn, Antoine Fourriere, and Edouard Berrocal. Beyond mhz image recordings using leds and the frame concept. *Scientific Reports*, 10(1), October 2020.

[85] Chao Deng, Xuemei Hu, Jinli Suo, Yuanlong Zhang, Zhili Zhang, and Qionghai Dai. Snapshot hyperspectral imaging via spectral basis multiplexing in fourier domain. *Optics express*, 26(25):32509–32521, 2018.

[86] Elias Kristensson, Joakim Bood, Marcus Alden, Emil Nordström, Jiajian Zhu, Sven Huldt, Per-Erik Bengtsson, Hampus Nilsson, Edouard Berrocal, and Andreas Ehn. Stray light suppression in spectroscopy using periodic shadowing. *Optics express*, 22(7):7711–7721, 2014.

[87] Lokenath Debnath. A short biography of joseph fourier and historical development of fourier series and fourier transforms. *International Journal of Mathematical Education in Science and Technology*, 43(5):589–612, 2012.

[88] Naol Tufa Negero. Fourier transform methods for partial differential equations. *International Journal of Partial Differential Equations and Applications*, 2(3):44–57, 2014.

[89] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644, April 1928.

[90] Richard G Lyons. *Understanding digital signal processing, 3/E*. Pearson Education India, 1997.

[91] Eero P Simoncelli and Bruno A Olshausen. Natural image statistics and neural representation. *Annual review of neuroscience*, 24(1):1193–1216, 2001.

[92] David J Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379–2394, 1987.

[93] Daniel L Ruderman. The statistics of natural images. *Network: computation in neural systems*, 5(4):517, 1994.

[94] Vassily Kornienko, Yupan Bao, Joakim Bood, Andreas Ehn, and Elias Kristensson. The space-charge problem in ultrafast diagnostics: an all-optical solution for streak cameras. *Ultrafast Science*, 4:0055, 2024.

[95] Ek, Simon. *Optical Multiplexing for Ultrafast Videography*. PhD thesis, Lund University, 2025.

[96] Simon Ek, Vassily Kornienko, and Elias Kristensson. Long sequence single-exposure videography using spatially modulated illumination. *Scientific Reports*, 10(1):18920, 2020.

[97] Kornienko, Vassily. *Structured Light for Ultrafast Videography*. PhD thesis, Lund University, 2024.

[98] Wang Yongqing, Gu Zongqing, Wang Shuonan, and He Ping. The temperature measurement technology of infrared thermal imaging and its applications

review. In *2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, pages 401–406. IEEE, 2017.

[99] Eric W. Hansen and Phaih-Lan Law. Recursive methods for computing the abel transform and its inverse. *J. Opt. Soc. Am. A*, 2(4):510–520, Apr 1985.

[100] JW Andersen and RS Fein. Measurements of normal burning velocities and flame temperatures of bunsen flames. *The Journal of Chemical Physics*, 17(12):1268–1273, 1949.

[101] KA Trowell, S Goroshin, DL Frost, and JM Bergthorson. Aluminum and its role as a recyclable, sustainable carrier of renewable energy. *Applied Energy*, 275:115112, 2020.

[102] Wu, Zhiyong. *Aluminum droplet combustion studies using spatiotemporal diagnostics*. PhD thesis, Lund University, 02 2025.

[103] Shankar Krishnan, Louis Hennet, Sandro Jahn, Thomas A Key, Paul A Madden, Marie-Louise Saboungi, and David L Price. Structure of normal and supercooled liquid aluminum oxide. *Chemistry of materials*, 17(10):2662–2666, 2005.

[104] Mohammed Raad Abdulwahab, Yasir H Ali, Fatima J Habeeb, Abdoulhadi A Borhana, Ahmed M Abdelrhman, and Salah M Ali Al-Obaidi. A review in particle image velocimetry techniques (developments and applications). *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, 65(2):213–229, 2020.

[105] Andreas Ehn, Joakim Bood, Zheming Li, Edouard Berrocal, Marcus Aldén, and Elias Kristensson. Frame: femtosecond videography for atomic and molecular dynamics. *Light: Science & Applications*, 6(9):e17045–e17045, 2017.

[106] Alexander Selvikvåg Lundervold and Arvid Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift fuer medizinische Physik*, 29(2):102–127, 2019.

[107] Lei Ma, Yu Liu, Xueliang Zhang, Yuanxin Ye, Gaofei Yin, and Brian Alan Johnson. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS journal of photogrammetry and remote sensing*, 152:166–177, 2019.

[108] Erick Moen, Dylan Bannon, Takamasa Kudo, William Graf, Markus Covert, and David Van Valen. Deep learning for cellular image analysis. *Nature methods*, 16(12):1233–1246, 2019.

[109] Christopher J Fluke and Colin Jacobs. Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2):e1349, 2020.

[110] David Frantz, Tugba Karamahmutoglu, Allison J Schaser, Deniz Kirik, and Edouard Berrocal. High contrast, isotropic, and uniform 3d-imaging of centimeter-scale scattering samples using structured illumination light-sheet microscopy with axial sweeping. *Biomedical Optics Express*, 13(9):4907–4925, 2022.

[111] David Frantz, Courtney J. Wright, Allison J. Schaser, Deniz Kirik, Elias Kristensson, and Edouard Berrocal. Image post-processing for silmas: structured illumination light sheet microscopy with axial sweeping. *Biomed. Opt. Express*, 15(8):4943–4962, Aug 2024.

[112] Frantz, David. *SILMAS Structured Illumination Light-sheet Microscopy with Axial Sweeping: 3D imaging of brain tissue*. PhD thesis, Lund University, 2024.

[113] Kevin M Dean, Philippe Roudot, Erik S Welf, Gaudenz Danuser, and Reto Fiolka. Deconvolution-free subcellular imaging with axially swept light sheet microscopy. *Biophysical journal*, 108(12):2807–2815, 2015.

[114] Kelvin C. Luk, Victoria Kehm, Jenna Carroll, Bin Zhang, Patrick O'Brien, John Q. Trojanowski, and Virginia M.-Y. Lee. Pathological alpha-synuclein transmission initiates parkinson-like neurodegeneration in nontransgenic mice. *Science*, 338(6109):949–953, 2012.

[115] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[116] Thierry Gys. Micro-channel plates and vacuum detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 787:254–260, 2015.

[117] Timo Zimmermann, Jens Rietdorf, and Rainer Pepperkok. Spectral imaging and its applications in live cell microscopy. *FEBS letters*, 546(1):87–92, 2003.

[118] Timo Zimmermann. Spectral imaging and linear unmixing in light microscopy. *Microscopy Techniques: -/-*, pages 245–265, 2005.

[119] PB Canham and Alan C Burton. Distribution of size and shape in populations of normal human red cells. *Circulation research*, 22(3):405–422, 1968.

[120] Monica Diez-Silva, Ming Dao, Jongyoon Han, Chwee-Teck Lim, and Subra Suresh. Shape and biomechanical characteristics of human red blood cells in health and disease. *MRS bulletin*, 35(5):382–388, 2010.

[121] Chunlei Yu, Xue Li, Bo Yang, Songlei Huang, Xiumei Shao, Yaguang Zhang, and Haimei Gong. Noise characteristics analysis of short wave infrared ingaas focal plane arrays. *Infrared Physics & Technology*, 85:74–80, 2017.

[122] Abhishek Vembadi, Anoop Menachery, and Mohammad A Qasaimeh. Cell cytometry: review and perspective on biotechnological advances. *Frontiers in bioengineering and biotechnology*, 7:147, 2019.

[123] Paul Rees, Huw D Summers, Andrew Filby, Anne E Carpenter, and Minh Doan. Imaging flow cytometry. *Nature Reviews Methods Primers*, 2(1):86, 2022.

Summary of scientific publications

Author contributions

Co-authors are abbreviated as follows, in order of appearance:

Elias Kristensson (EK), Yupan Bao (YB), Vassily Kornienko (VK), Zhiyong Wu (ZW), David Frantz (DF), Simon Ek (SE)

Paper I: A light-efficient and versatile multiplexing method for snapshot spectral imaging

This paper presents a novel optical design for spectral multiplexing using FRAME, improving the light-efficiency compared to previous iterations. Making use of this improvement, the setup is used in combination with an intensified camera to capture multispectral images of nano-second plasma discharges. It is also used with a high-capture rate camera, as well as a monochrome CMOS-camera to perform pyrometry imaging on a propane flame.

I designed and built the experimental setup, including the 3D-printed design, and collected all data used for the paper, with occasional experimental assistance from EK, YB and VK. I also performed the analysis and wrote the manuscript with help and input from all authors.

Paper II: Multispectral coded light for time-resolved temperature imaging of aluminum combustion

This paper builds on the pyrometry imaging methodology developed in Paper I by investigating the combustion process of aluminium droplets. The same type of optical multiplexing setup is used in combination with a high capture-rate camera to

create multispectral images of 550 µm diameter aluminium droplets at 50 kHz. The multispectral images are then used to create temperature distributions of the droplet and its surrounding reaction zone.

I designed and built the optical setup. I and ZW collected the data I performed the data analysis with input from ZW. I have written the majority of the manuscript, with input and help from ZW to put the research into context.

Paper III: Simultaneous multiple time scale imaging for kHz–MHz high-speed accelerometry

This paper demonstrates the use of temporal multiplexing to make any camera system capable of burst imaging, and how this can be used to achieve snapshot accelerometry data. By combining a structured pulse train with a high capture-rate camera, FRAME-triplets (three exposures) are captured in the same camera image. By comparing positions in the three exposures, the velocity and acceleration can be calculated from a single camera image, resulting in accelerometry data at 21 kHz. This was applied to spray dynamics.

I assisted VK in collecting the data, and I performed the data analysis, including calculation of projected velocity- and acceleration vectors. I also wrote the part of the manuscript concerning said analysis, and assisted with input for the rest.

Paper IV: Using SILMAS to improve machine learning-assisted quantification of pathology

This paper demonstrates the importance of data quality in the context of machine learning. In a research climate where a lot of focus is placed on improving the capabilities of neural networks, our goal was to highlight their dependence on the quality of the underlying data used for training. In the paper, two datasets are used, captured using two measurement techniques - SILMAS, which reduces the effect of scattering in images using structured light, and ASLM, a non-structured light equivalent. A well-known network architecture is then trained on both datasets separately, and the relative performance is evaluated.

I, along with DF, performed the manual data labeling, and I performed the analysis utilizing neural networks and post-processing of the output. I and DF were jointly responsible for the writing of the manuscript, with input from the other authors.

Paper V: Snapshot multicolor fluorescence imaging using double multiplexing of excitation and emission on a single detector

This paper presents a novel method of multispectral capture using FRAME which combines structured illumination and the passive encoding of light, simultaneously modulating the intensities of exposures in two different ways. This is used to spectrally distinguish between 9 different fluorophores, some of which have near identical emission spectra.

I participated in collecting data and provided input for the writing of the manuscript.

Paper VI: Simultaneous multispectral microscopy using passive FRAME multiplexing with seven spectral channels

This paper demonstrates the combination of passive multispectral FRAME with microscopy. The setup features 7 spectral channels, and is used to image a number of different sample types - microscopic beads, a light-emitting display, and a blood sample. The microscope is used in two different configurations - brightfield using a broadband light source and using laser illumination - and demonstrates the ability to resolve features down to a size of 2 µm.

I designed and built the setup together with EK, and we collected the data with assistance from SE. I have analyzed the data and written the manuscript together with EK.