Structured Light for Ultrafast Videography

Kornienko, Vassily

2024

Document Version:
Publisher's PDF, also known as Version of record

Structured Light for Ultrafast Videography

Structured Light for Ultrafast Videography

by Vassily Kornienko

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Ass. Prof Elias Kristensson
Faculty opponent: Ass. Prof Daniel R. Richardson

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism at Rydberg Lecture Hall (Rydbergssalen) at the Department of Physics on Friday, the 12th of April 2024 at 09:15.

Organization	Document name
LUND UNIVERSITY Department of Physics Box 118 SE–221 00 LUND Sweden	**DOCTORAL DISSERTATION** Date of dissertation 2024-04-12 Sponsoring organization
Author(s) Vassily Kornienko	

Title and subtitle

Structured Light for Ultrafast Videography

Abstract

Monitoring natural events at ultrafast timescales is a prerequisite for our continued understanding of physical, chemical and biological processes. Due to exponential advancements in knowledge and optical technology, the field of ultrafast videography has seen a sudden uptick in activity beginning in the mid-2010s. Today, the field has matured past its infancy and is slowly becoming a standard tool used to solve both fundamental and applied scientific questions. However, there are constant improvements to be made in terms of speed, video lengths and spatial resolution in order to reach full accessibility.

This thesis work presents, discusses and puts the applied method for ultrafast videography, Frequency Recognition Algorithm for Multiple Exposures (FRAME), into a larger context of videography as a whole and ultrafast videography in particular. FRAME is then both stress-tested in terms of illumination, video length and speed and furthermore, applied to novel scientific questions such as plasma diagnostics and light matter interactions. In addition, the underlying principle, structured light, was applied to alleviate the electro-optical problems of passive ultrafast detection seen most prominently in streak cameras.

In short we have shown and pinpointed the whens and wheres that the FRAME technique is most advantageous for ultrafast videography. Firstly, the technique is quite insensitive to illumination source, which, being an active technique, is key for its accessibility. Secondly, since the method of encoding involves the use of spatial frequency modulations, FRAME can be made sensitive to e.g., spectroscopic information alongside its temporal sensitivity, unlocking potential utilisation in fields that are dependent on the detection of other photon degrees of freedom. Finally, new ways of approaching the optical setup (e.g., with diffractive optical elements) and the post-processing (e.g., with a phase sensitive algorithm,), have shown that FRAME is a robust and versatile way of observing natural events with ultrafast temporal resolutions.

With that, I believe that the work presented herein will contribute to the bank of knowledge that will help propel FRAME as an ultrafast videography method to the far reaches of monitoring nature at its limits.

Key words

Structured Light, FRAME, Coherence Lifetime Imaging, Optical Diagnostics, Ultrafast Videography

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language
	English
ISSN and key title ISSN: 1102-8718	ISBN 978-91-8039-985-2 (print) 978-91-8039-984-5 (pdf)
Recipient's notes	Number of pages 245 / Price
	Security classification

Structured Light for Ultrafast Videography

by Vassily Kornienko

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: An artist's impression of a FRAME setup (Credits: Grigory Kornienko).

Cover illustration back: A hand structuring a laser wave at the Physics and Lasershow.

Funding information: The thesis work was financially supported by the European Research Council, Grant Number: 803634 and Vetenskapsrådet Grant Number: 2019-05183

Faculty of Engineering, Department of Physics

ISBN: 978-91-8039-985-2 (print)
ISBN: 978-91-8039-984-5 (pdf)
ISSN: 1102-8718
LRCP: 252

Printed in Sweden by Media-Tryck, Lund University, Lund 2024

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

MADE IN SWEDEN

Table of Contents

Abstract iv
List of publications v
Acknowledgements vii
Popular Summary x
Populärvetenskaplig sammanfattning xii
Résumé Populaire xiv

Structured Light for Ultrafast Videography **1**

1 Introduction **3**

2 Optical Diagnostics **5**
1 The Photon 6
2 Detection, Intensity and Discrimination of Photons 7
2.1 Optical technology as decoders of information 10
2.2 Formalising optical diagnostics 14
3 The Category of Time 17
3.1 Criteria of ultrafast videography 18

3 Current State of Technology **21**
1 Focal Point Arrays 22
1.1 FPA technology from 1970 to 2023 22
1.2 Speed limit of digital sensor technology 23
2 Ultrafast Videography - History and Terminology 24
2.1 Bookkeeping of videography terms 26
3 Overview of Existing Technologies 28
3.1 Passive videography solutions 29
3.2 Active videography solutions 33

4 Fourier Analysis - In Theory and Experiment **41**
1 The Fourier Formalism 42
1.1 Convenient theorems and properties 44
1.2 Sampling of continuous signals 45
1.3 Constructing images with elementary functions 47
2 Imaging 49

2.1 Image propagation through free space . . . 50
2.2 Fourier transforming properties of lenses . . . 53
2.3 Image formation . . . 54
2.4 Spatial resolution . . . 55
3 Natural Images . . . 59
3.1 Intensity histograms and Fourier transforms . . . 60
3.2 Scale invariance . . . 61
3.3 1/f Fourier spectrum dependence of natural images . . . 63

5 Storing Images in the Fourier Domain 65
1 Manipulating Image Frequency Coordinates . . . 66
2 Spatial Resolution vs Specificity . . . 68
2.1 Choosing a filter size . . . 69
3 Discretising Spatial Modulations . . . 72
4 Storage Efficiency of Image Multiplexing . . . 73
4.1 Packing strategy . . . 73
4.2 Spatial efficiency . . . 74
4.3 Dynamic range . . . 78
4.4 Concluding remarks on multiplexed image storage efficiency . . . 79
5 Exotic Packing Strategies . . . 80
5.1 Doubly modulated packing strategy . . . 80
5.2 Encoding light-matter interactions with structured light . . . 81
6 Pushing to the Limits . . . 83

6 Outlook and Final Remarks 87
1 Temporal Resolution . . . 88
2 Outlook . . . 89
3 Concluding Remarks . . . 90

7 Appendix 93

References 110

Summary of Scientific publications 123
Paper I: Beyond MHz image recordings using LEDs and the FRAME concept . 127
Paper II: Long sequence single-exposure videography using spatially modulated illumination . . . 139
Paper III: Simultaneous multiple time scale imaging for kHz-MHz high-speed accelerometry . . . 151
Paper IV: High-speed videography of transparent media using illumination-based multiplexed schlieren . . . 163
Paper V : Improved temporal contrast of streak camera measurements with periodic shadowing . . . 175

Paper VI : The space-charge problem in ultrafast diagnostics: an all-optical solution for streak cameras 181
Paper VII : A versatile, low-cost, snapshot multidimensional imaging approach based on structured light 193
Paper VIII : Snapshot multicolor fluorescence imaging using double multiplexing of excitation and emission on a single detector 209

Abstract

Monitoring natural events at ultrafast timescales is a prerequisite for our continued understanding of physical, chemical and biological processes. Due to exponential advancements in knowledge and optical technology, the field of ultrafast videography has seen a sudden uptick in activity beginning in the mid-2010s. Today, the field has matured past its infancy and is slowly becoming a standard tool used to solve both fundamental and applied scientific questions. However, there are constant improvements to be made in terms of speed, video lengths and spatial resolution in order to reach full accessibility.

This thesis work presents, discusses and puts the applied method for ultrafast videography, Frequency Recognition Algorithm for Multiple Exposures (FRAME), into a larger context of videography as a whole and ultrafast videography in particular. FRAME is then both stress-tested in terms of illumination, video length and speed and furthermore, applied to novel scientific questions such as plasma diagnostics and light matter interactions. In addition, the underlying principle, structured light, was applied to alleviate the electro-optical problems of passive ultrafast detection seen most prominently in streak cameras.

In short we have shown and pinpointed the whens and wheres that the FRAME technique is most advantageous for ultrafast videography. Firstly, the technique is quite insensitive to illumination source, which, being an active technique, is key for its accessibility. Secondly, since the method of encoding involves the use of spatial frequency modulations, FRAME can be made sensitive to e.g., spectroscopic information alongside its temporal sensitivity, unlocking potential utilisation in fields that are dependent on the detection of other photon degrees of freedom. Finally, new ways of approaching the optical setup (e.g., with diffractive optical elements) and the post-processing (e.g., with a phase sensitive algorithm,), have shown that FRAME is a robust and versatile way of observing natural events with ultrafast temporal resolutions.

With that, I believe that the work presented herein will contribute to the bank of knowledge that will help propel FRAME as an ultrafast videography method to the far reaches of monitoring nature at its limits.

List of publications

This framework of this thesis is based on the following publications:

I **Beyond MHz image recordings using LEDs and the FRAME concept**

V. Kornienko, E. Kristensson, A. Ehn, A Fourriere, E. Berrocal
Scientific Reports, 10 (1) pp. 16650, 2020

II **Long sequence single-exposure videography using spatially modulated illumination**

S. Ek, **V. Kornienko**, E. Kristensson
Scientific Reports, 10 (1) pp. 18920, 2020

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) pp. 1712–1722, 2022

IV **High-speed videography of transparent media using illumination-based multiplexed schlieren**

S. Ek, **V. Kornienko**, A. Roth, E. Berrocal, E. Kristensson
Scientific Reports, 12 (1) pp. 19018, 2022

V **Improved temporal contrast of streak camera measurements with periodic shadowing**

Y. Bao, **V. Kornienko**, D. Lange, W. Kiefer, T. Eschrich, M. Jäger, J. Bood, E. Kristensson, A. Ehn
Optic Letters, 46 (22) pp. 5723–5726, 2021

VI **The space-charge problem in ultrafast diagnostics: an all-optical solution for streak cameras**

V. Kornienko, Y. Bao, J. Bood, A. Ehn, E. Kristensson
Ultrafast Science 4, pp. 55, 2024

VII **A versatile, low-cost, snapshot multidimensional imaging approach based on structured light**

K. Dorozynska, **V. Kornienko**, M. Aldén, E. Kristensson
Optics Express, 28 (7) pp. 9572–9586, 2020

VIII **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 (1) pp. 20454, 2021

Acknowledgements

Here we are, that looming deadline is finally upon me. It's a weird feeling to sit here and type out this piece of text as there are nearly more expectations on this section of my thesis than any other. But is a thesis such a big feat that it requires a text of this weight? After the last couple months, the answer is, pardon my French, oh my god yes... Therefore, a proper acknowledgement section doesn't feel so strange anymore since (1) I am in a honeymoon phase and (2) this could be the only time an opportunity arises where I can write out my thank yous in permanent form. I would like you to keep in mind that, being human, I have been super irritated at everybody within this non-exhaustive list at least once, but you all understand that. With that, here we go!

Firstly, Elias, your simultaneous hands on and hands off approach to supervision has been perfect. It is as simple as that. Your door has always been open for fun, learning and discussions and without you this journey would have been mediocre at best, while simultaneously I would not have had the chance to really explore what I really burn for. So for all, Thank You.

To my co-supervisors, Andreas, Edouard and Joakim. Your doors have always been open for listening to crazy ideas, open for discussing and general banter. This has been quintessential for my journey in finding what really makes this job amazing. To all, for all, Thank You.

Now comes the economy class: the people I have worked closest with. Simon, you know I always go to you for those super tricky technical chats and for updates on your crazy projects. I feel like our teamwork has developed into an amazing thing! Yupan, the insane teamwork, nothing is impossible for us in the lab. You want to improve streak cameras, hell yeah, staple water to a tree, gotcha. My times in the lab have been an absolute pleasure to spend with you! Dario is actually on my lap at the moment, and he obviously says hi. Sam and Klara, I learnt so much when we worked together and what amazing journeys you put me through. Sebastian, Jonas and David (as Yupan would say with a sigh, the boys) I don't know if I have become crazier with you guys around or if your presence and constant chit-chatting about literally nothing of substance has kept me sane. Nonetheless, how could I survive without you.

No acknowledgement section would be complete without a shoutout to my office mate! Saeed, even though we were never at the office at the same time, being alone for real has really been empty; a hole that is impossible to fill. Keeping the acknowledgement section complete: my closest friend, Eric (Herrgårn), all I can do is laugh at our great memories that have been and are to be made. Love you babes.

To all the Combustion Physics family (feels kind of dystopian to call colleagues family but that's really the way I have felt). There are way too many of you that have meant so much

to me throughout this time. All that sweet lunch banter, the discussions, the open doors, general kindness, understanding and of course the willingness to listen and help out both personally and work related has really been central to the amazing working environment I have been lucky enough to have spent time in. And for creating such an environment, a special thanks to all the senior and administrative staff, you do such amazing work to promote and keep this down-to-earth culture that is so rare within workplaces and academia in particular.

To Leading Edge, driving 5000 kilometres with you every year, living and gigging from the shittiest places (I could not think of another word to describe these places) to those one has dreamt of will never get old. I also didn't know there were that many types of onions. Anna, Lewis, Lina, Eric, Emma, Elias, John, Saga, Sebastian, Folke, Adam, thank you for amazing times!

To Fysik o Lasershowen. Do I even need to say anything? All that lifting, deathly exhaustion, insanely good banter, dinners and beer! You mean too much for a general thanks, so see the back of the front cover. I will however write a thank you (along with a personal note) to Johan and PO that have given me the opportunity to explore myself, grow into the shoes I stand in today and for providing the jumping board that has allowed me to meet all these great people. Thank You.

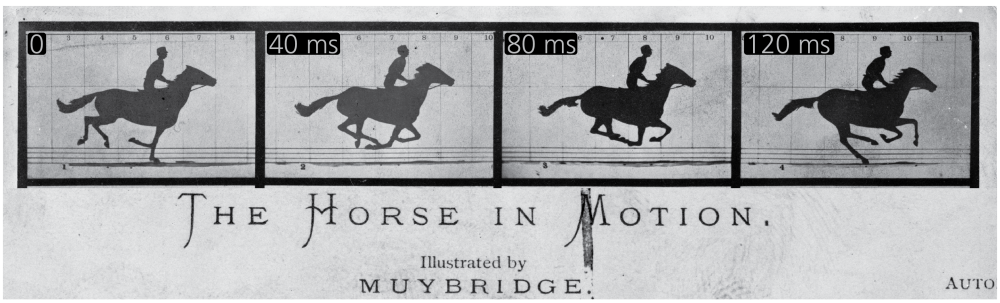
My grandmothers: a pianist from Astrachan and a Geneva seamstress, incredible strong women that, through extremely hard times, drove the doors wide open for their grandchildren. My grandfathers: a radar and antenna physicist that had to build his measuring apparatus out of birch trees and a builder of a still existing bus company that drove through thick snow in the alps during the day and repaired buses during the night, again through incredibly hard times (that still go on) have allowed us grandchildren to really explore whatever pleases us. Merci, Спасибо. Down the generational tree, my parents: being professional figure skaters one would think that they forced me to skate. None of that! You two have been so supportive of anything I do and really allowed me to flourish in my endeavours. Unconditional love, there it is, and I will for once actually say: Thank You Mama and Papa.

Griguette, all those crazy mischievous things we did together growing up, there is not even enough room in my head to remember it all. You will always be my little brother, but the tenacity, drive and dedication you put into your interests really has inspired me. Keep it up and Thank You

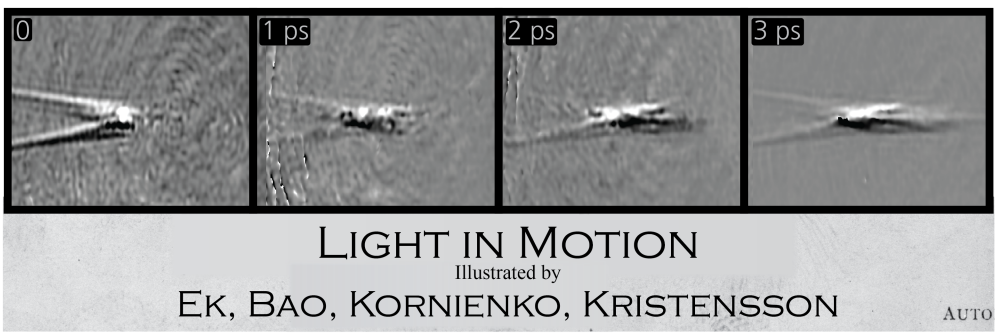
Finally, Lina. How was I able to meet such an amazing person, a person that fights for what's right, a person that jumps off the deep end with me and a person that without fail ALWAYS supports me. Thank You for All.

After this piece of text, I realise that I really do stand on the shoulders of all you giants!

1877: 25
frames per second / bilder per sekund / images par seconde



2024: 1 000 000 000 000
frames per secomd / bilder per sekund / images par seconde



Eng: From the horse to light in motion. What a long way videography has come, eleven orders of magnitude, i.e., 10^{11}, faster!

Swe: Från hästan till ljus i rörelse. Vad långt filmning har kommit, elva storleksordningar snabbare, dvs 10^{11} gånger snabbare!

Fr: Des chevaux à la lumière en mouvement. Quel chemin long la vidéographie a parcouru, onze ordres de grandeur, c'est-à-dire 10^{11} fois plus rapide!

Popular Summary

Humankind's wish to be able to record and film events has probably existed for millennia, however, it wasn't until the late 19th century when this was made possible for the first time. At that point one had for a long time admired the gracious way horses galloped, but a single question was under heated debate: does a horse completely leave the ground when it gallops? It was the industrialist, horse enthusiast and ex-governor of California, Leland Stanford, who decided to solve this debate, not only out of curiosity but legend has it that he had a $25000 bet placed on his theories[1]. So he employed the photographer Eadward Muybridge, and in 1878 they were able to capture a four image movie of the horse, Sallie Gardner, and jockey, Gilbert Domm, finally showing that horses leave the ground as they gallop (see the figure on page ix). The general public had a hard time accepting these findings as they found the images of the horse's movements to be quite ungracious. Furthermore, many even questioned the reason to interfere with God's created beauty of the world. Eadward Muybridge went on to publish an encyclopaedia type book about Animal Locomotion where animal and human subjects[2] were filmed with his tripwire based shuttering contraption.

This technological feat opened up for the possibility to quench humans' innate curiosity of seeing natural events in slow motion, and today, the ability for slow motion capture or *high-speed videography* exists, to a certain degree, within everybody's pockets. However, as is always the case in science, we feel the need to do better and in this case that would mean to film faster. Indeed, what if we could film fast enough for nothing to happen, such that even light, which travels at the fastest speed of which anything can travel, is stopped in its tracks? Surely that would be the end of the line? This is where the work presented within this thesis comes into play, as during my PhD I have worked on developing techniques for attaining video speeds that are capable of "stopping light". I then worked on applying them to various natural events, such as plasmas, where things go at such speeds that one does not have a choice but to film with a "stopping light" type of technique (an example video taken by us in Lund is shown in the lower panel of the figure on page ix).

The title of this thesis, **Structured Light for Ultrafast Videography** can be divided into two parts: structured light and ultrafast videography, where the intertwining of these two concepts forms the basis of the discussions within this thesis.

Ultrafast Videography entails the filming of ultrafast phenomena. And no, I am not

[1]There is no evidence to this legend however I always like to include it because he actually ended up spending about $50000 in total, spending more than he bet. For context $50000 in 1880 is worth $1 500 000 in 2024, i.e., about 15 years' worth of what I cost my supervisor as a PhD student per year.

[2]It is debated if the human subjects were really filmed out of scientific curiosity as men were filmed performing physical tasks while women were filmed e.g., "opening a parasol and turning around" in revealing clothing. Whatever it was, the nudeness did stir up quite a bit of controversy!

joking, the scientific term is indeed ultrafast. This type of event, (such as the rotation of molecules, or the creation of plasma light channels called filaments) occur on timescales faster than 1 picosecond, i.e., 0.000000000001 seconds (that's 11 zeros after the decimal point). Therefore in order to capture such events we need to film at corresponding speeds i.e., above 1 000 000 000 000 frames per second[3]. This is so fast that if one compares the necessary shutter time of the camera to one second, that would be equivalent to the time comparison between a day and the age of the universe!

These shutter speeds are so fast that the speed of electrons (the carriers of information in cameras) within the electronic circuitry of cameras is simply too slow for this approach to be useful. Therefore, one needs to use light instead, in particular ultrashort laser pulses. In this part of the title, **Structured Light**, light pulses are created and manipulated such that they (1) illuminate the event like a stroboscope would, but at an ultrafast speed and (2) they are structured such that they have unique striped patterns imprinted onto them. As will be shown within this thesis, together these two manipulations are enough to be able to open up the realm of ultrafast videography.

Within this thesis work, paid by you the taxpaying reader, I firstly delve into the magical world of optics, lasers, imaging and videography and try to compare all these technologies to what we have closest to us, our eyes and vision (I would encourage even the non-physicist to read this). Then I talk in depth about existing technologies and their inner workings, such as lasers systems and high-speed camera systems. I follow with a discussion in the necessary mathematics before coming out the other end and showing the reader that during my time as a PhD student we have been able to develop, not only niche ultrafast videography systems, but also a high-speed video system that is commercially viable. With that, I hope you had and will have a good read.

[3]For reference, standard movies are filmed and projected on your television or computer screen at 24 or 144 frames per second respectively.

Populärvetenskaplig sammanfattning

Mänskligheten önskan att kunna spela in och filma händelser har förmodligen funnits i årtusenden. Det var dock inte förrän sent 1800-tal som detta möjliggjordes för första gången. Vid den tiden hade man länge beundrat hästens graciösa galopp, men en fråga förblev föremål för intensiv debatt: lämnar en galopperande häst någonsin marken? Det var industrimannen, hästentusiasten och den före detta guvernören i Kalifornien, Leland Stanford, som beslutade sig för att lösa denna debatt, inte bara av nyfikenhet utan också för att han enligt legenden hade han satsat $25000 på sina teorier[1]. Han anlitade fotografen Eadward Muybridge och 1878 lyckades de fånga en fyrbildsfilm av hästen, Sallie Gardner, och jockeyn, Gilbert Domm, och visade en gång för alla att hästar lämnar marken när de galopperar (se figuren på sida ix). Allmänheten hade svårt att acceptera dessa fynd eftersom de ansåg bilderna av hästens rörelser vara tämligen ograciösa. Dessutom ifrågasatte många anledningen till att blanda sig i Guds skapade skönhet i världen. Eadward Muybridge gick vidare till att publicera en encyklopediartad bok om djurs rörelse där djur och människor[2] filmades med hans trådutlösta slutaranordning.

Denna teknologiska prestation öppnade upp möjligheten att stilla människans medfödda nyfikenhet att se naturliga förlopp i slow-motion, och idag finns förmågan att filma i slow-motion eller utföra höghastighetsvideografi, i någon mån, i allas fickor. Men, som alltid inom vetenskapen känner vi behovet av att göra bättre, och i det här fallet skulle det innebära att filma snabbare. Vad sägs om vi kunde filma så snabbt att ingenting händer, så att även ljus, som färdas med den snabbaste hastighet som någonting kan färdas, stoppas i sina spår? Säkerligen skulle gränsen då vara nådd? Det är här arbetet som presenteras inom denna avhandling kommer in i bilden, eftersom jag under min doktorandtjänst har arbetat med att utveckla tekniker för att uppnå filmhastigheter som är kapabla till att "stoppa ljuset". Jag arbetade sedan med att tillämpa dem på olika naturliga händelser, som exempelvis plasma, där saker rör sig med sådana hastigheter att man inte har något annat val än att filma med en teknik av typen "stoppa ljuset" (en exempelvideo taget av oss i Lund visas i figuren på sidan ix).

Titeln på denna avhandling, **Strukturerat Ljus för Ultrasnabb Videografi**, kan delas upp i två delar: strukturerat ljus och ultrasnabb videografi där samspelet mellan dessa två begrepp utgör grunden för diskussionerna inom denna avhandling.

[1]Det finns inga bevis för detta påstående men jag gillar att inkludera det då det slutade med att han spenderade totalt $50000 för att få fram ett bevis, alltså att han spenderade mer än vad han satsade. För att sätta detta i sammanhang är $50000 år 1880 värt $1 500 000 år 2024, alltså ungefär kostnaden av 15 år av min PhD för min handledare.

[2]Det är debatterat vare sig människorna verkligen var filmade i vetenskapligt syfte eller ej, då männen filmades utförandes fysiska aktiviteter medans kvinnorna filmades medan de exempelvis "öppnar ett parasoll och vänder sig om" i lätta kläder. Vad det än anses vara, så lyckades nakenheten dra igång viss kontrovers.

Ultrasnabb Videografi innebär att filma ultrasnabba fenomen. Och nej, jag skojar inte, den vetenskapliga termen är verkligen "ultrasnabb". Denna typ av händelse (som rotationen av molekyler eller skapandet av ljuskanaler av plasma, kallade filament) inträffar på tidskalor snabbare än 1 pikosekund, det vill säga, 0.000000000001 sekund (det är 11 nollor efter decimaltecknet). För att fånga sådana händelser behöver vi därför filma med motsvarande hastigheter, det vill säga över 1 000 000 000 000 bilder per sekund[3]. Det är så snabbt att om man jämför den nödvändiga slutartiden för kameran med en sekund skulle det vara ekvivalent med tidsjämförelsen mellan en dag och universums ålder!

Dessa slutartider är så snabba att elektronernas hastighet (bärarna av information i kameror) inom kamerors elektronik helt enkelt är för långsam för att denna metod ska vara användbar. Därför behöver man istället använda sig av ljus, specifikt ultrakorta laserpulser. I denna del av titeln, **Strukturerat Ljus**, skapas och manipuleras ljuspulser så att de (1) belyser händelsen som ett stroboskop skulle göra, men i ultrasnabb hastighet, och (2) de är strukturerade så att de har unika randmönster inpräntade på sig. Det som kommer visas inom denna avhandling är att dessa två manipulationer tillsammans är tillräckliga för att kunna uppnå ultrasnabb videografi.

Inom ramen för detta avhandlingsarbete, finansierat av dig som skattebetalande läsare, fördjupar jag mig först i den magiska världen av optik, laser, bildbehandling och videografi och försöker jämföra alla dessa teknologier med det vi har närmast oss, våra ögon och synen (jag skulle uppmana även den som inte är fysiker att läsa detta). Sedan går jag djupt in i befintliga teknologier och deras innersta mekanismer, såsom lasersystem och höghastighetskameror. Jag följer upp med en diskussion om den nödvändiga matematiken innan jag kommer ut på andra sidan och visar läsaren att under min tid som doktorand har vi kunnat utveckla inte bara specialiserade system för ultrasnabb videografi, utan också ett höghastighetsvideosystem som är kommersiellt gångbart. Med det sagt hoppas jag att du har haft och kommer att ha en trevlig läsning.

[3]För att jämföra, standard filmer är filmade och projicerade på TV-apparater och datorskärmar i 24 respektive 144 bilder per sekund.

Résumé Populaire

Le souhait de l'humanité de pouvoir enregistrer et filmer des événements existe probablement depuis des millénaires. Cependant, ce n'est qu'à la fin du XIXe siècle que cela a été rendu possible pour la première fois. À ce moment-là, on admirait depuis longtemps la manière gracieuse dont les chevaux galopaient, mais une seule question faisait l'objet d'un débat passionné: est-ce qu'un cheval quitte complètement le sol lorsqu'il galope? C'est l'industriel, amateur de chevaux et ancien gouverneur de Californie, Leland Stanford, qui a décidé de résoudre ce débat, non seulement par curiosité, mais la légende raconte qu'il avait misé 25000 dollars sur ses théories[1]. Ainsi, il engagea le photographe Eadward Muybridge, et en 1878, ils parvinrent à capturer un film composé de quatre images du cheval, Sallie Gardner, et du jockey, Gilbert Domm, montrant enfin que les chevaux quittent le sol lorsqu'ils galopent (voir le figure sur la page ix). Le grand public eut du mal à accepter ces conclusions, car il trouvait les images des mouvements du cheval assez disgracieuses. De plus, beaucoup remirent même en question la raison d'interférer avec la beauté crée par Dieu dans le monde. Eadward Muybridge publia par la suite un livre de type encyclopédique sur la locomotion animale, où des sujets animaux et humains[2] ont été filmés avec une caméra, dont l'ouverture et la fermeture sont contrôlées par des fils de déclenchement.

Cet exploit technologique a ouvert la voie à la satisfaction de la curiosité innée des êtres humains à observer les événements naturels au ralenti. Aujourd'hui, la capacité de capturer des vidéos au ralenti ou de réaliser de la *vidéographie haute vitesse* est disponible, dans une certaine mesure, à portée de main pour chacun, offrant ainsi une expérience visuelle captivante. Toutefois, comme c'est souvent le cas en science, nous aspirons toujours à aller plus loin. Dans cette perspective, cela impliquerait de capturer des images à une vitesse encore plus élevée. Imaginez un instant que nous pouvons capturer des images à une vitesse si vertigineuse que tout semble figé: même la lumière qui se déplaçant à la vitesse maximale. Ce serait véritablement la fin de la course, n'est-ce pas? C'est là que le travail présenté dans cette thèse entre en jeu, car au cours de mon doctorat, j'ai travaillé sur le développement de techniques permettant d'atteindre des vitesses de vidéo capables de "stopper la lumière". J'ai ensuite travaillé à les appliquer à divers événements naturels, tels que les plasmas, où les choses se déplacent à des vitesses telles qu'on n'a pas d'autre choix que de filmer avec une technique du type "stopper la lumière" (une vidéo exemple réalisée par nous à Lund est montrée dans la figure sur le page ix).

[1]Il n'y a aucune preuve de cette légende, cependant, j'aime toujours l'inclure car il a fini par dépenser environ 50000 dollars au total, soit plus que ce qu'il avait misé. Pour mettre les choses en perspective, $50000 en 1880 valent $1 500 000 en 2024, soit environ 15 années de ce que je coûte à mon superviseur en tant qu'étudiant en doctorat.

[2]Il est débattu si les sujets humains ont réellement été filmés par curiosité scientifique, car les hommes ont été filmés en train d'effectuer des tâches physiques tandis que les femmes ont été filmées, par exemple, öuvrant un parasol et se tournant"en vêtements révélateurs. Quoi qu'il en soit, la nudité a suscité pas mal de controverses!

Le titre de cette thèse, **Lumière Structurée pour la Vidéographie Ultra-rapide**, se décompose en deux parties : la lumière structurée et la vidéographie ultra-rapide. L'entrelacement de ces deux concepts constitue la pierre angulaire des discussions développées dans ce document de recherche.

La vidéographie ultra-rapide implique la capture de phénomènes ultra-rapides. Et non, je ne plaisante pas, le terme scientifique est effectivement ultra-rapide. Ce type d'événement (comme la rotation des molécules ou la création de canaux lumineux au plasma appelés filaments) se produit sur des échelles de temps plus rapides qu'une picoseconde, c'est-à-dire 0.000000000001 secondes (c'est 11 zéros après la virgule). Par conséquent, pour capturer de tels événements, nous devons filmer à des vitesses correspondantes au-dessus de 1 000 000 000 000 images par seconde[3]. C'est tellement rapide que si l'on compare le temps d'obturation nécessaire de l'appareil photo à une seconde, cela équivaudrait à comparer une journée à l'âge de l'univers!

Ces vitesses d'obturation sont tellement rapides que la vitesse des électrons (les porteurs d'informations dans les caméras) à l'intérieur des circuits électroniques des caméras est tout simplement trop lente pour que cette approche soit utile. Par conséquent, il est nécessaire d'utiliser la lumière à la place, en particulier des impulsions laser ultracourtes. Dans cette partie du titre, **Lumière Structurée**, des impulsions lumineuses sont générées et contrôlées pour (1) éclairer l'événement de manière similaire à un stroboscope, mais à une vitesse ultrarapide, et (2) elles sont structurées de manière à présenter des motifs rayés uniques imprimés sur elles. Comme le montrera cette thèse, ces deux manipulations réunies sont suffisantes pour ouvrir le domaine de la vidéographie ultra-rapide.

Au sein de ce travail de thèse, financé par vous, lecteur taxé, je plonge d'abord dans le monde magique de l'optique, des lasers, de l'imagerie et de la vidéographie, cherchant à comparer toutes ces technologies à ce que nous avons de plus proche, nos yeux et notre vision (j'encouragerais même les non-physiciens à lire cela). Ensuite, je discute en profondeur des technologies existantes et de leur fonctionnement interne, telles que les systèmes laser et les systèmes de caméras haute vitesse. Je poursuis avec une discussion sur les mathématiques nécessaires avant d'aboutir et de montrer au lecteur qu'au cours de mon doctorat, nous avons été capables de développer non seulement des systèmes de vidéographie ultra-rapide de niche, mais aussi un système vidéo haute vitesse commercialement viable. Avec cela, j'espère que vous avez eu et aurez une bonne lecture.

[3]Pour référence, les films standard sont filmés et projetés sur votre télévision ou écran d'ordinateur à 24 ou 144 images par seconde respectivement.

Structured Light for Ultrafast Videography

C'est le métier qui rentre!
– Remo Borini

Chapter 1

Introduction

Within classical scientific philosophy there exists two primary avenues that lead to the emergence of groundbreaking discoveries. Either one accidentally stumbles upon something new and subsequently devises an explanation for it, or alternatively, one looks at existing theory, and makes predictions such that a measurement can be performed to corroborate said theory. These can be likened to two types of "Eureka" moments. This type of scientific discourse is what has propelled humanity through discoveries and paradigm shifts enabling us to reach our current understanding of the fundamentals of nature. This way of performing science constitutes what I would call Nobel Prize-type research. In contrast, the large majority of modern research and scientific work leverages Nobel Prize-calibre discoveries as *tools* for the continued answering of, most usually, quite niche scientific questions. Fortunately, due to the sheer bulk of these niche questions[1] humanity's bank of knowledge and technology grows exponentially. Moreover, the expansive reservoir of knowledge, full of questions that are capable of sparking anybody's curiosity, has facilitated the evolution of science and research into a more inclusive vocation, propelling even further exponential growth.

The work presented within this thesis, **Structured Light for Ultrafast Videography**, relies on said exponential development in knowledge and technology to even be feasible. Indeed, the invention of the maser, the chirped pulse amplifier and attosecond pulses, all Nobel Prize discoveries [1–3], have set the stage for ultrafast illumination with the ability to produce flashes of light that are fast enough to monitor even electron dynamics. The invention of the charged coupled device (the chip embedded in a camera with the purpose of recording images), also a Nobel Prize [4], made two dimensional digital recording of light fields possible, not only enabling image acquisition, but also the possibility to di-

[1] A quick google puts the number at about 5 million articles, reviews, conference proceeding etc. that were published in 2022.

gitally process them. These two classes of discoveries serve as prerequisites for the advent and maturation of a new category of measurement techniques that can be leveraged for monitoring nature at extreme spatiotemporal resolutions.

Nature has made the real time filming of the propagation of light and its interactions with matter quite a daunting task. Indeed, the speed at which such events unfold demands more than what technology itself has to offer. Instead, subtle ways of outsmarting both nature and technology are necessary and the discipline where such work is carried out is the field of ultrafast videography. Apart from research conducted by Nils Abrahamsson in the late 70s [5], the field has seen a sudden uptick since the early 2010s, revealing that ultrafast videography has moved on from its infancy. Indeed, in the mid-2010s each advancement merited a Nature article, while now, a mere decade later, the time is ripe both technologically and knowledge wise for ultrafast videography to transition into a standard tool for addressing scientific questions within the field of ultrafast science.

This thesis is divided into five chapters wherein I try to describe the pieces of the puzzle that are finally mature enough to render this work possible. We will go from a general description of optical diagnostics, through the current state of technology, via the mathematics and practicalities of Fourier analysis and finally the nitty gritty of structured light and how it can be manipulated with respect to imaging. With that, I hope that this thesis work can put a dent in the field of ultrafast videography, a field that can potentially uncover questions we may not have been aware we could ask.

Chapter 2

Optical Diagnostics

Have you tried throwing a laser at your problem?

The field of optical diagnostics has turned out to become a central part of physics as it is tailor-made (by definition) for extracting information from the photon, the fundamental force carrier of electromagnetism. With the ability to answer both fundamental and applied questions ranging from the photoelectric effect to detecting the remains of ancient civilisations in the Amazon rainforest, the field of optical diagnostics stretches across the reaches of human curiosity. But how can this field achieve these feats? What makes photons so special and how can we extract information from them? In this chapter we will start by introducing the photon then try to formalise how the field of optical diagnostics approaches, detects and categorises the information embedded within. Here we go!

1 The Photon

Light, and its corresponding fundamental particle, the photon, has incredible information carrying abilities. Simply look at the internet, the backbone of our connected society, where photons are exploited to carry 150 000 GBytes per second through optical cables that stretch around continents and across oceans [6]. Photons can also carry information through free space, bouncing off satellites to reach even the most remote areas of the planet and indeed our solar system. However, it is not often pondered upon how broad these information carrying abilities extend and how essential they have been for human cultural and economic development throughout the millennia. Neolithic cave paintings have been able to conjure emotions throughout generations of humans, where photons of light are the intermediate conveyor of information between the paints and our eyes. Light that is sent to us from distant stars and reflected off planets allowed the likes of Tycho Brahe, Copernicus and Galileo to understand our place in the universe. Different colours of light and the way it interacts with cells is essential for decoding the inner workings of e.g., cancer or Alzheimer's disease. All this is possible due to the information carrying abilities of photons and the novel ways in which natural and human inventions have been able to extract said information.

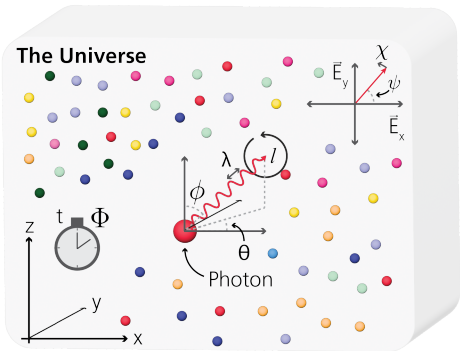
Let us begin with the most obvious type of photon information, one that most of us humans encounter every day, namely colour. Indeed, we can easily tell the difference between a ripe red strawberry compared to its sour white-greenish hue. Evolution has chosen for this capability for our survival - the ability to distinguish between photons of light that carry unique information in terms of colour. Light can also come from a certain position in space, i.e., a photon of light originates from an x, y, z position in Cartesian spatial coordinates. Even though less intuitive, this is in fact a piece of information that the photon carries with it as it travels away from said position. Evolution has given us the ability to distinguish between this type of information by choosing for the eye, or more specifically the lens in conjunction with the retina. A single eye however can only capture x, y information, hence, evolution gave us two of them. That way the brain can stitch together two individual images and extract the z-information (depth) from incoming photons as well[1].

It turns out that light can carry information in many more dimensions than we humans are capable of seeing. The photon as a fundamental carrier of information can in fact carry eleven different types of information [7][2]: wavelength (colour), x, y, z (position in space), polarization and ellipticity angles (the direction of the electric field), polar and azimuthal

[1] Try to touch your two index fingers together with one eye closed as compared to both eyes open. You will notice the change in difficulty. If you're feeling real mischievous try closing both and use that as a reference.

[2] Note that for this thesis we omit any type of quantum information that a photon can carry, we will only be working with classical information and furthermore I have added dimensions to the ones already present in [7].

angles (two photons may come from the same x, y, z position but are emitted at different angles), time of emission, orbital angular momentum and the electric field phase (specifically for coherent sources). Figure 2.1 illustrates these along with a table which includes short descriptions. Even though we, as humans, cannot directly detect all these dimensions, we have in fact been able to make use of them quite efficiently. Going back to the internet, in order to transfer digital information we have designed a light based infrastructure that uses a combination of polarization, angle, colour, time of emission and x, y position in order to maximize the speed and efficiency of data transfer [8–10].



Type	Short Description
x	The x Cartesian coordinate (spatial position)
y	The y Cartesian coordinate (spatial position)
z	The z Cartesian coordinate (spatial depth)
t	Time (when the photon emanated)
θ	Polar angle (with reference to Cartesian coordinates)
ϕ	Azimuthal angle (with reference to Cart. coordinates)
λ	Wavelength (colour of photon)
ψ	Polarisation angle (with reference to direction of travel)
χ	Polarisation ellipticity (ratio between rotation axes)
l	Photon orbital angular momentum
Φ	Electric field phase (related to time)

Figure 2.1: Photons as carriers of information.
A schematic along with the corresponding descriptions of the eleven degrees of information a photon can carry.

2 Detection, Intensity and Discrimination of Photons

Before digging deeper into the different types of information, let us address the issue of detection of photons, how this relates to what we would call brightness, or *intensity*, and the practicalities involved in the discrimination between the degrees of information that they carry. Indeed, without the ability to detect photons or their effect on the physics around us they could just as well not exist[3]. Even though humans can detect single photons at a rate above chance [11], in general we rely on detecting a huge number of photons at a time. On a sunny day at the summer solstice about 4×10^{16} photons enter our eyes every second[4], where each one of these photons is unique with respect to the aforementioned eleven degrees of freedom (disregarding quantum phenomena such as e.g., downconversion). The job of the eye, or any optical detection system in general, is to sort all these photons into categories such that useful information can be extracted. Fig. 2.2a illustrates a breakdown of how optics and optical systems in general perform this task. Photons of interest are captured then funnelled into a discriminating step, where they are categorised into the buckets 1

[3]After all, if a tree falls in a forest and nobody is there to hear it… well you know the drill!

[4]Solar irradiance at the top of the atmosphere is 1360 J/(sm^2) with a central wavelength of 555 nm. A pupil of diameter 4 mm is assumed.

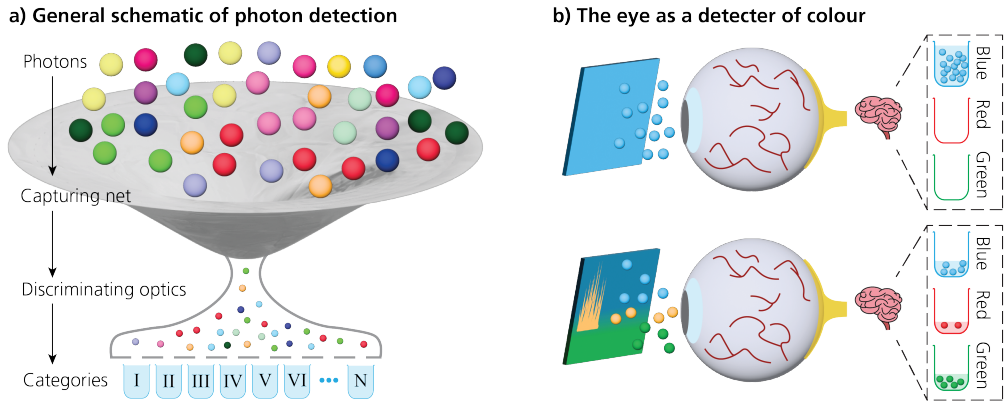


Figure 2.2: Optical detection - the act of sorting photon information.
a) Photons of different properties are captured by an optical system (capturing net) and funnelled into a discriminating step (discriminating optics). Here they are categorised into the categories 1:N for which the system is designed. **b)** Looking at a bright (high intensity) blue sky on a summer's day the eye and brain will categorise photons as blue (upper panel). On a winter afternoon the number of blue photons is significantly lower (lower panel) hence less photons require categorisation. Looking over a scene of a green forest and a campfire, our optical system will further categorise photons into green and orange (red and green).

through N that the element or system is designed for. Note that this breakdown of multiple steps can all be included in a single optical element, an example being the lens. In this case its size and shape captures the photons and its material and curvature discriminates photons depending on their spatial information. We will look into this specific case in significantly more detail throughout this thesis.

Detection

Let us go back to the example of colour detection by humans. The eye has three colour sensors in the form of cone cells that are sensitive to three different colours, red, blue, and green. Out of all these $\sim 4 \times 10^{16}$ photons that enter the eye on a sunny day, the ones emanating from the blue sky will only activate the blue cones and will hence be sorted into the corresponding blue category, to be interpreted by the brain as, you guessed it, blue. This spelled out chain of events exemplifies the way our optical detection system categorises blue photons into a hypothetical blue bucket (Fig. 2.2b, upper panel). If we were to look down from the blue sky, over a forest, then the photons that emanate from the green trees are categorised into the green bucket, while the sky photons will still fall into the blue bucket, exemplifying the simultaneous detection of multiple colours. Now if the there is a campfire with orange flames in the scene, both the red and the green categories would fill up with photons[5], and the ratio between the filling rates will be interpreted by our brain as orange

[5]Note that a single photon has the ability of exciting a red and green cone but can only excite one of them per event.

(Fig. 2.2b, lower panel). This thought experiment can be performed for all the colours most humans can see. To put this in our newly developed terminology: seeing colours amounts to our optical detection system, the eye, categorising photons depending on the wavelength information they carry.

Intensity

What about the difference between the blue sky of a bright summer day as compared to the darker blue sky of a late afternoon in winter, indeed they both have the same colour (let us assume this) but one is obviously different from the other. We now naturally turn to a definition of brightness or as it will be referred to in this thesis: *intensity*. The intensity of light detected by an optical system can be defined as the total number of photons that are categorised into its buckets[6]. In the case of said blue skies at different times of year, the number of photons that end up in the blue bucket during the summer day (Fig. 2.2b upper panel) is much larger than the number of photons that end up in the blue bucket during the winter afternoon (Fig. 2.2b lower panel) - the intensity of light categorised as blue will be higher on the summer day than on the winter afternoon.

Discrimination

If one has a set of unique photons with respect to the information they carry in the eleven dimensions, it should be clear that by choosing a certain type of photon within the set, say that emanates from a certain position in space, the rest of the photons must be discarded. In other words, in order to discriminate for a certain type of information, we must discard all other photons that do not carry said information of interest. Hence, the more constrained our buckets are, the more intensity we must sacrifice. On the extreme end, if our bucket is intended for the information that is carried by a single photon within the whole set, then the optical detection system as a whole must be constrictive enough to discard all photons except the one of interest while simultaneously being sensitive enough to handle an intensity of one photon.

Putting it all together

This act of discarding photons in the hunt for the subset that holds the information of interest, is the only way of extracting information from light. It is the job of the experts of a given field to identify which photons are of interest and design an optical system that can

[6]The buckets can span an arbitrary number of degrees of freedom. They can e.g., accept all photons or be so specific that they only accept photons of a given wavelength to the third decimal.

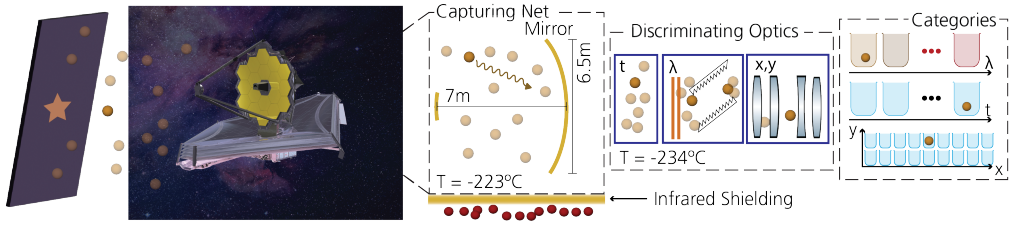


Figure 2.3: The James Webb Space Telescope - an extreme of optical detection and discrimination.
Photons originating from distant stars are captured by the massive gold-plated mirrors. They are then led through timing electronics, spectral filters, spectrometers and lens systems, which, one-by-one discard photons that do not contain the information of interest. The remaining photons, in this case one, are then categorised into their intrinsic categories. Extreme cooling and shielding is necessary to keep unwanted infrared photons out of the system such that the few signal photons of interest can be detected and categorised. Image of the JWST: courtesy of NASA and STScI.

accommodate for this need. For example, say an astronomer needs to know the temperature of a star lying in the far reaches of the universe. An optical system that can detect the colours of photons emanating from the position in space and time where the star is located is then necessary. This might seem like a daunting task to the reader, and it is, since stars that we can see are themselves quite dim and small, let alone stars that are on the edges of our visible universe. Therefore, the intensity of photons that are of interest to the astronomer will be extremely low. It turns out that the intensities are so low that huge, shielded, supercooled optical systems set in space, such as the latest James Webb Space Telescope are necessary in order to fulfil this task of identifying and categorising the few photons that are of interest (Fig. 2.3). A massive capturing net consisting of a honeycomb structure of gold plated mirrors lead photons through timing electronics, spectral filters, spectrometers and lens systems such that unwanted photons are thrown away while photons of interest are categorised in the bins of interest! For a more relevant example with respect to this thesis, if a plasma is formed in a laboratory setting and we are interested in seeing how the molecular nitrogen along the edge of the shockwave acts 1 µs after breakdown, an optical setup that can discriminate in time, x, y, z and wavelength is necessary (see Appendix F for details). As the reader will see, at the present moment in history, there exists multiple layers of discriminating optical elements and tricks such that the buckets necessary for this specific task can fill up to a point where this question is answerable.

2.1 Optical technology as decoders of information

As we have seen, optical elements, optical diagnostic techniques and all the technologies included in this branch of science are aimed at capturing photons' information in novel ways. This branch includes technologies, whether direct or indirect, linear or non-linear, that sort incoming photons into buckets that are of interest to the scientist. In the following section, we will look into some simple examples, formalise them and then apply these concepts to the context of the work presented herein.

Linear optical elements for decoding photon information

The first optical technology I would like to address is probably the oldest type, the pinhole[7]. This optical element essentially consists of a small hole in a piece of paper (or a wall) where light can pass through. By tightly constraining the incoming light so that it can only enter the hole at a specific angle given its spatial point of emission, the pinhole allows for an image to be created on a screen behind it (Fig. 2.4a). After some post-processing in the form of simple geometry, the angle at which the photon emanated can be calculated. Hence, the pinhole is able to capture, x, y, θ and ϕ information from a set of incoming photons. This fact has been used in practice for centuries in technologies such as the camera obscura.

As a continuation of the pinhole, and one of the most important optical elements to exist, is the lens. The first lens, used in order to magnify objects[8], was constructed with water inside a tube up to 4000 years ago in China. The oldest existing glass lens is the Nimrud lens created by craftspeople in the Assyrian empire during the 8th century BC, while one of the oldest and most intact collections was found in a Viking grave in Visby, Gotland [12]. These archaeological findings indicate that glass lens technology has existed for a long time. The combination of its curvature, along with the material's ability to slow down light, allows the lens to, like the pinhole, extract x, y information from photons emanating from a scene - an act that is termed *imaging* (Fig. 2.4b). Since the lens does not extract polar angle information, its strong benefit is that it can provide x, y information in the form of an image at a much higher intensity than the pinhole, in other words its photon *throughput* is larger. Indeed, by not creating unique buckets for polar angles, i.e., not discriminating for this information, the x, y buckets will be filled with more photons than for the case of the pinhole. The reader might be alarmed as to why we would want to use the lens instead of the pinhole if the acquired information is vaguer in comparison? It turns out that in scientific fields where imaging is of interest (from biomedicine to energy science) the x, y, information is the information the scientists are after and the θ, ϕ information is most often redundant. Hence the lens is the preferred optical element for imaging scenes.

What happens when we start combining optical elements such as the pinhole and the lens? By placing a pinhole in front of the lens, all that happens is a decrease in intensity[9] and a deterioration of the image (we will get back to this later). However, if the pinhole is placed at the focus of the lens, the photon polar angles will be constrained in a way that one can extract z, or depth information. Notice how this combination has allowed us to create a new bucket within which we can sort photons, a bucket that is unique compared to the

[7] Of course this is impossible to know, however it would not surprise me if some Neolithic human noticed the perfect silhouette of a bird passing over a thick rainforest canopy on the ground in front of them, where the space between the leaves functioned as a pinhole.

[8] Or burn ants, a timeless child activity.

[9] The system will capture less light from the scene since it subtends a smaller solid angle, i.e., its capturing net is smaller.

possible buckets that the pinhole and the lens themselves can supply. This does not need to end here, now we place a piece of green stained glass (in technical terms, a colour filter) after the pinhole lens setup and tada, we now have wavelength information. Hence, with these three optical elements we have constrained four different photon dimensions, x, y, z, and λ (Fig. 2.4c).

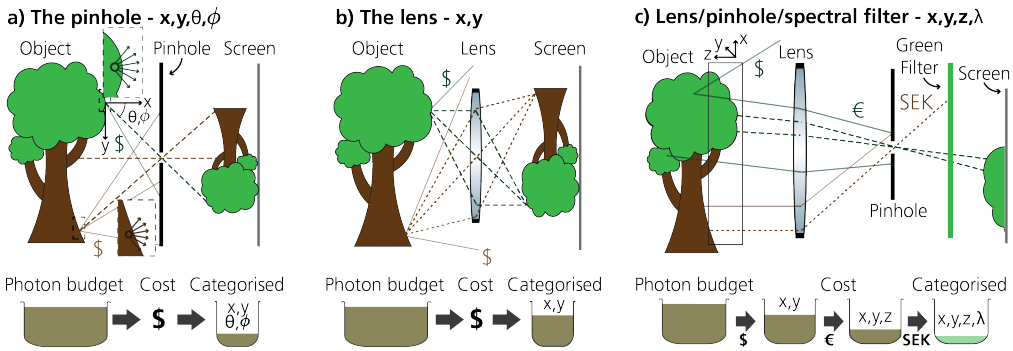


Figure 2.4: Linear optical elements and their categorising abilities.
a) The pinhole confines photons with specific angles in order to extract x, y, θ, ϕ information. This process is quite costly as the photons are constrained tightly by the small hole. **b)** The lens is also able to extract x, y information albeit at a much lower cost than the pinhole (in terms of discarding photons). **c)** The combination of multiple optical elements in novel ways can be utilised to discriminate multiple degrees of freedom, albeit at a price. In this case the lens pinhole and green filter are able to discriminate for x, y, z, λ. This is done by discarding photons that do not fall within the categories of interest.

Non-linear optical elements for decoding photon information

The above optical elements and their combinations exemplify linear optical technologies - they are technologies that map photons directly, or linearly, to their corresponding buckets. An example of a non-linear optical technology is the microchannel plate (MCP). This technology was first developed for military use and can be recognised in pop culture as the "see in the dark" goggles that produce green images for spies and soldiers[10]. The MCP consists of a slab of resistive material dotted with an array of tubes over which a high voltage is applied. Simply put, when a photon enters a given tube, it will excite an electron from the material which will then be accelerated by the voltage difference and, via collisions with the tube walls, release more and more electrons in a so called avalanche amplification (Fig. 2.5a). This has the effect of amplifying the intensity of a single incoming photon to up to millions of electrons, that can ultimately be read by a phosphor screen. The information that an MCP extracts from photons is spatial x, y information, indeed the individual tubes provide an array within which photons are confined depending on their spatial posi-

[10]The MCP is only one of the components included in these goggles, as a phosphor screen is then necessary to produce the green image.

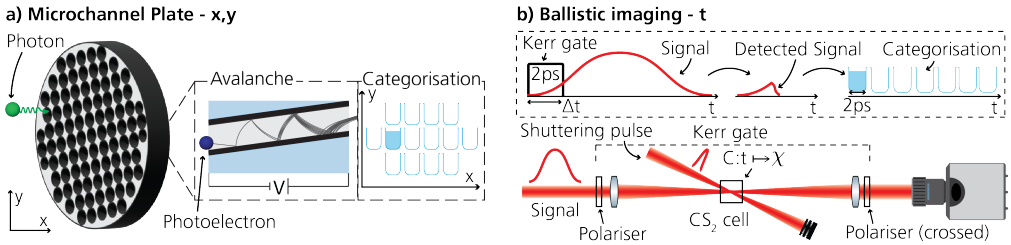


Figure 2.5: Non-linear optical elements and their categorising abilities.
a) The microchannel plate (MCP) categorises photons depending on their x, y information. It is specifically designed for low-light usages, amplifying the number of photons in x, y bins by leveraging other degrees of freedom such as $\lambda, \theta, \phi, \psi, \chi$. This process is extremely non-linear. **b)** Ballistic imaging uses a Kerr medium (CS_2) in order to couple photons' temporal information with their polarisation ellipticity. This non-linear detection system can hence categorise photons that pass through the Kerr medium with picosecond sensitivity.

tion at the entrance of the MCP[11]. The non-linearity of this optical technology stems from three reasons. Firstly, the intensity within the resulting x, y buckets does not only depend on the x, y information embedded within the incident photons, but also on wavelength, polar angles and polarisation. Secondly, the dependence of the intensity within these x, y buckets on the aforementioned degrees of freedom is highly non-linear and to some extent stochastic. However, the sheer number of electrons created in the avalanche amplification averages this randomness out. Finally, the initial electron kinetic energy depends on the work function of the material which is described by a rectified linear unit function, which is itself highly non-linear.

Another example of a non-linear optical technique is ballistic imaging (BI). BI is a technique that exploits the so-called non-linear Kerr effect within a carbon sulphide cell in order to create an ultrashort shuttering mechanism (Fig. 2.5b). It exists in order to solve the issue of categorising photons emanating from a scene within time windows as short as 2 ps [13]. Briefly, a high power femtosecond pulse (the shuttering pulse) enters the cell and aligns the carbon sulphide molecules such that when signal photons pass through, their polarisation will rotate. By placing crossed polarisers on either side of the cell, only light that has interacted with the molecules as they are aligned will be able to pass through the gate. The trick is that this state of molecular alignment has a lifetime of about 2 ps, hence functioning as an ultrafast shutter. This is an example of a technique that performs an operation on the signal photons and converts or *couples* the temporal degree of photon information into one that is easier to measure, polarisation. This consists of a new development within our introduced terminology, one that needs to be formalised in order to compare and understand different optical diagnostic techniques. Firstly, however we need to discuss one last method for decoding photon information, namely post-processing.

[11]Note that this spatial information is only the x, y information that the photon has at the entrance of the MCP since photons originating from different positions at a scene can land on identical tubes given that they have a certain angle. In order to extract the x, y information from the *scene*, a lens in front of the MCP is necessary.

Post-processing for decoding photon information

There is actually one more way of discriminating between photon information. These operations are performed in the so-called *post processing* step, that is, once the photons have been optically captured, detected and categorised into bins and one is left with a detected signal. This process can range from something as easy as deleting the first part of a time-dependent one dimensional signal, i.e., discriminating for photons that arrived at the detector after a certain time, to something as difficult as extracting the edges from an image of a scene with complicated edge detection algorithms[12]. No matter the purpose, this process will still entail the discarding of detected photons in order to isolate the desired quantity. It turns out that no matter what the application, some post-processing is close to always necessary.

2.2 Formalising optical diagnostics

In order to formalise and categorise the various optical elements and techniques we have discussed so far, we turn to state-space formalism. A vector space **U** can be constructed around the aforementioned eleven degrees of freedom, dubbed the photon space (Fig. 2.6a). A measurement on this space, with the aim of extracting information from the photons within, can be seen as a mapping, **M**, from photon space **U** to a real, measured space **V**:

$$\mathbf{M}:\mathbf{U}\mapsto\mathbf{V}. \tag{2.1}$$

An example of an element that performs such a mapping would be a screen, where x, y information from photon space, is mapped onto x, y positions in real, measured space. In this case, the mapping performed by the screen, $\mathbf{M}_{\text{screen}}$, is a linear mapping as there is a direct and unique correspondence between values in **U** and **V**. **M** can also be non-linear, as is the case with the MCP. For such an operation, $\mathbf{M}_{\text{MCP}}$ maps $x, y, \lambda, \theta, \phi, \psi, \chi$ to x, y. These are two examples of *direct* optical diagnostic methods, i.e., the photon information of interest is directly measured by the optical element (Fig. 2.6b).

Another type of mapping that can be performed is a mapping from photon space to photon space dubbed **C**:

$$\mathbf{C}:\mathbf{U}\mapsto\mathbf{U}. \tag{2.2}$$

C can be seen as a transfer of information or a coupling operation within or between photon degrees of freedom. For example, in the case of the lens in an imaging configuration it will

[12] Even though this can nowadays be performed with metasurfaces!

map photon x, y positions in the object plane to x, y positions in the image plane, with a certain magnification, thus functioning as an information "relay". If the lens is instead set in a collimation configuration, i.e., the photons of interest emanate from its focal point, the lens will map photon's angular information θ, ϕ to a spatial distribution x, y, in effect coupling these two photon dimensions via the operation $\mathbf{C}_{\text{lens}} : \theta, \phi \mapsto x, y$. By knowing $\mathbf{C}_{\text{lens}}$ and measuring the x, y information with e.g., a screen ($\mathbf{M}_{\text{screen}}$), one can backtrack and figure out the angular information that the photons carried to begin with. The real strength of this coupling operation is that one does not need to resort to directly measuring difficult quantities, but one can instead operate on photons and convert their information into photon dimensions more conveniently measured (such as x, y information). These are examples of *indirect* optical diagnostic elements and techniques (Fig. 2.6b).

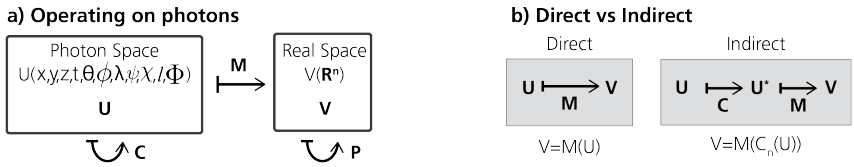


Figure 2.6: Formalising optical diagnostics.
a) The mathematical formulation of optical diagnostics in terms of state spaces where the detection of photon information entails a mapping, **M**, from photon space to real, measured space. Coupling operations, **C**, and post-processing operations, **P**, can be performed within the individual spaces for further versatility. **b)** Direct optical diagnostic techniques are characterised by a direct mapping of the photon dimension to real space while indirect optical diagnostic techniques entail one, or multiple, coupling operations within photon space, $\mathbf{C}_n$, before detection. Most often than not, post processing, **P**, is applied to detected signals, hence it is omitted in these descriptions.

Further examples - measuring time

Let us now turn to two linear techniques that measure the temporal information of photons in **U**, namely a mechanical shutter, an example of a *linear direct* technique and Sequentially Timed All-Optical Mapping Photography (STAMP), a *linear indirect* technique.

The mechanical shutter, when placed within an optical setup, lets light through when it is open and stops light when it is closed, thus performing the direct measurement of the temporal information embedded within photons (Fig. 2.7a). This can be described by a measurement on the photon dimension of time directly to real, measured space, i.e., $\mathbf{M} : t \mapsto t$.

STAMP (in particular for this case, spectrally filtered STAMP, SF-STAMP) is an ultrafast videography technique that uses the angular dependence of a band-pass filter in conjunction with a chirped broadband pulse in order to extract videos of ultrafast events (Fig. 2.7b). A broadband pulse is incident on a glass rod, creating a controlled chirp such that red photons arrive first and blue photons arrive last in time. This rod acts as the first coupling operation within the setup, namely $\mathbf{C}_1 : t \mapsto \lambda$. This operation couples the photon's

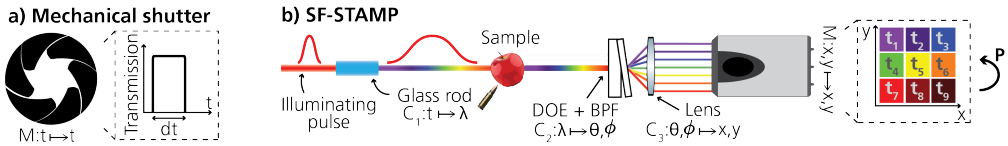


Figure 2.7: Linear techniques in terms of mapping photons' temporal information.
a) The camera shutter, where photon time is mapped directly to time in detection space by a temporally dependent "opening" and "closing" of the transmission of photons to the detector. **b)** STAMP exemplifies an indirect technique where time is initially coupled to wavelength then, after further coupling operations, is converted to a spatial distribution on a camera. Hence, time is encoded into the spatial dimension of measurement space to then be divided into a video via operation **P**.

wavelength information to its temporal information such that $\lambda = \lambda(t)$. This pulse is then incident on an event such that the dynamics are imprinted onto its profile. A diffractive optical element (DOE) and tilted spectral band-pass filter (BPF) together perform $\mathbf{C}_2 : \lambda \mapsto \theta, \phi$ such that, in the case of θ: $\theta = \theta\big(\lambda(t)\big)$. A lens is then applied to the setup (in a collimation configuration) in order to convert the polar and azimuthal angles into spatial coordinates: $\mathbf{C}_3 : \theta, \phi \mapsto x, y$ for a camera to then read it out: $\mathbf{M} : x, y \mapsto x, y$. Hence, the temporal information of the initial signal, which has been imprinted on by the event, is placed in unique x, y positions on the camera. With the knowledge of the mapping $x, y = \mathbf{M}\big(\mathbf{C}_3\big(\mathbf{C}_2(\mathbf{C}_1(t))\big)\big)$, one can trace back to the temporal information of the signal photons and in effect the event of interest. This is a prime example of a linear indirect optical diagnostic method, that is able to distinguish between photons that are within 100s of femtoseconds from each other [14]. It might seem like all these operations on $\mathbf{U}$ to attain x, y information is redundant. Indeed, why not just stop at λ? However, as we will see later in this thesis, converting t to x, y brings with it many benefits in terms of bandwidth, use of resources and versatility [15].

Post-processing step

There is one more operation that is nearly always necessary, and that is the post-processing step, $\mathbf{P}$. It is a mapping within measurement space $\mathbf{V}$, such that:

$$\mathbf{P} : \mathbf{V} \mapsto \mathbf{V}. \tag{2.3}$$

In the case of STAMP, the post-processing step encompasses the spatial separation of timestamps t_1 to t_9 (in Fig. 2.7b). Masking each sub image constitutes the final step to separating the photon temporal information into individual images. There can exist more complicated algorithms, such as those used for compressive ultrafast photography (CUP) where an algorithm based on compressive sensing is employed [16]. In this case, the mapping $\mathbf{P}$ is paramount to separating the photon's temporal information as all the operations prior to measurement are performed specifically with the algorithm in mind.

3 The Category of Time

We now move on to discussing the details of categorising photons with respect to their temporal information. After all, the ultrafast videography part of this thesis' title implies that time[13] has constituted a central part of my PhD work.

Causality

Categorising photons depending on their timestamps is of fundamental importance when addressing *causality* within an event. To illustrate, Fig. 2.8a shows four images: a campfire, a forest fire, a burnt down forest and a healthy forest. These images are taken by a fire department for monitoring forest fires in the area, with the catch that all the time information has been lost, i.e., we do not know the order of the four images. For the fire department this information is essential to addressing the root cause of dangerous events and preventing future fires. The causality relations between these images might seem obvious to the reader, however this thought experiment is nonetheless included in order to illustrate the importance of timestamps on the images. For all we know, the images might be taken years in between - the forest could have burnt down years ago and the image of the campfire in a healthy forest is quite recent!

Motion Blur

Other than being able to look into causality relations, sometimes one wants to "freeze" a certain event and see what occurred at a given moment in time. In technical terms this would amount to capturing and discriminating for photons that emanate from the event within the interval $[t, t + dt]$. In fact, this was the motivation for Edward Muybridge's first high-speed camera where freezing the motion of a galloping horse allowed the interested parts to settle the longstanding debate of whether all four hooves actually left the ground during a single cycle (see figure on page ix). Fig. 2.8b illustrates this for the event of a bullet penetrating an apple. The blur is caused by the fast motion of the bullet and the ensuing explosion, where all emanating photons fall within a single temporal category. If the detection window in time is constrained enough, then the bullet does not move while the detection system collects and categorises the photons, resulting in this "freezing" in time. In technical terms this amounts to the elimination of *motion blur*.

[13] Or the lack of it...

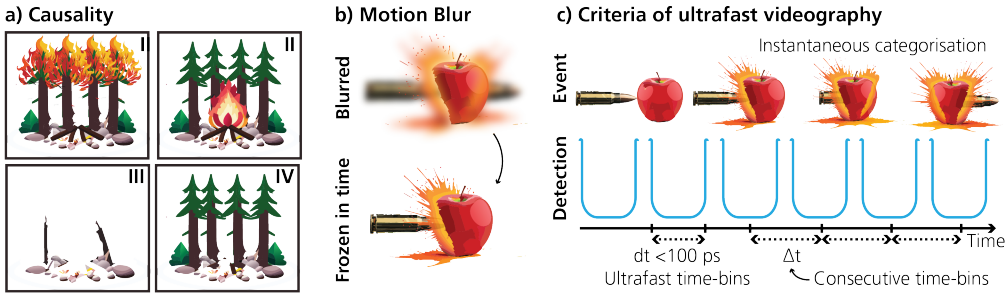


Figure 2.8: The category of time - its significance and considerations.
a) Four pictures of a forest are acquired over a finite timespan. In order to follow the causality relations between the events, a timestamp is necessary. Indeed, how do we know if e.g., picture IV was taken before or after the fire? **b)** Motion blur induces a blurriness on the bullet and explosion. By detecting the event with a constrained photon temporal categorisation, it can be frozen in time for a clear view of the bullet penetrating the apple. **c)** Ultrafast, consecutive time-bins and instantaneous categorisation are fundamental criteria of ultrafast videography. Indeed, one only has a single attempt to film the exploding apple (unless of course you own an apple tree)

3.1 Criteria of ultrafast videography

The basis of the work performed in this thesis is regarding the development of an indirect linear technique dubbed the Frequency Recognition Algorithm for Multiple Exposures (FRAME). It is an *ultrafast videography* technique, which, in the context of the discussions and developed terminology throughout this chapter, amounts to: the instantaneous categorisation of photon information within consecutive, ultrafast time-bins. There are three key concepts to unwrap within this definition (all illustrated in Fig. 2.8c):

1. **Ultrafast time-bins** reflects the fact that the discrimination of photons with respect to their temporal information is on the order of 100s of picoseconds down to femtoseconds and beyond. That is, two photons emanating from an event with a time difference of dt ($\lesssim$100 ps) should be categorised uniquely. This is directly related to the concept of motion blur.

2. **Consecutive time-bins** reflects the videography aspect of this thesis. Namely that it is not enough to have a short temporal width, but the time-bins must also follow each other back to back. This means that two photons emanating from an event with a time difference of Δt (e.g., $\Delta t = 100$ fs) should be captured and uniquely categorised.

3. **Instantaneous categorisation** reflects the fact that this temporal categorisation occurs all at once, i.e., the photons emanating from a single event are categorised into the consecutive time-bins immediately. This is in contrast to repeating the event and shifting the time-bins in order to catch photons of different temporal stamps for each repetition of the event[14].

[14]This is typically what is performed in pump-probe techniques, hence this constraint is necessary.

It is the combination of all these three criteria along with two-dimensional imaging within a self-contained diagnostic technique that necessitates an entire PhD thesis (and much more) to develop. Furthermore, it is said combination, dubbed ultrafast videography, that has the potential to open up for the investigation of new physics, chemistry and biology on ultrafast timescales.

In order to accentuate the difficulties in combining all three criteria, understanding the technological context within which this work has been performed, and grasping the implications of combining these constraints, the following chapter will discuss existing technologies and the compromises that have to be made in order to achieve high-speed videography. I will then end with discussions on optical diagnostic techniques that are in fact able to combine all three criteria in various ways such that they are able to reach the realm of ultrafast videography. With that, let us start digging into these details!

Chapter 3

Current State of Technology

Gotta go fast
– zonic hegehog

In this chapter I will provide an overview of existing technology in the context of the aforementioned three prerequisites of ultrafast imaging, namely: ultrafast time-bins, consecutive time-bins and instantaneous acquisition. This will be followed by a presentation of the field of ultrafast videography, i.e., existing techniques that satisfy all three criteria. But first, we must discuss the technological workhorse that makes this thesis possible – digital imaging devices, collectively known as focal point arrays.

1 Focal Point Arrays

Other than temporal categorisation, the field of videography is fundamentally dependent on imaging, i.e., the act of capturing and categorising spatial information from photons emanating from a scene. A central device in the field of imaging is the Focal Point Array (FPA). The FPA consists of a discrete array of photosensitive components, called pixels, that convert photon energy to electronic signals to then be read out in either digital or analogue form. There are many types and subtypes of FPAs, however in this thesis we will be concentrating on Charged Coupled Devices (CCD) and Complementary Metal Oxide Semiconductor (CMOS) technology [17]. These are by far the most widely used devices within the field of imaging due to their relatively low price, their extreme spatial bandwidths and the ease at which they can be mass produced[1].

1.1 FPA technology from 1970 to 2023

The first CCD sensor, dubbed a charge 'bubble' device, was proposed and experimentally demonstrated in 1970 by Gil Amelio and Michael Francis Tompsett. The invention consisted of a row of closely spaced metal squares placed on a bed of oxidised silicon (called pixels) and connected by simple wires – a design for which they were awarded the Nobel Prize in 2009 [4, 18]. This invention was closely followed by more sophisticated designs until in 1974 the first commercially available CCD *array* was invented, consisting of 100×100 pixels (0.01 megapixels) and a readout circuit including an amplifier and 8-bit shift register [19]. This set the stage for the development of 2D digital photosensitive arrays, a technology that has developed exponentially with respect to sensor size, pixel size and, the more difficult to quantify, quality[2] (Fig. 3.1a). Furthermore, with the 2023 announcement of the 200 MP Samsung ISOCELL HP2 and its individual pixel size of 600 nm [20], digital sensors are today more spatially sensitive, or continuous, than medium grade photographic paper[3].

CCD and CMOS technology

Both CCDs and CMOS are based on Metal Oxide Semiconductor (MOS) technology and differ mainly via the readout process (Fig. 3.1b). In short, for a CCD, the photoelectrons embedded within each pixel are read out through a single amplifier and analogue to digital converter (ADC). This process is slow and necessitates large amounts of power. In con-

[1] This fact is directly related to the mass production of silicon based chip technology.

[2] This would include noise characteristics, shutter time, dynamic range etc.

[3] The grain size of medium-grained photographic emulsion (consisting of photosensitive silver halide crystals) are on the order of 350 - 700 nm [21].

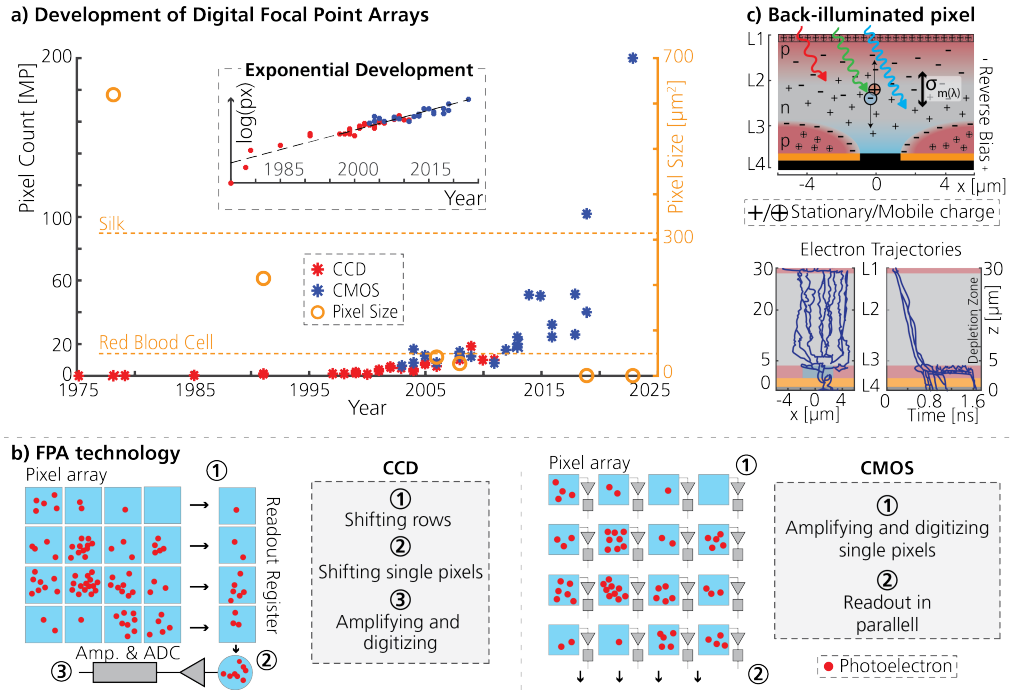


Figure 3.1: FPAs to measure photon spatial information.
a) The historical development of photosensitive array size, showing a clear exponential trend akin to Moore's law for transistor size. Simultaneously, the size of the individual pixels has decreased past the size of individual grains in photographic paper (right orange axis). **b)** The CCD and CMOS sensor differ by the readout approach, where, in the case of CCDs, the photoelectrons are all shifted through registers to a single readout node, while for CMOS technology, each pixel has its own readout circuit. **c)** The back-illuminated pixel is the pixel of choice for high-speed applications. Simulated electron trajectories (lower panel), show where and for how long electrons are stuck in the various pixel layers, fundamentally limiting the maximum frame rate of such technology. Adapted from [22]

trast, CMOS sensors have an individual readout circuit for each pixel which unfortunately results in a significantly increased number of noise sources as compared to CCDs. Even though this fact has historically kept the CMOS in the position of the underdog, huge technological improvements combined with ease of production, low power consumption, and fast readout, has propelled CMOS FPAs into the realm of scientific laboratories and more importantly nearly every pocket in the world [23].

1.2 Speed limit of digital sensor technology

In the context of this thesis, we are interested in the speed limits of our imaging devices[4]. Unsurprisingly, the primary limiter is the readout circuitry, where signal electrons need to be physically moved to and from different positions and pass through various electrical components. Methods of limiting the electron travel distance by storing signal electrons

[4]This discussion is centred around the intrinsically faster CMOS technology.

close to the photoactive site for later readout do exist (see later section) however parasitic capacitance and electron repelling forces set a, for now, technological limit at speeds around 200 Mfps [24–26].

A more fundamental culprit is actually embedded within the single pixel itself, or more specifically in the photoelectron generation and transport process. The back-illuminated pixel (the pixel of choice for low-light, high-speed conditions due to the large photoactive area) is built around a deeply depleted pn-junction such that when a photon creates an electron-hole pair within it, the two are immediately separated and the electron travels downwards towards the readout circuitry (Fig. 3.1c) [27]. Monte Carlo simulations show that if the electron-hole pairs are created at the top end of the pixel (L1 or upper L2), the most amount of time is spent in section L3 where photoelectrons perform a horizontal random walk around the p-wells before they drop into the circuitry, L4. This sets a limit at 1 Gfps for this type of technology. However, by applying waveguides such that electron-hole pairs are created at $x = 0$, this drift can be eliminated. In conjunction with a smaller depth of 10 µm, simulations have shown that 10 Gfps is attainable [24].

The actual insurmountable limit is however related directly to the penetration depth of photons before they produce an electron-hole pair. The average penetration depth of green light at 550 nm is 1.73 µm resulting in a travel time to the circuitry of 11.1 ps, pushing the limit to 90.1 Gfps. However, this penetration depth varies with colour resulting in travel times ranging from 4 ps for 450 nm light to 30 ps for red light, hence a temporal broadening of $\sigma_{m(\lambda)}$ on a delta pulse consisting of differently coloured photons is always unavoidable [22].

2 Ultrafast Videography - History and Terminology

We have now looked into the fundamental building block of imaging, the FPA. Before investigating how these are put to use in conjunction with other types of technology for high-speed and ultrafast videography, let us delve into a little history lesson with Fig. 3.2a. In the end of the nineteenth century, the industrialist, horse enthusiast and ex-governor of California, Leland Stanford, wanted to once and for all settle the debate of whether a galloping horse ever completely leaves the ground[5]. Therefore, he hired the convicted (and acquitted) murderer and natural photographer Edweard Muybridge with the aim of solving this conundrum. From the years 1873 to 1877, Muybridge developed sensitive photographic emulsion, a fast trip-wire based shutter system and a "film studio" consisting of a white backdrop and the strong California sun as a light source. This setup was situated on what

[5]There are rumours that he also had a $25000 bet on all four legs leaving the ground in a gallop, in contrast to a trot. The true irony is that he ended up spending around $50000 to settle the debate. Classic rich 1800s men, am I right.

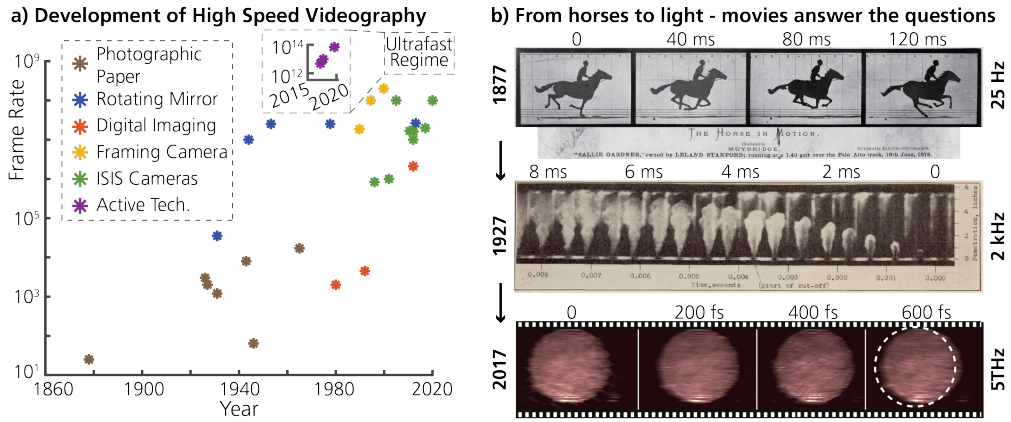


Figure 3.2: History of high-speed videography.
a) The history of high-speed videography with respect to frame rate is plotted for different technologies. A similar exponential behaviour as sensor size (Fig. 3.1) can be seen with sudden jumps upon the invention of new technologies. For the bibliography of this plot, see Appendix B. **b)** 25 Hz in 1877 through 2 kHz in 1927 to 5 THz in 2017, videography has answered fundamental questions in many fields thanks to technological advances in sensor and illumination technology.

is now the Stanford University campus. The captured movie of a jockey[6] galloping past the tripwires, acquired at a speed of about 25 frames per second (top panel of Fig. 3.2b is an earlier version, with only four images) ended up settling this debate and launched Muybridge's success in filming the locomotion of humans and all sorts of animals [28].

For the first instance of kHz imaging, one must wait, albeit not too long, to 1927, where Edward G. Beardsley was able to illuminate a fuel spray from an oil injection with 20 individual high speed consecutive sparks at the Langley Research Center of the National Advisory Committee for Aeronautics (NACA, the precursor of NASA) [29]. Being top-secret work, the public only saw this type of technology 10 years later when it was being applied to the filming of, among other objects, sprays and propellers [30]. After this pioneering work, MHz imaging was first achieved by rotating mirror cameras – technology that had been developed to film hydrogen bomb detonations after World War II [31]. Today, with a category of filming dubbed active videography[7], the THz limit has been breached (bottom panel of Fig. 3.2b) and is here applied to the filming of light travelling through a Kerr medium[8] [32].

[6]The jockey and the story behind him has been adapted in the movie Nope by Jordan Peele. The jockey's name was Gilbert Domm and the horse Sallie Gardner. Furthermore, the original movie of the galloping horse exists on IMDB, listed as Sallie Gardner at a gallop.

[7]We will see later that this is essentially going back to how Beardsley was able to reach kHz frame rates but with faster illumination.

[8]This is work published by my supervisors.

2.1 Bookkeeping of videography terms

We will now delve into the world of videography with an emphasis on the three criteria mentioned at the end of the previous chapter, namely time-bins (hitting the ultrafast regime at the end of the section), consecutive time-bins and instantaneous categorisation. All techniques and technologies presented will already satisfy the property of instantaneous categorisation such that the interplay between the criteria of ultrafast and consecutive time-bins can be thoroughly discussed. Furthermore, we will see that the size of the recorded images (remember we must have images for it to be classified as videography) exhibit a non-trivial relation towards these two criteria.

However, before diving into said beautiful world of videography, some bookkeeping terms need to be agreed upon. The reader might notice that I love to categorise everything that can be categorised, I should probably become an evolutionary biologist. But here we are...

Passive vs active videography

There are two possible ways to divide temporal information carried by photons into time-bins [33]. The most straightforward way is to actually catch the photons from the event in one's net and then divide them using linear/nonlinear discriminating optics or electronics. An example of such a discriminator would be the aforementioned mechanical shutter. This type of categorisation is called *passive videography* (Fig. 3.3a). Passive videography techniques perform a direct mapping from photon space to detection space, $\mathbf{M}: t \mapsto t$, while still maintaining sufficient x, y information in order to satisfy the videography criterion. FPAs with temporally dependent shuttering capabilities are prime examples of passive videography equipment.

Most of the time, as we will see later on, passive videography techniques simply cannot

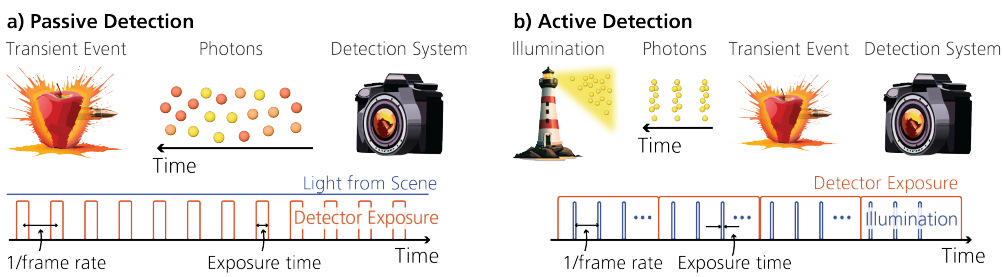


Figure 3.3: Two types of videography approaches - passive and active.
a) Passive videography captures and discriminates photons directly emanating from the scene. The frame rate and exposure times are uniquely set by the detector. **b)** In active videography, the scene is usually illuminated with a specially designed pulse train to then be captured by the detector. Most often, the detector captures multiple light pulses within a single exposure to then be divided in post-processing. The frame rate and exposure time are set by the illumination.

satisfy all three ultrafast videography criteria. In such cases, one can resort to performing an indirect measurement via a coupling or mapping within photon space, $\mathbf{C} : t \mapsto \gamma$, for γ one or more of the eleven photon dimensions. Then by knowing the characteristics of $\mathbf{C}$ and discriminating based on the coupled degree of freedom, γ, one can indirectly categorise the original temporal information. This usually entails the exploitation of specially designed light pulse trains which illuminate the event, hence these techniques fall under the category of *active videography* (Fig. 3.3b). There are pitfalls that must be considered when applying active videography techniques, mainly embedded within the coupling operation. For example, if γ is the x, y degrees of freedom, one is limited in the coupling by the imaging constraint of videography. If on the other hand γ is wavelength, λ, then one is limited by finite spectral bandwidths of the incoming photons and furthermore in the detection of spectral information.

Examples of techniques and technologies that fall within these two categories will follow within this chapter.

Frame rate, sequence depth and image size

In order to demystify and quantify what is meant by ultrafast time-bins and consecutive time-bins, we will turn to the terms *frame rate* and *sequence depth*. These two terms are directly related to the criteria that we must satisfy and their values are strongly dependent on the type of videographic technology that is put to use in a given measurement situation. Another less obvious term that is intrinsically connected to the aforementioned criteria is the *image size*. Indeed, depending on the applied technique the image size can vary over a wide range of values. For some definitions:

1. **Frame rate:** The rate at which consecutive, unique images,[9] or frames, can be acquired within a single measurement. Its value is given in Hz or frames per second (fps). To put it in numbers: if two unique images can be acquired within one second then the frame rate is 2 Hz or 2 fps[10]. Note that, in the limit where the frame rate is maximized for a given system (essentially always the case in this thesis) it is oftentimes directly related to the exposure time (dt). Therefore, the frame rate is inherently connected to both the time-bin size and the consecutive time-bins criteria. In the context of this discussion only technology that attains frame rates above 100000 Hz (100 kHz) will be considered.
2. **Sequence Depth:** The sequence depth of a videography system is the number of consecutive, unique images that the system can acquire at a given frame rate. A min-

[9] We will discuss what is meant by unique in the last chapter. Until then, it is enough to understand that the content of an image is unique with respect to other images within the video.

[10] For comparison, nowadays computer screens run at 144 Hz.

imum sequence depth of four should be necessary in order to enter the videography regime. This, since a minimum of three consecutive images are necessary to calculate the accelerations (or forces) acting upon an event, therefore, in order to see the result of such quantities one more image will be needed. Due to the instantaneous categorisation criteria and lower limit of 100 kHz frame rates, the sequence depth will most often be a finite number and quite limited for the majority of presented techniques.

3. **Image size:** The image size is given as the number of pixels (px) for which the technique or technology is able to acquire photons for every time-bin. This enters the category of fundamental videography quantities since in order to follow the spatial unfolding and correlations of an event, a certain spatial extent and resolution is necessary[11]. In a similar manner to the sequence depth, the stringent ultrafast criteria that we have set for this thesis result in quite small image sizes, however, due to the exponential growth of FPA sizes we will see that this facet is endowed with a huge potential for development.

3 Overview of Existing Technologies

These bookkeeping terms function as the basic properties that need to be considered upon application of a high-speed videography technique for the detection of a dynamic transient event. For example, if one is to look at the light emission of a nanosecond discharge plasma, one needs a *passive* solution (we just want to look at the light emitted from the scene) with sub nanosecond temporal resolution (>GHz *frame rate*). Furthermore, since this is a stochastic process, one wants to catch the entire event within a single acquisition, hence a *sequence depth* that can allow for a temporal coverage which covers the entire event is necessary. In this case, we would have to resort to a streak camera[12].

In the context of these terms let us finally look into the current state of the field of high-speed and ultrafast videography. To assist us in this discussion, I have taken the liberty[13] of creating Fig. 3.4 from a large scale literature review. It shows a selection of videography

[11]Furthermore, videography is defined as the practice of recording images, hence each frame needs to consist of an image. Let us keep going down the rabbit hole: an image is defined as the visual representation of [...] an object produced on a photographic material or electronic display. A representation is to bring clearly to the mind an artistic likeness or image of an object. I hope I have made myself clear here that the image is a fundamental aspect of videography, hence image size is a fundamental variable in the characterisation of videography systems.

[12]Even though the image spatial information is confined to a single dimension, this is unfortunately the only solution that exists.

[13]And a lot of salary money from my supervisor Elias.

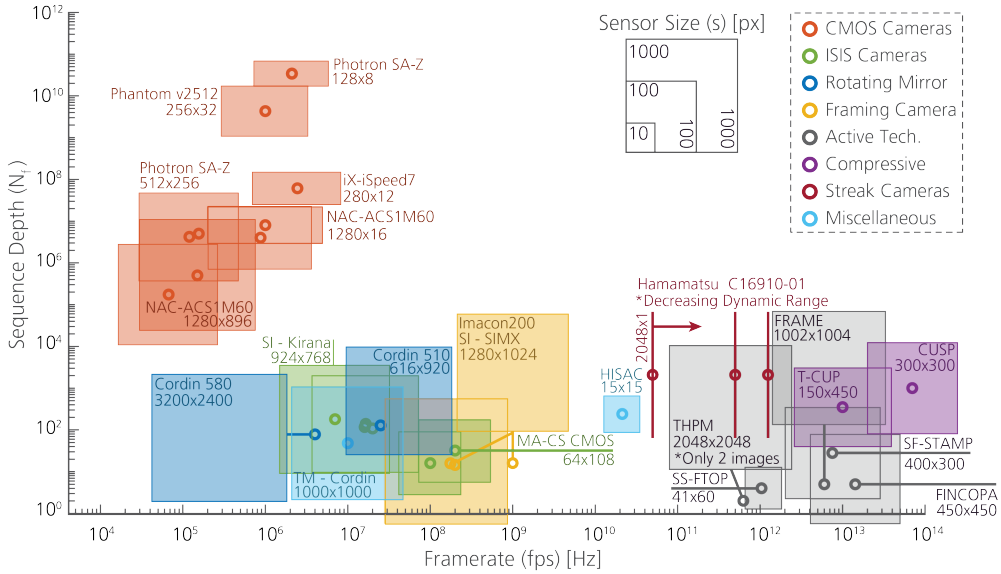


Figure 3.4: An overview of the current state of high-speed and ultrafast videography.
A log-log-log diagram with the literature values for a representational selection of technologies and techniques in terms of frame rate and sequence depth (represented by the placements of the circles) and image size (represented by the boxes). These are colour coded with respect to the type of technology and most of the model/technique names are included beside the corresponding boxes. Note the empty space in the upper right corner: the holy grail of ultrafast videography where ultrafast speeds along with long sequence depths are simultaneously achieved! The inspiration for this plot is taken from Phillip Reu, Sandia National Laboratories, who has created similar plots in the past found in [34] or by a quick google of "Phillip Reu high-speed imaging".

technologies and techniques in terms of frame rate, sequence depth and sensor size where the colours represent various classes of technologies.

3.1 Passive videography solutions

We begin by looking into technologies that fall under the umbrella of passive videography solutions. These consist of the red, green, blue, yellow and burgundy boxes of Fig. 3.4.

High-speed CMOS cameras

High-speed CMOS cameras (the red boxes in Fig. 3.4) are probably the most widely used type of high-speed videography solution. In contrast to the sequential readout architecture of CCD sensors, the parallelised readout of CMOS sensors allows for pixel readout speeds on the order of 10^{10} pixels per second. After readout the images are directly transferred to a high-speed onboard storage drive, hence, this technology is characterised by extremely long sequence depths, bordering on continuous acquisition [35]. The interplay between

directly writing to memory and a constant pixel readout rate results in two characterising traits for these types of solutions. Firstly, as the frame rate increases, the acquired images necessarily become smaller where a rule of thumb is that they can attain 10^4 fps for 10^6 pixels and vice versa [36]. This is illustrated as the negative linear correlation of Fig. 3.5a for four state of the art models. Secondly, as these images become smaller, more can be stored on the memory, hence a positive correlation between frame rate and sequence depth can be observed in Fig. 3.4. Consequently, a single acquisition runs for, in general, equivalent time lengths irrespective of frame rate. Overall, this class of high-speed videography solutions is adequate for many applications ranging from the life sciences [37, 38] to energy, combustion and plasmas [39, 40]. The main con is the hard speed limit of the electronics, in particular the direct readout and writing to memory, placing their current upper boundary at 2.1 Mfps [35].

In situ storage image sensor cameras

In order to overcome these limiting factors associated with high-speed CMOS technology, Walter Kosonocky invented the *in situ* storage image sensor (ISIS) in 1996 [41] (green boxes in Fig. 3.4). Since then, swift development with the aim of reaching GHz frame rates has been led by Takeharu Goji Etoh and colleagues [42], where the electron transport research presented in Fig. 3.1c is a part of said development. The basic principle is that around each photosensitive pixel there is a group of storage pixels to which photoelectrons from a given exposure are moved and stored within until they are read out [41, 43–48]. The key here is that the slow readout process occurs *after* the entire acquisition has been completed. The location geometry of these temporary storage pixels is key for the attainable frame rate as momentum changes of electrons (quickly switching their direction) affects the fidelity and speed at which frames can be acquired. For example, creating a curvilinear CCD stack with a slight slant results in a smoothly varying potential for electrons to travel through, a design that was able to double the frame rate of ISIS technology [49]. Nonetheless, there exists negative correlations between frame rate, sequence depth and sensor size all attributed directly to geometry constraints. Indeed, a certain geometry of pixels is necessary in order to move electrons fast enough, constraining the number of storage pixels, hence sequence depth. Equivalently, higher frame rates necessitate larger pixels, hence, in conjunction with geometry, the sensor sizes become smaller. Nevertheless, multiple commercial options are currently available with large image sizes and MHz speeds adequate for use within fundamental [50] to applied research [51, 52]. However, as before the hard limit with this type of technology is photoelectron transport through the pn-junction and further electronics.

Rotating mirror cameras

The inner workings of the rotating mirror camera (dark blue boxes in Fig. 3.4) involves relaying the light emanating from a scene onto a gas turbine-driven mirror that then rotates and *streaks* the event over individual FPAs (Fig. 3.5b). In order to avoid motion blur from this streaking mechanism, each FPA is equipped with an imaging lens that passively guarantees image formation at discrete positions. The image from each FPA is hence a temporally resolved representation of the event at the time the mirror is at a certain angle. Due to its mechanical nature, the frame rate and sequence depth of these systems is directly limited by mechanical constraints on the rotating mirror[14] in conjunction with challenges involving image conservation and the bulkiness of large relay structures. Furthermore, the multiple FPAs necessary can introduce variations with respect to e.g., sensitivity and linearity between images within a single acquisition. The lack of correlation between frame rate and sequence depth for the two exemplary models in Fig. 3.4 goes to show the brute force nature of the technique; if you want more of one or the other, just add sensors, spin the mirror faster or add more pixels. With the Brandaris research facility in Twente, Netherlands [53] these types of cameras have been utilised, among other things, for performing pioneering work within the field of microbubble acoustics for drug delivery [54, 55]

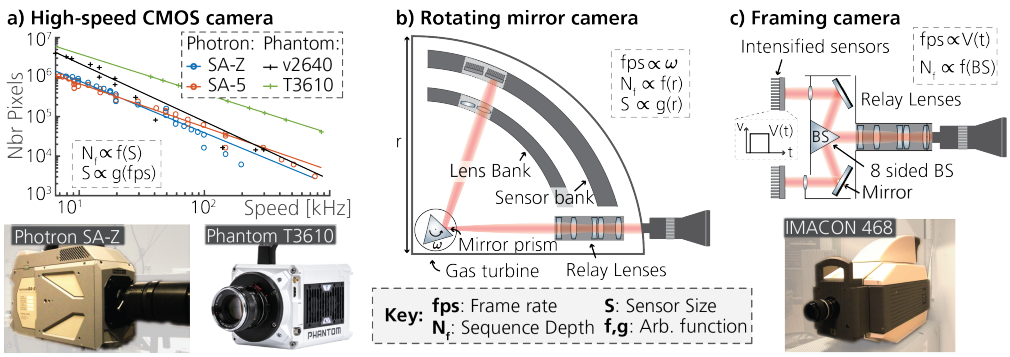


Figure 3.5: Passive high-speed videography solutions.
a) The design of high-speed CMOS cameras results in a negative correlation between the frame rate and image size (number of pixels) set by the constant 10^{10} px/s. Hence, the individual image size (S) is dependent on the frame rate (fps) and the sequence depth (N_f) is dependent on the image size. **b)** The rotating mirror camera relays the scene onto a prism, rotating at an angular frequency of ω. The image will be streaked via a lens bank onto individual sensors. The frame rate is dependent on ω, while the sequence depth and sensor size are functions of the radius of the camera. **c)** The framing camera relays the scene onto a beamsplitter prism (BS) which splits the light onto multiple intensified sensors. The frame rate is dependent on the voltage gating function of the sensors while the sequence depth is dependent on the number of sides on the beamsplitter prism.

[14] It is interesting to note that the rotation speed of gas pressure driven rotation devices is significantly higher when helium is the working gas as compared to nitrogen.

Ultrafast framing cameras

When light emanating from a scene enters an ultrafast framing camera (yellow boxes in Fig. 3.4) it is first divided by a beamsplitter prism (BS) to then be imaged onto multiple intensified sensors (iCCD or iCMOS where the "i" stands for intensified and implies the use of an MCP) with, in some cases, sub nanosecond gating capabilities (Fig. 3.5c) [56–58]. This allows framing cameras to reach up to GHz frame rates limited only by the electronics of the MCPs in terms of jitter and rise time. The sequence depth, on the other hand, is strictly limited by the number of sides on the prism, where the most common is an octagonal type arrangement for eight images[15]. Unfortunately, the division of light in this manner strains the system in terms of intensity, a scarce commodity for high-speed imaging, hence the practical sequence depth is usually limited by the lighting characteristics of the scene [59]. Pushing these cameras to their extremes can end up affecting the sensor size, even though the FPAs are fixed. This is due to the spatial gating of the MCPs which "open" in an iris like fashion. Hence, if the frame rate is pushed beyond the complete opening of the MCPs, the effective sensor size will decrease. The design of ultrafast framing cameras is comparable to the rotating mirror configuration, hence there is no clear correlation between frame rate and sequence depth, making these cameras perfect for applications within fusion and combustion applications [60, 61]. Approaches to increasing the frame rate of this type of technology resort to applying light-based shuttering in non-linear crystals, a technique dubbed FINCOPA [62]. Even though it technically falls under the category of active videography techniques, the basic principle is akin to the passive version.

Streak cameras

Streak cameras (burgundy lines in Fig. 3.4) are one dimensional temporally resolved imaging systems. In short, photons enter through a slit and are converted to electrons in a photocathode to subsequently be accelerated towards a so called streak tube where an ultrafast voltage ramp is applied (Fig. 3.6, adapted from [63]). This voltage ramp has the effect of deflecting electrons at a given angle depending on their time of arrival at the tube, and by extrapolation the time of arrival of photons at the photocathode, essentially converting a temporal distribution into a spatial distribution on a camera. This constrains the images to a one dimensional line where spatial information is acquired along the x-axis and temporal information is acquired along the y-axis. These types of cameras were invented in 1949 [64] and, although Hamamatsu breached the ps barrier in 2023 with their M16911 − 01 sweep unit [65], streak camera temporal resolution has always stayed quite stagnant around the 1 − 2 ps mark. The main con is that they are severely affected by the so-called space-charge effect, i.e., the Coulomb repelling force between electrons as they travel through

[15] This type of beamsplitter prism 'cuts' the spatial Fourier domain resulting in quite a lot of astigmatism.

the system. This smears out the spatiotemporal response of the system affecting the linearity, the dynamic range and the frame rate of streak camera systems. Furthermore, in stark contrast to previously mentioned high-speed videography solutions these three features are entwined, an effect that is not visible in the plot of Fig. 3.4a even though it directly affects the plotted variables. Streak cameras are widely used in applications where the temporal evolution of the total light emanating from a scene is of interest such as laser and plasma physics or fluorescence lifetime imaging [66–69]. There is a plethora of details that I cannot go through in such a short paragraph, hence the interested reader should be referred to Paper v and vi.

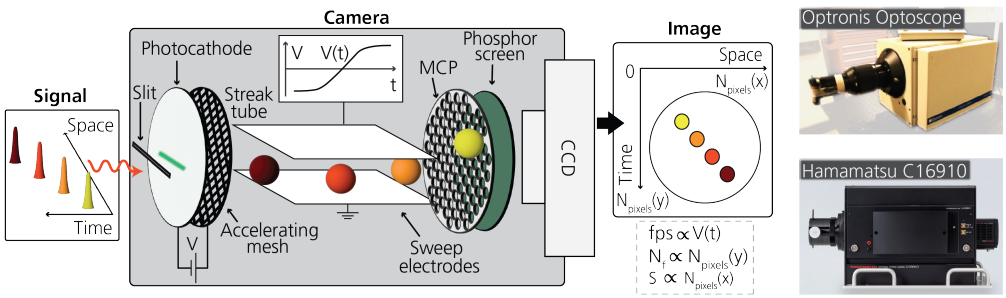


Figure 3.6: Inner workings of the streak camera.
Four pulses separated in space and time are incident on a photocathode, creating photoelectrons. These are then accelerated and incident onto a so-called streak tube. Here, an ultrafast voltage ramp deflects them at an angle dependent on their time of arrival at the tube, converting a temporal distribution into a spatial distribution in the final image. Schematic adapted from Hamamatsu's Guide to Streak Cameras [63]

3.2 Active videography solutions

Up until now we have concentrated on so-called passive methods of videography where the frame rate is dependent on the shuttering speeds, either via electronics (in its extreme: the inertial mass of the electron) or moving mechanical parts. Even if we find a way to circumvent this, there is one more limiting factor: namely that passive approaches measure the light that is emitted from an event (see Fig. 3.3a). Let us make a back of the envelope calculation on how much light would be necessary in order to passively reach THz frame rates.

The sun is a good reference for a luminous object to be filmed with a passive high-speed videography solution. The solar irradiance at the top of the atmosphere, with the sun directly above the measurement site, is 1360 W/m^2. At a wavelength of 555 nm (the candela and lumen is defined at this wavelength) this corresponds to 1360 $Js^{-1}m^{-2}/3.58 \times 10^{-19}$ J $= 3.80 \times 10^{21}$ photons per second per square meter ($s^{-1}m^{-2}$).

For a reasonable spatial resolution and light sensitivity, we use a camera pixel of size 10 µm $\times$10 µm (this quite a widespread size for high-speed videography solutions). At THz frame

rates, the exposure time is 1 ps or 10^{-12} seconds. Hence each pixel will receive:

$$N^{\text{photons}}_{\text{px,1THz}} = 3.42 \times 10^{21}\ \text{ph s}^{-1}\text{m}^{-2} \times 10^{-10}\ \text{m}^2 \times 10^{-12}\ \text{s} = 0.38\ \text{ph/(px fr)}, \qquad (3.1)$$

where the units are ph: photons, px: pixels and fr: frame (the result is in units of photons per pixel per frame). With this simple calculation, we see that, statistically speaking, one photon will be incident on the pixel every three frames. Including a value for the quantum efficiency will lower this number even further. Therefore, the frame rate will have to be lowered from THz speeds to about 300 GHz in order to hand-wavingly speaking obtain any signal at all from the sun within a single take. The reader might already be thinking, but Vassily, this calculation is flawed, your pixels can be made bigger! Indeed, at a single pixel size of 100 µm $\times$100 µm there will be 38 photons per pixel per frame incident upon it. For a reasonably sized 1 MP image with a sensor fill factor of 1 this would correspond to a sensor size of 1 $\times$ 1 cm, a completely manageable feat. The difficulties arise when considering the engineering challenges for the creation of the low noise conditions that are necessary for detection of these few photons (noise scales with pixel size), the fact that we are looking straight at the sun hence there are no losses in the imaging stage[16] and that there are not many natural processes that emit that much light passively. Hence, a purely passive system would be limited in its applications. Furthermore, the dynamic range of such an image will be quite lousy (< 100 for even quite large pixels) hence even the use of ultrafast amplifiers that could achieve such speeds would not make a difference.

A possible solution around these problems is by illuminating the object of interest with a strong light source and look at the shadow, an example of active videography. Indeed, if we were to focus the sun's light to a small area and illuminate an object, then there would in fact be enough light for THz videography of the shadow of the event. In a more realistic scenario, the use of high-fluence laser pulses less than 1 ps long, opens up for the ultrashort illumination of dynamic scenes with enough light to reach beyond THz imaging speeds.

There are however challenges related to active videography mainly caused by the need to record the video with a regular, "slow" camera. Furthermore, the repetition rate of these pulses should be in the THz regime in order to reach said THz videography. In short, these problems are solved by tailoring ultrafast optical pulses into pulse trains and performing coupling operations in photon space such that their temporal information can be captured on regular cameras. Due to the incredible development of FPAs in terms of number of pixels and pixel size, i.e., spatial bandwidth (Fig. 3.1a), these coupling operations are usually performed such that the final mapping to detector space ends in the projection of photons' temporal information onto spatial (or spatial frequency) coordinates of the cameras, i.e.,

[16] One could focus the image onto the sensor, however for a given spatial resolution one would simultaneously have to make the pixels smaller.

we always want some kind of $\mathbf{C} : t \mapsto \gamma_i \mapsto x, y$ before performing $\mathbf{M}$. In the following section we will investigate how this is achieved in some existing techniques.

Spatial division techniques

This category of ultrafast videography is characterized by the direct coupling between photons' temporal dimension to their spatial dimension, i.e., $\mathbf{C} : t \mapsto x, y$. A camera at the end of the setup then performs the mapping $\mathbf{M} : x, y \mapsto x, y$[17]. A prime example of this is the SS-FTOP technique (Fig. 3.7). A femtosecond laser beam is led towards a spatially stepped quartz echelon and is then focused onto the probe volume. The echelon performs two operations. Firstly, as the femtosecond beam passes through, different spatial parts of it will be delayed, dividing the incoming photons into different, discrete time-bins, i.e., creating a pulse train. Secondly, due to its spatial structure different spatial areas on the incoming beam are delayed by different amounts. Hence, the echelon performs the mapping $\mathbf{C} : t \mapsto x, y$. Upon detection by the camera, different x, y, areas on the sensor will correspond to different time-stamps, creating (after a post-processing step that spatially masks the images), in its current version a 1.04 THz, four frame video [70]. Due to the low spatial extent of the femtosecond pulse, and the trade-off in FOV vs spatial resolution of the imaging optics, the individual image sizes are quite small, only 41 × 60 pixels. The main con of this type of active technique is the trade-off between energy density of the individual pulses and the sequence depth. Furthermore, the use of a quartz echelon for the pulse-train creation limits this technique in terms of frame rate due to the introduced chirp as the femtosecond pulse passes through the dispersive optical component. The technique is also extremely sensitive to the spatial profile of the pulse in order to maintain an equal illumination between illuminating pulses.

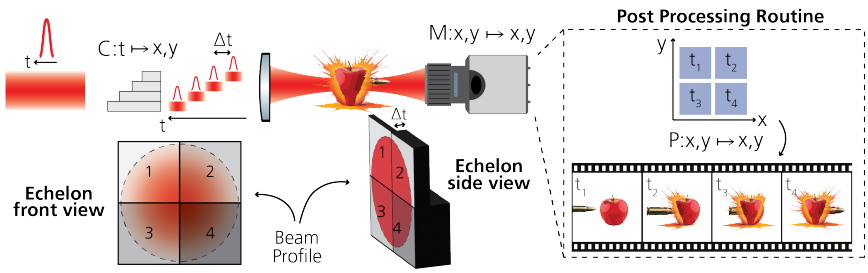


Figure 3.7: SS-FTOP - spatial division.
SS-FTOP utilises an echelon of spatially distributed thicknesses in order to simultaneously delay and divide the beam profile into four temporally separated pulses. Upon detection, the individual time-delayed areas of the image are extracted forming a video.

[17]Remember that it is always beneficial to perform the mapping $\mathbf{C}$ to x, y in photon space due to the high spatial bandwidth of modern FPAs.

Wavelength division techniques

This subclass of active videography techniques utilises the high spectral bandwidths of femtosecond pulses as the coupling dimension to time. Once time has been coupled to wavelength, many options open up in terms of optics that are able to turn this into a spatial division for the creation of a video. The best example of wavelength division is sequentially timed all-optical mapping photography (STAMP) [71] and in particular its follow-up spectrally filtered (SF-) STAMP [14] (Fig. 3.8). Here, a high bandwidth pulse is chirped in a glass rod coupling time to wavelength $\mathbf{C}_1 : t \mapsto \lambda$. This entire pulse is then incident onto the sample, recording its transient dynamics. It then passes through a diffractive optical element (DOE) and a spectral band-pass filter (BPF) tilted at an angle such that only light of certain wavelength will pass through at a given diffractive order of the DOE. This coupling operation can be described as $\mathbf{C}_2 : \lambda \mapsto \theta, \phi$. A lens then converts these angles to a spatial position to be imaged on an FPA resulting, in its current configuration, in a 7.52 THz video with a sequence depth of 28 [14]. SF-STAMP's exposure time and hence frame rate depends on the band-pass filter properties and its sequence depth on the pulse bandwidth. The main problem with wavelength division techniques is that, by "locking" the wavelength dimension of the photons, these techniques cannot detect spectroscopic information from dynamic scenes. This is not unique to STAMP as in general any active technique relies on coupling photon dimensions to time, sacrificing the bandwidth of said dimension in order to increase the technique's temporal bandwidth. However, oftentimes wavelength is considered a "more important" dimension, hence it constrains STAMP's application to physical events more severely than other techniques. Similar designs have been adapted to various applications [72–75].

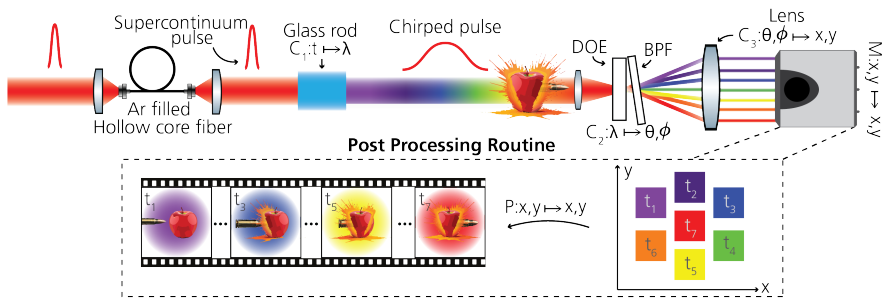


Figure 3.8: SF-STAMP - wavelength division.
SF-STAMP is performed by creating a supercontinuum pulse that is then chirped, effectively coupling wavelength to time. The pulse is then incident on the event and divided into individual angles with a diffractive optic element (DOE). A tilted band pass filter (BPF) subsequently chooses distinct colour bands depending on the angle of incidence of the beam, resulting in a spatial separation between the colours on the camera. Since the colours were coupled to time, the result is a spatially separated video of the event to be extracted in the post-processing step.

Light-in-flight holography

By using the streaking effect that a pulse of light generates when it is incident on a detector at an oblique angle, one can, in principle, attain a temporal resolution that is determined by the speed of light passing a single detector element [76]. To be more specific, a pulse of light is incident on a detector and functions as an ultrafast shutter as it interferes with light emitted from a sample. Extraction then occurs by undoing the created interference pattern with a continuous wave laser (incident at the same angle) or, in the digital case, by algorithmic extraction (Fig. 3.9a). Technically, this is a form of space division however the separating detail is the use of the electric field phase, Φ, as the information carrier resulting in an interference pattern in the x, y domain forming the basis of extraction [77]. The first example of single shot videography of a travelling light pulse was actually performed using this approach in 1978 by Nils Abrahamsson in KTH, Stockholm [5]. Since this technique is dependent on the speed of light and the coherence length of the reference pulse, in its current state it achieves speeds of 11.4 THz and sequence depths of 7 [78]. The main con, is that one does not directly image the light but rather one is dependent on the light-matter interaction within a scattering medium (diffusor plate in this case). Furthermore, this limits its applicability to the measurement of other transient events.

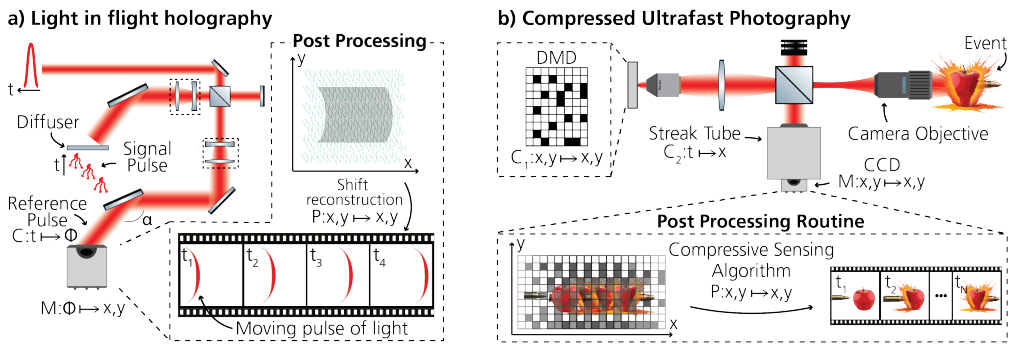


Figure 3.9: Holography and Compressive Sensing.
a) Light in flight holography utilises an obliquely directed reference beam that arrives at the sensor simultaneously with the signal, creating a hologram where the signal's temporal information is coupled to the phase information of the electric field. After reconstruction the moving pulse of light can be extracted. **b)** When applying the compressed ultrafast photography technique (CUP), light from an event is superimposed with a random binary structure to then be streaked in time and recorded with a streak tube and camera. A compressive sensing algorithm extracts the recorded movie of the event.

Compressive ultrafast photography

Compressive Ultrafast Photography (CUP) has its basis in the mathematics of compressive sensing. Light emanating from a scene is imaged onto a digital micromirror device (DMD) such that a binary pseudorandom pattern is superimposed onto the signal of interest, i.e., a coupling of the form $\mathbf{C}_1 : x, y \mapsto x, y$ is performed (Fig 3.9b). These photons are then incident on a streak tube with a wide open slit that has the ability to shear the entire image across a CCD sensor, i.e., performing operations $\mathbf{C}_2 : t \mapsto x$[18] and $\mathbf{M} : x, y \mapsto x, y$. After acquisition, a compressive sensing algorithm is employed in the mapping within measurement space, $\mathbf{P} : x, y \mapsto x, y$. CUP [16, 33] and in particular its follow-up CUSP [79]. (which combines CUP with a wavelength division approach) are able to reach speeds of 70 THz and sequence depths of 1000 frames. Technically, they are classified as passive techniques, since no illuminating light is necessary, however, due to the low luminosity of most objects when filmed at such speeds, it is difficult to monitor anything but laser-based light-matter interactions. The main con to this class of techniques is the speed limit that is set by the streak camera electronics and the limiting space-charge effect. However, there does exist approaches that would be able to employ light-matter [80] and light-light interactions [81] in order to perform the streaking, thus eliminating this limiting factor from the CUP technique. Further applications of similar nature are found in [82, 83].

Spatial frequency division techniques - FRAME

Finally, we are here, the first time I introduce the centre of this thesis, a technique called Frequency Recognition Algorithm for Multiple Exposures, FRAME. FRAME, in this context, is an active ultrafast videography technique that categorises photons into instantaneous, ultrafast time-bins. This categorisation is performed in an indirect manner with the coupling $\mathbf{C} : t \mapsto x, y$ in photon space such that the detection can occur with the mapping $\mathbf{M} : x, y \mapsto x, y$ (Fig. 3.10). This is quite akin to spatial division, however the coupling between t and x, y is performed such that in detection space all pulses are superimposed, hence one cannot divide them in a trivial manner. Instead, $\mathbf{C}$ is performed such that when the x, y information undergoes a spatial Fourier transform to be represented in terms of f_x, f_y[19], discrete separated peaks show up. Each peak corresponds to the time-bins within which photons are categorised. Performing a spatial lock-in detection algorithm in the post-processing, $\mathbf{P}$, the spatiotemporal information can be extracted.

The main advantage of spatial frequency division is the field-of-view that is attainable as

[18] In contrast to previous explanations where the temporal information is streaked along the y-axis, here I describe it as being projected on the x-axis. In practice this is achieved by physically flipping the streak tube.

[19] f_x, f_y indicates that the representation occurs in reciprocal or Fourier space, i.e., $\mathcal{F}(x, y)$ for $\mathcal{F}$ the Fourier transform operation.

compared to spatial division along with the lack of stereoscopic effects, immediate pixel to pixel correspondence and general versatility of the approach in terms of multidimensional imaging. The main con however, is the loss of spatial resolution upon application of the lock-in algorithm. The first example of spatial frequency division applied to imaging, albeit not connected to time, can be traced back to a computer science department in Louisiana [84] followed closely by the same author in Istanbul, Türkiye [85]. The first time it was applied to ultrafast videography was within the department of Combustion Physics here in Lund in 2017 [32]. To continue, within this division, the technique has been applied in multiple configurations, with applications on multiple dimensions of photon information – from depth information (z) to polarisation information (ψ, χ) [86–88]. Beyond these, the papers included within this thesis all constitute some version of spatial frequency division for encoding photon information within a snapshot. Finally, citations outside our group are necessary, as the approach seems to have caught on. The following list constitutes some of the more novel approaches within which this technique has been applied [89–96]

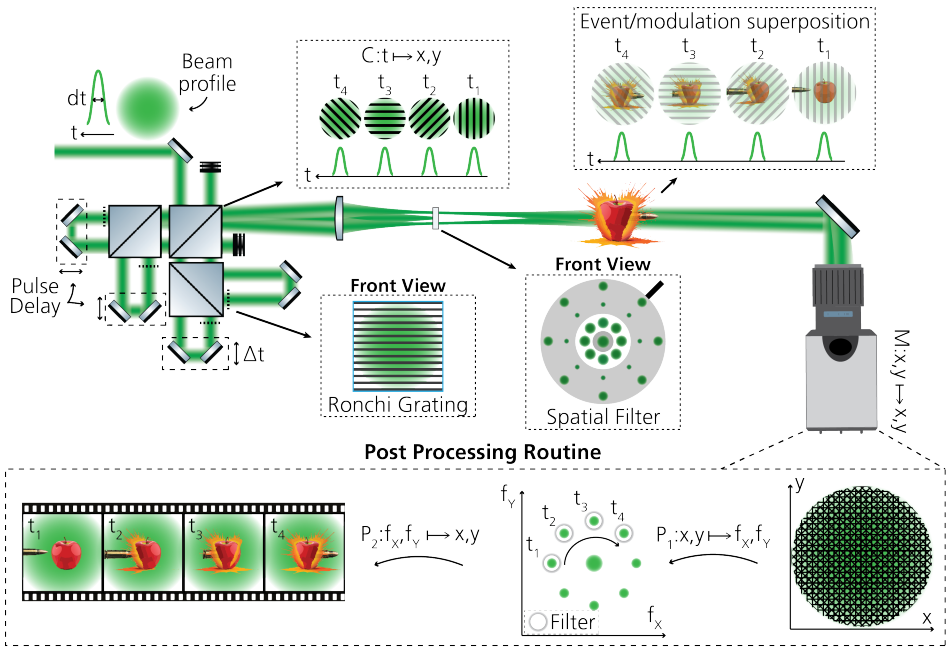


Figure 3.10: FRAME - spatial frequency division.
An optical pulse enters a beamsplitter/delay stage setup where each arm is imprinted with a unique spatial intensity modulation using a Ronchi grating. Passing through a spatial filter, only certain orders of the Ronchi grating's diffraction pattern are allowed through to the sample. Snapshots of the event, illuminated in a stroboscopic manner by means of the tailored pulse train, are then superimposed upon the modulations and detected within a single exposure. The total image, consisting of four spatially modulated snapshots of the event, can be Fourier transformed in the post processing where a lock-in detection algorithm is applied. In this step, they are separated and filtered such that a video of the event can be extracted.

Chapter 4

Fourier Analysis - In Theory and Experiment

When in doubt, perform the Fourier transform.
– at least one lecturer per semester during my physics education

The Fourier domain, what a concept, is simply put a fancy way of representing signals of various kinds. If the reader has heard of Fourier transforms, you might also know that, according to many physicists, it seems to explain and solve all the world's problems: from the Heisenberg uncertainty principle in quantum mechanics[1] to the prediction of climate change[2]. Typically, at university physics courses, lecturers throw the term around in the spirit of "obviously the Fourier transform is the way to go" and then they look at their notes and copy the results of said transform. This discourse is similar to the way the Taylor expansion is thrown around willy nilly. I definitely do not blame anybody as I am guilty of this as well. As a student however, this is most often magic as it gives the impression that all the questions are already solved.

Before beginning, I would like to discuss the word *intuitive*. This word is used quite often in physics discourse when trying to explain things to each other. This can be compared to the way "obviously" is used in physics textbooks when discussing equations and causalities. The definition of intuition is "the power or faculty of attaining to direct knowledge or cognition without evident rational thought and inference" (definition from Merriam-Webster

[1]This might not count as a world problem in most eyes.
[2]Unfortunately this counts as a problem in way too few eyes.

dictionary). In the field of experimental physics (I can only speak to this field), intuition can be compared to a gut feeling and can solve surprisingly many problems. However, something that is intuitive for me might not be intuitive to others. Therefore, when I use the word intuitive in this chapter it should be taken with a tablespoon of salt, Fourier transforms and the Fourier representations of images are NOT at all intuitive, they are in fact strange and mesmerising. Once you get the hang of it though, it opens up a completely new way of seeing and understanding signals, or within this thesis, images.

With that in mind, in this chapter I will not attempt to be too mathematically formal in my description of Fourier mathematics, nor will I explain it in the most correct way. For such an in depth and complete text I will refer the reader to [97], a text I have used thoroughly for the creation of this chapter. Instead I will try to give the reader some gut feeling or intuition as to how the Fourier transform is applied in the context of optics and imaging, what properties so-called natural images have in terms of their Fourier transforms and how the Fourier transform can be manipulated in the lab[3], and on the computer in the post-processing phase. But first what is it?

1 The Fourier Formalism

Optical imaging in terms of the wave equation can be described by linear mappings from object light distributions to image light distributions. This property opens up for the use of linear systems theories in their descriptions and brings with it the property of being able to explain complicated light distributions with simple elementary functions. In the context of optical imaging, i.e., the propagation of spatial distributions of light waves, the elementary function of choice, $h(\mathbf{r})$, is the complex exponential[4]:

$$h(\mathbf{r}) = \exp(i\,\mathbf{f}\cdot\mathbf{r}), \tag{4.1}$$

where $\mathbf{r}$ is a 2-dimensional vector with Cartesian spatial coordinates as the entries, $\mathbf{f}$ is a 2-dimensional vector with so-called spatial frequencies as the entries and i is the imaginary unit. The set of these elementary functions form a complete orthonormal basis, hence can be used to fully describe spatial signals. All the notations used within this chapter are summarised in Appendix C.

A spatial distribution of light, or an image, $g(\mathbf{r}) = g(x, y)$, can be decomposed into a linear combination of said (simple) elementary functions with weights $G(\mathbf{f}) = G(f_x, f_y)$, such that:

[3] Yes it is not only a mathematical construct but we can actually see it in the laboratory.

[4] This group of functions form the basis of the solution to the Helmholtz equation in free space, i.e., the solution to Maxwell's equations for electromagnetic radiation in the absence of a source or matter.

$$g(x,y) = \iint_{-\infty}^{+\infty} G(f_x, f_y) \exp\left[i\,2\pi(f_x x + f_y y)\right] df_x df_y, \tag{4.2}$$

where the factor of 2π enters as the frequency normalisation factor. As previously mentioned, this linear decomposition opens up for the analysis of the complicated signal $g(x,y)$ via the analysis of much easier to handle elementary functions, multiplied by their weights, $G(f_x, f_y)$. The reader might have gotten ahead of me at this point, but the next step is the determination of these coefficients by performing the so-called Fourier transform of g, $\mathcal{F}\{g\}$:

$$G(f_x, f_y) = \mathcal{F}\{g(x,y)\} = \iint_{-\infty}^{+\infty} g(x,y) \exp\left[-i\,2\pi(f_x x + f_y y)\right] dxdy. \tag{4.3}$$

The only difference between Eq. 4.3 and Eq. 4.2 (also called the inverse Fourier transform, $\mathcal{F}^{-1}\{G\}$), is the negative sign in the argument of the elementary function and the variables over which the integration is performed. In the context of this thesis, the function $g(x,y)$ will be said to inhabit real or image space and the values $G(f_x, f_y)$ will be said to inhabit Fourier, frequency or reciprocal space. These concepts are illustrated in Fig. 4.1a, where the signal in real space (facing towards the left) can be described as a summation of four frequency components with different amplitudes (facing to the right). This goes to show that a Fourier transform is in essence a rotation of state space to a different basis spanned by the complex exponentials.

All of us have in fact encountered Fig. 4.1a and the "rotation" of a signal into its Fourier representation. When playing the piano one presses a selection of keys and out comes a sound wave described by some periodic function, such as the left facing part of Fig. 4.1a. On the other hand, the keys that were pressed can be described as the right facing part. In essence, the keys of the piano are the Fourier transform of the sound waves that are emitted from the instrument, and the pressure at which one presses said keys corresponds to the coefficients G, or the *amplitude* of the sound wave components. I would recommend the reader to find an instrument and play around with the idea that you are performing a Fourier transform just by taking a tone. Quite neat right?

Before continuing with more mathematics, we must establish that these coefficients, G, in fact exist and are meaningful to work with. Without getting into too much details we can take it from the experts that "physical possibility is a valid sufficient condition for the existence of a transform" [98]. Hence, in the work presented within this thesis, the Fourier transform always exists!

1.1 Convenient theorems and properties

Below I will list and discuss some theorems and properties of the Fourier transform and its corresponding elementary functions that are central to the discussions within this thesis:

- **Directionality and extent of the elementary functions:** For a given frequency pair, (f_x, f_y), the complex exponential has a phase of 0 or $n2\pi$ (for n, an integer) along straight lines which lie at an angle $\theta = \arctan(f_x/f_y)$ in relation to the x, y axes of a Cartesian coordinate system (this can be compared to how a sine or cosine consists of crests and troughs). This means that they are in a sense "directional", a property that is essential in imaging considerations. Furthermore, they have a spatial period given by $1/\sqrt{{f_x}^2 + {f_y}^2}$, hence image features of different scales can be described with different frequencies (discussed in depth later).

- **Translation:** Given that $\mathcal{F}\{g(x,y)\} = G(f_x, f_y)$

$$\mathcal{F}\{g(x-a, y-b)\} = G(f_x, f_y) \exp\left[-i2\pi(f_x a + f_y b)\right], \tag{4.4}$$

  i.e., a translation in real space introduces a phase shift in Fourier space.

- **Energy Conservations:** Parseval's theorem introduces the important energy and energy density concepts in the context of the Fourier transform. Given that $\mathcal{F}\{g(x,y)\} = G(f_x, f_y)$, then:

$$\iint_{-\infty}^{+\infty} |g(x,y)|^2 dxdy = \iint_{-\infty}^{+\infty} |G(f_x, f_y)|^2 df_x df_y. \tag{4.5}$$

  The left hand side can be understood as the energy contained within g, and since energy is conserved, $|G(f_x, f_y)|^2$ can be interpreted as the energy density of the signal in the frequency domain.

- **Convolution:** Given that $\mathcal{F}\{g(x,y)\} = G(f_x, f_y)$ and $\mathcal{F}\{h(x,y)\} = H(f_x, f_y)$, then:

$$\mathcal{F}\left\{ \iint_{-\infty}^{+\infty} g(\xi,\eta) h(x-\xi, y-\eta) d\xi d\eta \right\} = G(f_x, f_y) H(f_x, f_y). \tag{4.6}$$

  This is actually a strong statement, where the convolution in real space corresponds to a multiplication in Fourier space and vice versa. This has far reaching consequences as will be seen through the course of this thesis.

- **Localisation:** Each Fourier component represents a frequency that extends over the entire x, y space hence complete locality can only be attained via the summation of an infinite number of Fourier components. In practical applications though, a finite sum of Fourier components are sufficient for a spatial structure to become very small such that it falls under the noise floor at a rate of $1/f$ [99]. Hence, locality in both the Fourier and spatial domain can be experimentally attained simultaneously.

- **Order of operation:** Finally, at each point of continuity

$$\mathcal{F}\mathcal{F}^{-1}\{g(x,y)\} = \mathcal{F}^{-1}\mathcal{F}\{g(x,y)\} = g(x,y). \tag{4.7}$$

  This essentially means that no matter the order of performing the Fourier and inverse Fourier transform, the original function is always retained. This is directly related to the completeness of the orthonormal basis offered by the complex exponential elementary functions.

1.2 Sampling of continuous signals

FPAs, the main tool used to record light distributions in this thesis, are built up of discrete pixels which are given the task of *sampling*[5] light distributions at the plane of interest (most often the imaging plane). This piece of text concentrates on formalising this concept in terms of Fourier transforms and experimental images.

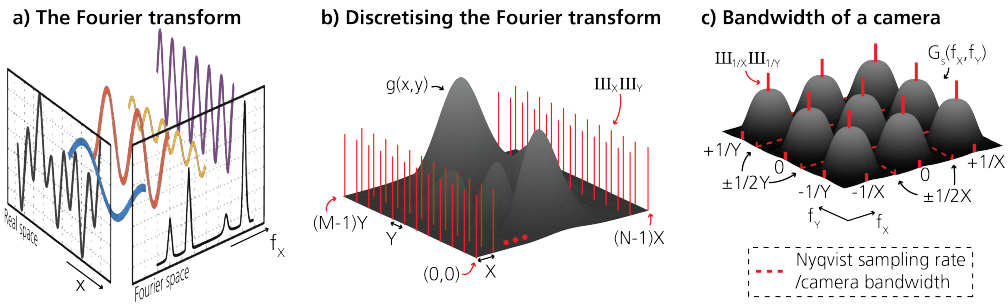


Figure 4.1: The Fourier transform and its discretisation.
a) A visualisation of how a signal in real space (left facing axis) can be built up of simple elementary functions (illustrated as the addition of the coloured sinusoidals). Upon rotation to the Fourier domain, one can instead represent the signal as discrete peaks where the heights are related to the amplitude of the modulation (right facing axis). In terms of music the real space represents what we hear while the Fourier space represents the tone (e.g., the key you press on the piano). **b)** A function g represented as the grey topology, can be finitely sampled at a spacing of X and Y by multiplication with the comb function (in red). **c)** When finitely sampled, the Fourier representation, G will be infinite in frequency space and the Fourier transform of g will be repeated at a spacing of $1/X$ and $1/Y$. Hence, the bandwidth of the detection system represented as $\mathrm{III}_X\mathrm{III}_Y$ is $1/2X$ and $1/2Y$, or the Nyquist sampling rate.

[5]The act of reducing a continuous signal to a discretised one by integrating and recording over a certain span.

The discrete Fourier transform

The discrete sampling of $g(x,y)$ (grey in Fig. 4.1b) by a camera can be mathematically represented by multiplying with the comb function, Ш (red in Fig. 4.1b), resulting in g_s:

$$g_s(x,y) = \text{Ш}_X \text{Ш}_Y g(x,y), \tag{4.8}$$

where the function Ш_T is a periodic function consisting of Kronecker deltas spaced by a distance T:

$$\text{Ш}_T = \sum_{n=-\infty}^{+\infty} \delta(x - nT). \tag{4.9}$$

In the context of a camera, n will go from pixel 0 to pixel $N-1$ where N is the number of pixels in the given direction. We insert this sampled version of g, Eq. 4.8, into Eq. 4.3 to derive its Fourier transform:

$$G_s(f_x,f_y) = \iint_{-\infty}^{+\infty} g(x,y) \exp\left[-i2\pi(f_x x + f_y y)\right] \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} \delta(x-nX)\delta(y-mY)dxdy, \tag{4.10}$$

where M is the number of pixels in the direction orthogonal to N. Due to the Kronecker deltas, this function will only be non-zero when $x = nX$ and $y = mY$, hence it boils down to:

$$G_s(f_x,f_y) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nX,mY) \exp\left[-i2\pi(f_x nX + f_y mY)\right]. \tag{4.11}$$

This is called the 2-dimensional Discrete Fourier Transform (DFT) and upon application of the Fourier transform on discretised signals, such as the images obtained from CCD or CMOS cameras, this is the approach that a computer will take. The algorithm of choice used to perform this transform is called the Fast Fourier Transform (FFT) which uses symmetries and matrix multiplicative magic in order to extract $G_s(f_x,f_y)$ from $g_s(x,y)$ in an optimised manner. In principle, the DFT can be evaluated at all frequencies, however since the sampling rate is restricted by $1/X$ and $1/Y$ (due to the structure of the comb function) only the harmonics, $1/nX$ and $1/mY$, (for $n \in [2, N-1]$ and $m \in [2, M-1]$) of the sampling frequency will be possible to probe. The astute reader may have noticed the number two as the lower bound within the brackets. This means that the maximum spatial frequency that can be probed is *half* the sampling rate. In real space, this implies that any structures that are smaller than $2X$ or $2Y$ will not be sensed by the camera in use. Let us look into why this is the case.

The Nyquist sampling rate

We perform Eq. 4.10. Since a multiplication in real space corresponds to a convolution in Fourier space it can be shown that:

$$G_s(f_x, f_y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} G\left(f_x - \frac{k}{X}, f_y - \frac{l}{Y}\right). \tag{4.12}$$

This means that the frequency spectrum of the sampled function g, is a periodic function consisting of the frequency content of g (i.e., G) around each point $(k/X, l/Y)$ in the infinitely sized frequency plane (an excerpt of said plane is shown in Fig. 4.1c). Examining $k = l = 0$, then we see that the next copy of G in the x direction lies around the frequency $\pm 1/|X|$ in the f_x direction and $\pm 1/|Y|$ in the f_y direction. The mid-point between these two is at the frequency coordinates $\pm(1/2X, 1/2Y)$, illustrated by the red dashed lines of Fig. 4.1c. This is the origin of the Nyquist sampling theorem: if a sampled signal is to be fully represented in frequency space without interference or distortion, then the sampling must occur at twice the bandwidth of said, band-limited signal.

In the context of imaging with FPAs, this means that the pixel spacing must be at least half the size of the bandlimited objects we are interested in for them to be represented without distortions. This makes intuitive sense because for e.g., a point to be visible, it is necessary to sample it once inside and once outside, i.e., two sampling points are necessary. With the current technological status in pixel size (and not number of pixels) this is almost always the case when imaging natural environments. In fact, since the imaging optics, which themselves function as low-pass filters, are quite often the limiting factors for the transmission of spatial frequency content of objects to images, it turns out that the Fourier domain of most images acquired with modern FPAs are quite void of information, i.e., the pixel spacing is tighter than what is necessary. As we will see, this is a key property for the functioning of the FRAME technique.

1.3 Constructing images with elementary functions

After all this mathematical formalism, let us now delve into a more intuitive approach to the Fourier transform. We will try investigate what the Fourier transforms of different functions $g(x, y)$ look like, or, to swap the order of explanation, let us look into how to build different $g(x, y)$ from the elementary functions. We start with a simple binary square as the function of interest (Fig. 4.2a). We know, from one-dimensional Fourier transforms, that a square waveform can be built up from a weighted sum of the odd frequencies[6]. In two dimensions, things are not too different. Let's have a look. The Fourier transform of

[6]Sorry I will let you google this one.

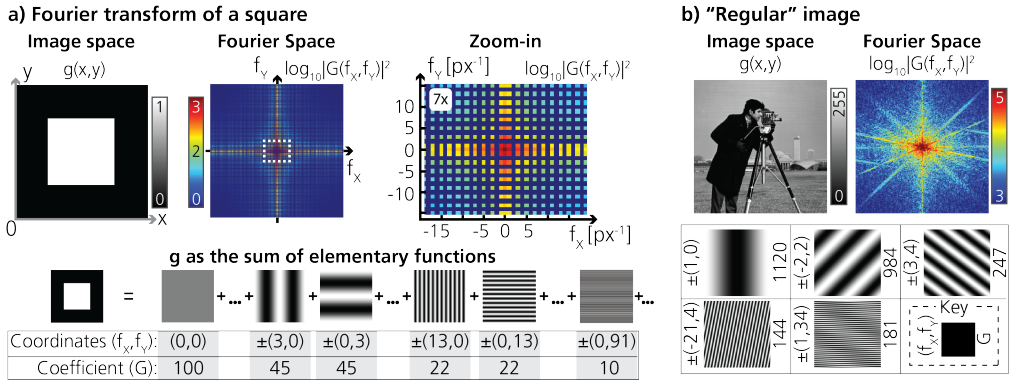


Figure 4.2: Building images with elementary functions.
a) A binary square in image space consists of odd frequency components in Fourier space. The zoom-in clarifies this concept where the coordinates represent the frequencies and the intensity values represented the coefficients, G. In other words, the square can be constructed out of elementary functions whose frequencies and coefficients can be found in Fourier space, as illustrated in the lower panel. Note that the depicted modulations are the real part of the complex exponential. **b)** The Fourier space representation of the cameraman is not as straightforward as the square in **a**. However, the same principle applies: it is built up of sinusoidals whose frequencies and coefficients are fully described by its Fourier representation, G.

the binary square in Fig. 4.2a, is depicted to its right along with a zoom-in. There are a couple of considerations that need to be taken into account in order to understand it:

1. It has been shifted, via the fftshift function, such that the low-frequency elementary function coefficients are shown in the centre and the highest frequencies, i.e., frequencies $1/2X$ and $1/2Y$ are placed on the ends of the coordinate system. This fits well into our Nyquist sampling rate formalism as the image in Fourier space is of the same size as in image space. Since it is mirrored about the centre, there are only half the number of unique frequency magnitudes compared to pixels on a given axis.

2. The pixel values are depicted as the logarithm of the absolute value of the coefficients, $\log_{10}|G|^2$, simply because this is the most convenient way to visualise them. Note that I have set the lower limit of G to be zero while the actual lower limit in this case is $-\infty$ ($\log_{10}(0) \to -\infty$). Again this is for visualisation purposes.

To summarise, in Fourier space, the *frequencies* of the elementary functions are depicted via their coordinates and the *coefficients*, G, are depicted via the pixel values.

Back to the image of the binary square. We see in the zoom-in that the Fourier coefficients of the square function alternate between positive values and 0, where non-zero coefficients are found at the odd frequencies, as we would expect. The meaning of this is that the image of the square, g, can actually be built up of the weighted sum of the odd frequency elementary functions plus the so-called DC component at $f_Y = f_X = 0$, i.e., the average power of the image, represented by a flat field. Stated in other terms, the square function

can be constructed via the constructive and destructive interference between the elementary functions with non-zero coefficients. The lower panel of Fig. 4.2a illustrates this concept with a subset of said elementary functions along with their coefficients.

If we instead consider a "regular" image, the Fourier domain is significantly more complicated, however the general concept remains true (Fig. 4.2b). Similarly to Fig. 4.2a, I have taken the liberty to extract a couple of the elementary functions that build up the cameraman along with their coefficients, presented in the lower panel. It is as simple as that.

At this point I must admit that when we work in the realm of Fourier optics and images, one does not always think in terms elementary functions at all. Instead, one looks into structures in G and through experience and intuition one has learnt to, in general terms, say how the image will respond to different operations, be it in the lab or on the computer. The ability to do this via intuition is actually not too far-fetched. Indeed, in real space one can track such image forming operations via convolutions, an arduous task if you ask me. On the other hand, due to the convolution theorem identical operations can be represented by a multiplication in the Fourier domain, a much lighter undertaking for human comprehension. Performing the correct multiplicative operations however requires an understanding into how Fourier components travel through an optical system. Therefore, in the following section I will try to give you, the reader, some useful tools that could assist in achieving this goal of laboratory intuition.

2 Imaging

As has previously been mentioned, but necessitates a reiteration, imaging is central to the field of videography. Most people, including the reader, know that in order to perform imaging, one needs a lens. But what does the lens actually do and how does it do it? There are many books into the exact mathematics of how image information is transferred through space and optical elements. For such an in-depth and mathematically correct explanation the reader is again referred to [97]. In this section however, I will offer a more intuitive approach that is, in my opinion, more useful for laboratory work and everyday life. We will start by looking into how the discussions of images as being built up of elementary complex exponentials can actually explain how image information travels in the form of photons through free space. We will then use this information to understand spatial filtering followed by a quick dive into lenses and how they actually perform a tangible Fourier transform.

2.1 Image propagation through free space

Consider coherent planar illumination[7] incident upon the binary square of Fig. 4.2a. As the light passes through the square, it will interact and be redirected in a well-defined manner. One can interpret this redirection as image information being transferred to the spatial and angular dimension of the photons as they interact with the object. This well-defined interaction can be described via diffraction theory and, in particular for this discussion, the *Fresnel diffraction integral*:

$$U(x,y) = \frac{\exp(ikz)}{i\lambda z} \iint_{-\infty}^{+\infty} \left\{ U(\xi,\eta) \exp\left[i\frac{k}{2z}(\xi^2+\eta^2) \right] \right\} \exp\left[-i\frac{2\pi}{\lambda z}(x\xi + y\eta) \right] d\xi d\eta, \quad (4.13)$$

where the coordinate system and symbols of the integral are included in Fig. 4.3. In words, the Fresnel integral says[8] that given a light intensity distribution in the object plane, $U(\xi,\eta)$ (equivalent to a function g), the intensity distribution at a point on another plane a distance z away, $U(x,y)$, will be the superposition of all the light waves emanating from all the distinct points in the object plane, multiplied by a spherical spatial phase and a "travelling" phase[9]. What a mouthful! In simpler terms, it essentially tracks the electric field phases of all the waves emanating from the object plane and then interferes them with each other, constructively, destructively and everything in between at the object plane. This integral fully describes (to the Fresnel approximation) the way a spatial distribution of light diffracts and travels through free space. We could therefore stop here and let the reader perform the mathematics. However, I would like to take a bit of an unconventional approach that fits in the context of images being built up of sinusoidal intensity modulations. This, since there are some key takeaways from such an approach that are extremely useful for practical applications.

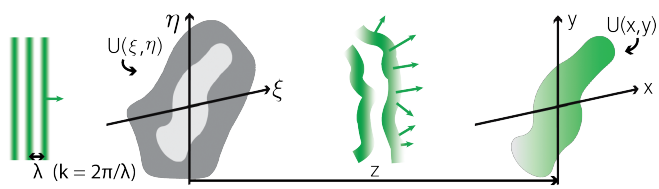


Figure 4.3: Image propagation through free space.
a) The coordinate system utilised when deriving and explaining the Fresnel diffraction integral. Monochromatic, plane waves, of wavenumber k, are incident from the left and are diffracted by the object in (ξ,η) space. The resulting intensity distribution in (x,y) space, at a distance z, is given, to the Fresnel approximation by Eq. 4.13.

[7] Coherent illumination means that the electromagnetic wave has a well-defined phase as it travels.

[8] This is my words skewed to the context of this thesis.

[9] For the spherical phase, think about this term as Huygens wavelets and for the travelling phase, this is to account the relative distances travelled by electric fields emanating from different points.

Fraunhofer approximation

In order to approach this nasty integral in a civilised manner there are three things we can realise about Eq. 4.13:

- The Fresnel approximation, also called the paraxial approximation, simplifies the Helmholtz equation for the propagation of light through free space, by introducing the condition that $z \gg (x-\xi)$ and $z \gg (y-\eta)$, i.e., the waves travel at a small angle with respect to the optical axis. This is in general true enough for the applications we are interested in.
- Aside from the multiplicative factors, we see that the Fresnel integral is actually the Fourier transform of $U(\xi,\eta)$ multiplied by a quadratic phase factor, $\exp[ik/2z(\xi^2+\eta^2)]$. This phase factor is conveniently only dependent on the distance travelled, z, and the original coordinates, (ξ,η).
- In terms of Fourier formalism, the transform (omitting the spherical phase) would have to be evaluated at frequencies $(f_x, f_y) = \left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right)$[10]. In other words, this means that the Fourier plane is in the (x,y) coordinate system and changes in size depending on the distance z.

With this knowledge, we can try to simplify the Fresnel integral with another approximation: the *Fraunhofer* approximation. If:

$$z \gg \frac{\pi(\xi^2+\eta^2)_{\max}}{\lambda}, \tag{4.14}$$

i.e., the distance away from the object is large compared to its spatial extent, then the spherical phase factor can be eliminated. Other than a multiplicative constant, Eq. 4.13 becomes exactly the Fourier transform of $U(\xi,\eta)$ evaluated at the frequencies $\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right)$:

$$U(x,y) = A \iint_{-\infty}^{+\infty} U(\xi,\eta) \exp\left[-i\frac{2\pi}{\lambda z}(x\xi + y\eta)\right] d\xi d\eta, \tag{4.15}$$

for all the prefactors combined into a constant A. In conjunction with the Fresnel approximation, for this to be true with red monochromatic illumination, $\lambda = 600$ nm, and an aperture width of 2.5 cm (1"), $z \gg 1600$ m. Even though this is an extremely stringent approximation, there are some useful lessons to be learnt.

[10] This part took me ages to understand so to avoid this unnecessary confusion for the reader, you get these frequencies by bringing in the $1/\lambda z$ from outside to inside the parenthesis.

Sinusoidal intensity patterns propagating through free space

As previously mentioned, even though we now have the tools to describe how any image's information travels through free space in the form of light, I would still like to digress via the formalism of sinusoidals. The reason will be obvious later on, so bear with me. Since an image can be built up of the weighted sum of complex exponentials, we will start by looking into how coherent light is affected by a sinusoidal spatial modulation. For this, we set:

$$U(\xi,\eta) = \frac{1}{2} + \frac{1}{2}\cos(2\pi f_0 \eta), \tag{4.16}$$

i.e., the real part of a complex exponential with frequency f_0 in the η direction (Fig. 4.4a). The function has been shifted by a factor of $1/2$ in order to guarantee non-negative values (in the context of imaging, U, being the light *intensity* distribution, cannot be negative). Inserting into Eq. 4.15, $U(x,y)$ will take the form:

$$U(x,y) = A\left(\mathcal{F}\left\{\frac{1}{2}\right\}^{(f_x,f_Y)} + \mathcal{F}\left\{\frac{1}{2}\cos(2\pi f_0\eta)\right\}^{(f_x,f_x)}\right) \tag{4.17}$$

$$\Rightarrow\ U(x,y) = \frac{A}{2}\delta(f_x,f_Y) + \frac{A}{4}\delta(f_x,f_Y+f_0) + \frac{A}{4}\delta(f_x,f_Y-f_0). \tag{4.18}$$

Hence we see that, in the far field, a sinusoidal intensity structure turns into three peaks located at positions $f_Y = 0$ and $f_Y = \pm f_0$. We know that $f_Y = \frac{y}{\lambda z}$ hence, the spatial position of the peaks in the far field will be at $(x,y) = (0,0)$ and $\pm(0, f_0\lambda z)$. This means that when the photons of a coherent plane wave interact with the sinusoidal, they contract an angle of propagation of $\theta = \arctan(f_0\lambda) \approx f_0\lambda$. Note the dependence of this angle on the spatial frequency of the sinusoidal intensity pattern, f_0. To spell it out further, a higher spatial frequency, f_0, results in a larger angle at which the corresponding information travels

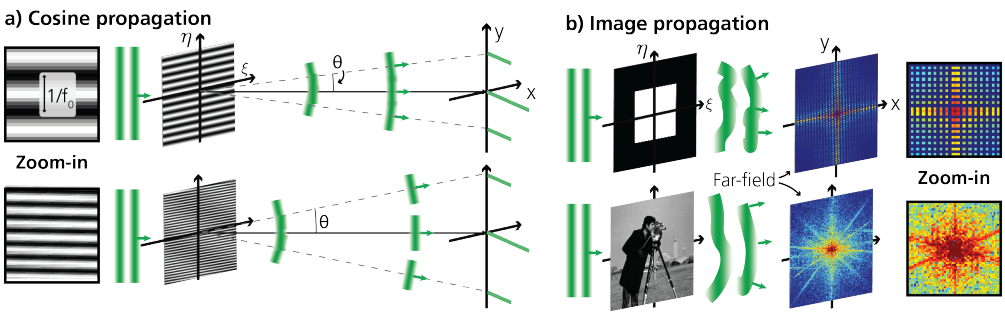


Figure 4.4: Propagation of structured light fields through free space.
a) When light passes through sinusoidal intensity structures of frequency f_0, it obtains a well-defined angle of propagation in the far field, θ, where larger f_0 results in a larger angle. **b)** Each sinusoidal component of an image will cause the incident light to diffract at a certain angle. In the far-field we therefore end up with the Fourier transform of the original light field.

through free space and vice versa[11] (Fig. 4.4a). This fact has far reaching consequences in the context of image formation and filtering in the laboratory, because if one can somehow discriminate for various angles (we have technically done this in the first chapter), the practical result is that one can filter out certain chosen frequency components that make up an image.

The lesson from these manipulations is that, when light interacts with e.g., the binary square image, it will essentially travel along paths fully characterised by angles with respect to the optical axis[12]. This is due to the fact that the square is built out of the sum of sinusoidals with varying frequencies and, as we have seen, each sinusoidal causes light to travel at a well-defined angle. Upon illumination with coherent light, we will therefore see its recognisable Fourier transform in the far field (Fig. 4.4b). The same goes for the cameraman, all the frequency components visible in the Fourier domain correspond to the angles and intensities of each part of the light field as it has interacted with the cameraman and travelled through free space.

2.2 Fourier transforming properties of lenses

We have seen in previous chapters that the lens maps photons' x, y information to x, y coordinates, functioning as a type of information relay. But it has one more strong characteristic, it can perform a physical Fourier transform.

Due to the higher refractive index of glass (usually what lenses are made of), a phase delay is introduced to light waves as they pass through the medium. For a lens, the amount of phase delay is dependent on spatial position and is given by the thickness function, t_l, where the higher up on the lens, the less glass the light goes through resulting in it picking up less of a phase delay. Under the paraxial and thin lens approximation, lenses are constructed such that the introduced phase delay on light waves passing through is given by:

$$t_l(x, y) = \exp\left[-i\frac{k}{2f}(x^2 + y^2)\right], \tag{4.19}$$

for a lens of focal length f. We recognize this spherical or quadratic phase factor from the Fresnel integral of Eq. 4.13. To continue, consider an object, with intensity profile $U_o(\xi, \eta)$, placed right up against a lens ($d_o = 0$ in Fig. 4.5) that is illuminated by monochromatic coherent planar waves. On the other side of the lens, the light will have picked up a phase

[11] Compare this to the diffraction patterns obtained from grating structures etched into e.g., CDs, a regular high-school experiment.

[12] This is not completely true as the approximation holds only in the far-field, but for illustrative purposes and in the lab, this is true enough.

as described in Eq. 4.19, transforming into:

$$U_l(x,y) = U_o(\xi,\eta)\exp\left[-i\frac{k}{2f}(x^2+y^2)\right]. \tag{4.20}$$

Inserting this into the Fresnel integral and allowing it to propagate a distance $z = f$ results in:

$$U_f(u,v) = \frac{\exp\left[i\frac{k}{2f}(u^2+v^2)\right]}{i\lambda f}$$

$$\times \iint_{-\infty}^{+\infty} U_o(\xi,\eta)\cancel{\exp\left[-i\frac{k}{2f}(x^2+y^2)\right]}\cancel{\exp\left[i\frac{k}{2f}(x^2+y^2)\right]}\exp\left[-i\frac{2\pi}{\lambda f}(\xi u+\eta v)\right]d\xi d\eta$$

$$= A\mathcal{F}\{U_o(\xi,\eta)\}, \tag{4.21}$$

for A, a phase factor to compensate for the travelling electric field, and U_f the light distribution at the focal plane. Hence we see that the lens eliminates the quadratic phase curvature in the same manner as the Fraunhofer approximation. This results in, to a multiplicative phase factor, a Fourier transform of the input light distribution at the focal point evaluated at $(f_x, f_y) = (\frac{u}{\lambda f}, \frac{v}{\lambda f})$ (Fig. 4.5). When performing measurements, or visualising the signal at the focal point of the lens (usually with a screen or camera), we see the intensity of the electric field. Hence, only the absolute squared of Eq. 4.21 or the *energy spectrum*, $|G(x,y)|^2$, of the input intensity distribution $U_o(\xi,\eta)$, will be measured.

If the input, $U_o(\xi,\eta)$, is at an arbitrary position, d_o, in front of the lens (instead of $d_o = 0$ for the previous calculation), this result still holds to a "travelling through free space for a distance d_o" phase factor. However, again, it will be eliminated upon the measurement of intensity. Hence, when performing imaging with a lens of e.g., the square or the cameraman of Fig. 4.4b, the energy spectrum shown in that figure will correspond to what is detected at the focal point of the lens (without the logarithm operation). In terms of imaging in the laboratory and the FRAME technique, this is an extremely strong asset, as Fourier domain manipulations can be performed just by placing various optical elements in the focal point of a lens.

2.3 Image formation

By placing an object, say the cameraman, at a distance d_o from the lens and following the Fresnel integral past the focal point, one can show that an image of it, $U_i(x,y)$, will be formed at a position d_i, where the relationship between d_o and d_i is entirely described by

the focal length of the lens[13]:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}\,. \tag{4.22}$$

This is called the thin lens formula and is the way one performs imaging in the lab (Fig. 4.5). The ratio between d_i and d_o is the magnification M:

$$M = \frac{d_i}{d_o}, \tag{4.23}$$

where its sign and value gives the orientation and size of the image, $U_i(x, y)$, with respect to the object $U_o(\xi, \eta)$.

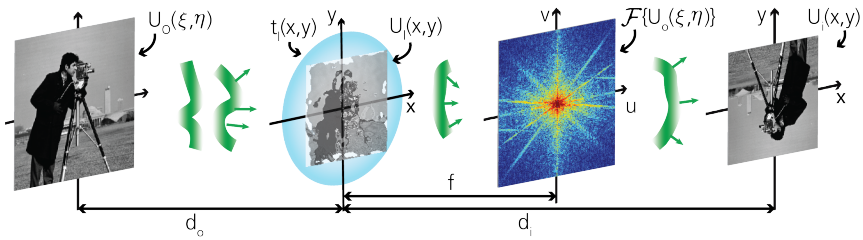


Figure 4.5: Imaging with a thin lens.
A light field, given by the cameraman image, $U(\xi, \eta)$, propagates a distance d_o through free space towards a lens with phase transformation function $t_l(x, y)$. It comes out on the other end as $U_l(x, y)$ to then travel a focal distance away where the measured intensity has the form, to a multiplicative factor, of the Fourier transform of $U_o(\xi, \eta)$, $\mathcal{F}\{U_o(\xi, \eta)\}$. This information continues to propagate until it reaches a distance d_i from the lens, where an image, $U_i(x, y)$, of $U_o(\xi, \eta)$, is formed. Note that its sign (if its upside down) and size depends on the magnification, M, of the imaging system. This whole process can be described by the Fresnel diffraction integral, Eq. 4.13 while the distances d_o and d_i are described by the thin lens formula Eq. 4.22.

I must digress and tell the reader that these formulae have been my friend and foe on COUNTLESS occasions. They always seem to end up solving my problems - who knew that mathematics is always right... Hence, when in doubt, even when you don't know what to cook for dinner, I strongly recommend the thin lens formula as your starting point!

2.4 Spatial resolution

We have now seen that, given coherent illumination, high frequency information, such as the edges of the square, will travel through space at a larger angle while the white middle part, which essentially consists of a DC component, will travel along the optical axis. This naturally leads to a discussion about the spatial resolution of an optical system.

Spatial resolution, in laymen's terms, is an optical system's ability to detect object detail. It can be defined in many ways, however here I will define it as the number of detectable

[13]The focal length is essentially a single value that fully describes the interplay between the curvature of the lens and its refractive index, i.e., it is a central component of the phase transformation function Eq. 4.19.

line-pairs per millimetre in units lp/mm. This definition, and in particular the way it can be measured in the lab, brings with it the advantage of a full characterisation of one's system within a snapshot. To break it down: a line-pair consists of a single dark and light line, e.g., a period of a sinusoidal function, and the idea is to quantify how many such line-pairs per millimetre at the object plane that are detectable by the optical system in question. A large value of detectable lp/mm means a higher spatial resolution and vice versa. To connect to our earlier discussions, more lp/mm at the object plane can be described by a higher f_0 of the corresponding sinusoidal, which in turn means that e.g., the edge of the square image will appear sharper. A higher lp/mm (or f_0) furthermore means that the given information will travel at a larger angle with respect to the optical axis. Therefore, if one were to place an aperture, such as a pinhole or a finitely sized lens, in the beam path, then high frequency information will be blocked from reaching the image plane. Features that include such frequency components will therefore be undetectable resulting in image blur, or lower spatial resolution (Fig. 4.6a).

Another aspect that this definition takes into account is the sampling rate of the camera in question. The camera can only detect a difference between a light and dark line if each part falls on individual pixels, i.e., at a rate of $1/2X$ or $1/2Y$ (Fig. 4.6b). Hence, the sampling rate of the camera will also act as a low pass filter resulting in images of, again, lower spatial resolution.

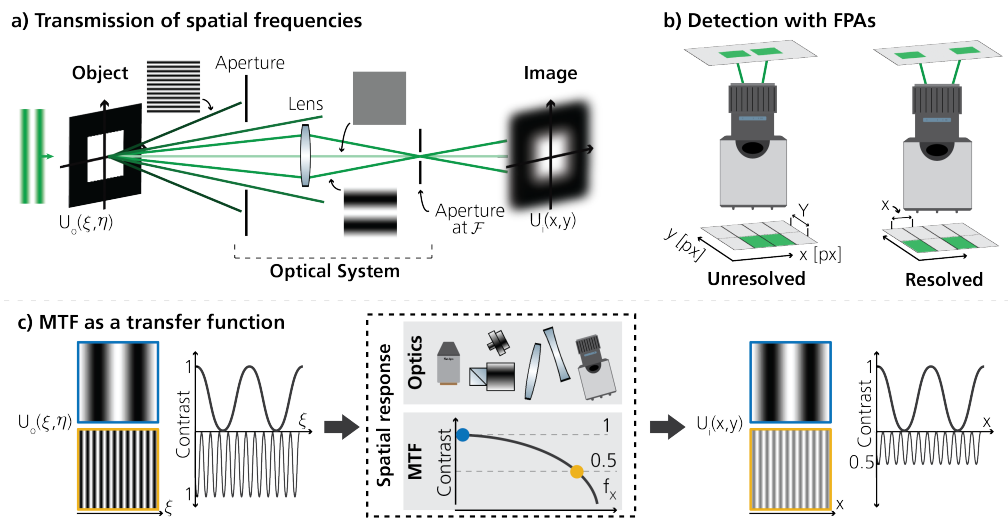
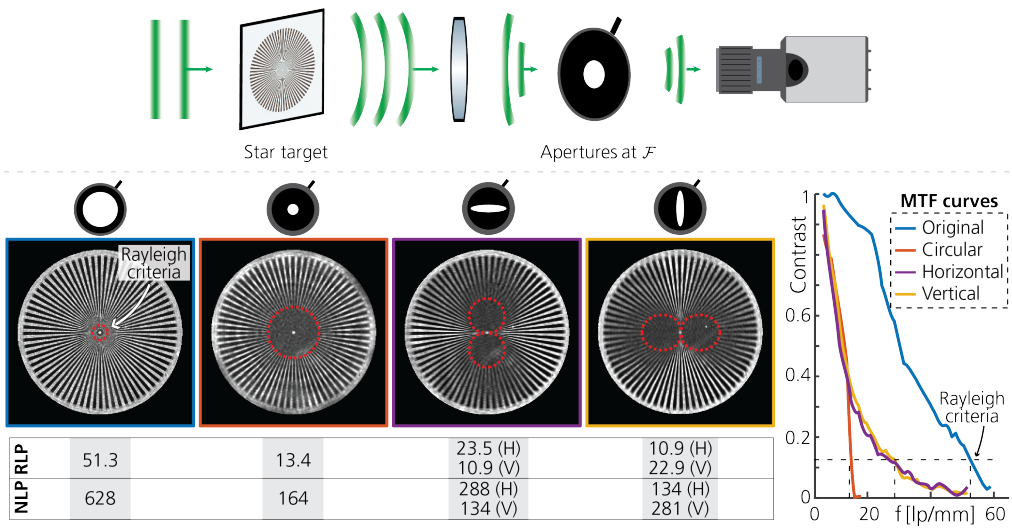


Figure 4.6: Spatial resolution of an optical system
a) As an image propagates through free space the various spatial frequency components can be blocked by the aperture limits of the optical system. When these block the high spatial frequencies, the resulting image becomes blurred. **b)** The finite pixel size of FPAs results in a resolution limit where two points can only be separated if they fall on two pixels separated by an empty one. **c)** Both considerations from **a** and **b** can be summarised into a single function: the Modulation Transfer Function (MTF) [100]. This curve (bottom centre box) functions as a summarised response (in units of transmitted contrast) of the optical system as a function of input frequency.

Quantifying the spatial resolution of an entire optical system can be performed via a single concept: the *modulation transfer function* (MTF). The MTF is a function that describes the ability of an optical system to transmit a given sinusoidal modulation frequency of contrast 1 in the object plane to the image plane [100]. For example, an MTF of 0.5 at a frequency of 10 lp/mm means that the measured contrast of the 10 lp/mm frequency component will be 0.5[14]. This concept can be boiled down to a response-based formalism, where one inputs a signal with spatial information, such as an image $U_o(\xi, \eta)$, and the output image $U_i(x, y)$ will be fully described by the MTF response function or how the optical system transmits the Fourier components, $G(f_x, f_y)$. Fig. 4.6c illustrates this where a low frequency component is transmitted through the optical system at contrast of 1 (in blue) and a higher frequency component is transmitted at a contrast of 0.5 (in yellow).

In order to measure the MTF of an optical system one can use a United States Air Force (USAF) calibration target or, as will be used for most figures herein, a Siemens sector star target. The sector star consists of 36 or 72 spokes of contrast 1 arranged in a wheel-like formation. This way, all spatial frequencies in all directions can be probed at once by fitting square functions at a given radius and recording their contrasts. This analysis results in an



NLP RLP	51.3	13.4	23.5 (H) 10.9 (V)	10.9 (H) 22.9 (V)
	628	164	288 (H) 134 (V)	134 (H) 281 (V)

Figure 4.7: Experimental spatial resolution of an optical system.
An experimental setup with an aperture in the Fourier plane shows how it affects the spatial resolution in the laboratory. By imaging through various apertures, various MTF curves of the given optical system are extracted, where the red dashed lines within the acquired images indicate the Rayleigh resolution. The values for the **RLP** are indicated in the MTF curve by dashed black lines and tabulated under their corresponding images alongside their **NLPs**.

[14]Note that, even though I have illustrated all these concepts in a binary way, the real transmission occurs over a continuum i.e., it is possible for a given frequency to be transmitted at any contrast between 0 and 1. This originates from the fact that I have concentrated my previous discussions in the context of the Fraunhofer approximation, i.e., each component travels along a well-defined narrow path, which is not completely true.

MTF curve, however usually one wants a single value for the spatial resolution cutoff of the system. Based on the Rayleigh criteria, the single value that we have decided to extract is the spatial frequency that is transmitted at a contrast of 15.7% and is dubbed resolvable line pairs per unit length (**RLP**). Fig. 4.7 includes a schematic and some laboratory images of how this is performed. Various apertures are placed in the Fourier domain to show the way different directional frequency components can be blocked resulting in different resolutions for different angles (in this case vertical, V, and horizontal, H). All articles that I have written include a measurement with a sector star target or USAF target and furthermore an in-depth analysis of various shaped apertures is included in Sam Taylor's master thesis [101].

In the context of ultrafast videography, the spatial resolution is not enough to assess imaging capabilities. Since the image needs to be taken in a snapshot, another characterising term should be included, namely the number of resolvable line pairs (**NLP**) across the field-of-view where the **RLP** is maintained (Fig. 4.7). Indeed, in the case of e.g., stimulated emission depletion microscopy (STED), the resolution can be incredibly high, however an image is built up of point measurements, hence such measurement techniques are irrelevant to ultrafast videography. Another aspect that is taken into account by the **NLP** is images that are stitched together from single acquisitions, such as google maps. Again, these are not relevant to videography as they are not acquired in a snapshot manner.

I could keep going with this really dry discussion forever, indeed I have been called a nerd on spatial resolution[15]. Therefore, I will end on a high-note: I have taken the liberty of finding and extracting a spatial resolution of the Google maps satellites over Arizona by using an MTF target that lies in the desert (Fig. 4.8). From this, we see that there is an RLP of about 2 lp/m and since I am not sure about the coverage of a single image, I have stuck to writing out how many resolvable line pairs are necessary to cover the entire earth if it were all taken in a snapshot. I hope this woke you up!

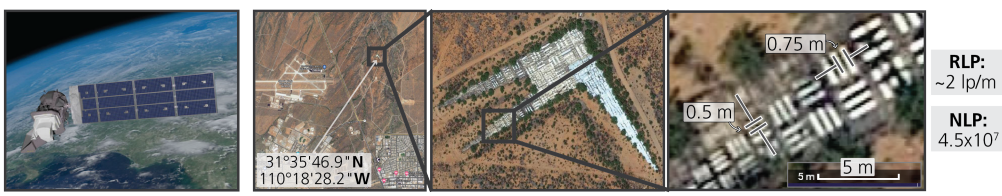


Figure 4.8: Google Maps' spatial resolution over Arizona
Upon zooming in on a Google Maps' satellite image of Fort Huachuca, Arizona, an old resolution target, built into the desert, can be found. By backtracking, i.e., using the scale bar, one can extract the resolution of the satellite which acquired the image to 2lp/m. Extrapolating to the area of the entire earth, the number of resolvable line pairs surpasses the 10^7 mark! The depicted satellite is an artist rendition of Landsat 9 which can image every part of earth every 16 days. Courtesy of NASA.

[15]Not going to point fingers, but you know who you are...

3 Natural Images

There is one last piece of the Fourier puzzle that is necessary to discuss, namely that of natural images. Imagine any 256×256 image. I am guessing just by saying image, you immediately imagine an image of a scene or an object. Try to widen your horizon and think of literally any combination of intensities on a 256×256-pixel matrix. The reader might have hit a roadblock, indeed even though I am the narrator of this thought experiment, I neither have a very wide horizon. Therefore, let us turn to the formalism of state-spaces. The state-space of images is built out of axes that represent the intensity values of each pixel. Hence it will be an N-dimensional space where N is the number of pixels within the image and where each pixel can take on any real intensity value (the $\mathcal{R}^N$ space of Fig. 4.9a)[16]. Every possible image consisting of N pixels will be represented as a point within this space, where its location (or vector) is its column stacked intensity values [102].

To narrow things down into human comprehension and give examples let us look at the subspace $\mathcal{W}^{256\times256}$ that consists of images with 256×256 pixels and whole number intensity values that are constrained within $I \in [0, 255]$, a case that is quite typical for experimentally recorded images (Fig. 4.9a). Three images, (i)-(iii), from within this set are depicted in Fig. 4.9b. The first one (in red) is simple white noise, the second one (in grey) is the cameraman image and the third (in green) is a checkerboard consisting of checkers with randomized intensities[17]. Even though all these images lie within $\mathcal{W}^{256\times256}$, they are not all relevant to videography. Indeed, in the context of experimental science and the monitoring of nature, the camera man would be the type of image we encounter most frequently. These types

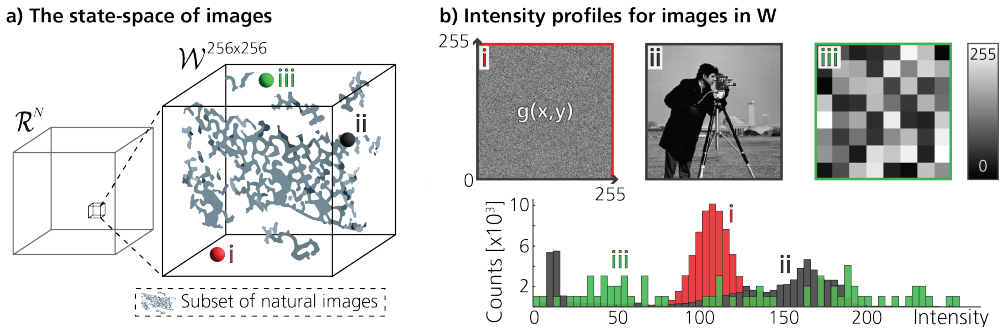


Figure 4.9: The state-space of images and their intensity profiles.
a) $\mathcal{W}^{256\times256}$, the vector space containing images of size 256×256, is an infinitesimally small subset of $\mathcal{R}^N$. Within $\mathcal{W}^{256\times256}$ there is a subset of images that we denote the natural images. Note that the spongy structure of this subset is used for illustrative purposes showing that it can have arbitrary complex shape. **b)** Three images within $\mathcal{W}^{256\times256}$ are depicted along with their intensity histograms illustrating significantly differing structures.

[16] This space includes negative values as well.

[17] I spent way too much time changing the seed in order to get a checkerboard of my liking.

of images are conveniently called natural images[18]. The aim of this section is to pinpoint what makes this subset/ensemble of images that we call natural images special. Does it for example have a specific shape within $\mathcal{W}$ (such as the spongy structure illustrated in Fig. 4.9) or can it be described by a certain statistical formalism? Once we have understood this, we will discuss how these properties can be utilised in ultrafast videography for the monitoring of natural events.

3.1 Intensity histograms and Fourier transforms

A first naive approach would be to look at the histograms of the intensity values of the images in Fig. 4.9b ($\text{HIST}\left[g(x,y)\right]$). As expected they differ widely and maybe so-called natural images are simply built of a predictable intensity distribution. In order to understand that this is not the case, three images from subspace $\mathcal{W}$ with identical intensity profiles (in red) are shown in Fig. 4.10a. The first one is obviously a natural image of flamingos (note that its histogram differs from that of the cameraman), while the other two are randomly reshuffled pixel values and sorted pixel values from the flamingo image. The identical intensity histograms show that these do not characterise the ensemble of natural images. Therefore, we turn to the Fourier domain of these images.

The Fourier transform of these three images show three distinct structures (Fig. 4.10a, bottom panel): the original image is characterized by a diamond shape, while the randomly reshuffled and sorted images are characterized by a flat and cross shape respectively. It

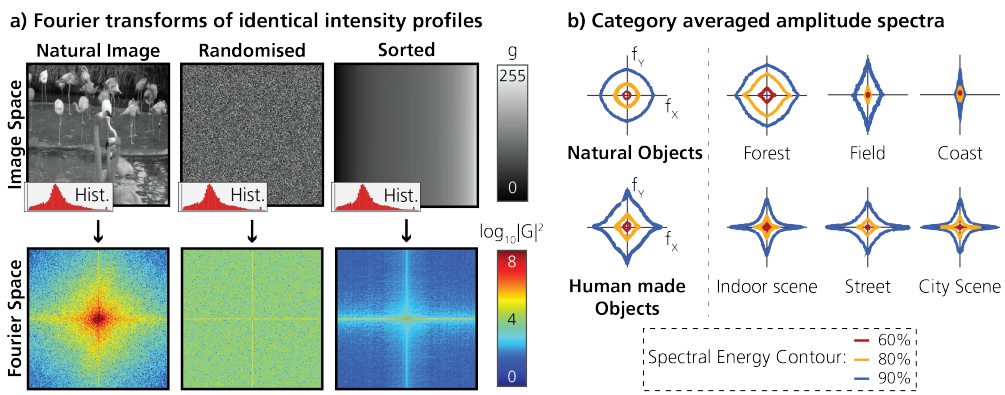


Figure 4.10: Fourier spectra of natural images.
a) The Fourier transforms of three images with identical intensity distributions: a so-called natural image, a randomised image and a sorted image, depict differing structures. **b)** Natural images (includes natural objects and human made objects) have a specific diamond like structure to their Fourier coefficients, illustrated by the coloured boundaries where most of the image power is confined. The higher kurtosis of human made objects originates in the existence of more edge like structures necessitating higher frequency components to create. This figure is adapted from [103]

[18]For the reader that wants to get to the point, unfortunately I am taking the route of intuition much like the pioneering work of neurologists that try to understand how the brain responds and encodes various scenes.

turns out that natural images can in fact, among other things, be characterized by their Fourier spectra, G. This fact has been thoroughly investigated by the computer vision and image compression community [104–107]. In particular, Fig. 4.10b shows a figure adapted from [103], where the average Fourier spectrum from a few hundred images from each labelled category are drawn. The contours, which represent 60%, 80% and 90% of the spectral energy signature, have a characteristic diamond shape that we will come to expect from the natural images. Furthermore, even though they differ in their e.g., kurtosis and directionality this shape seems to hold for all natural image categories. This fact fundamentally originates in natural images being *scale invariant* to a certain degree.

3.2 Scale invariance

Scale invariant features of objects are such that they do not change upon scaling of length, energy or other attributes but are multiplied by a common factor. In mathematical terms, for a function f to be scale invariant under the rescaling of its variable x it must satisfy:

$$f(\lambda x) = \lambda^\gamma f(x), \tag{4.24}$$

for scaling factor λ and to a degree γ ($\gamma > 0$ within the scope of this thesis). An example of scale invariant functions is the set of monomials $f(x) = x^n$ which are scale invariant to the n:th degree. Indeed:

$$f(\lambda x) = (\lambda x)^n = \lambda^n x^n = \lambda^n f(x). \tag{4.25}$$

This concept, which is oftentimes interchanged with the term self-similarity, widely exists within the field of physics from statistical mechanics to particle physics and is most often visualized by fractals. Within the field of image analysis and computer vision it is most often used within feature detection algorithms such as the scale invariant feature transform, SIFT, and its applications in e.g., 3D reconstruction [108]. Here I will give an intuitive understanding of how it applies to natural images and their Fourier spectra.

Consider a fixed imaging system consisting of e.g., a macroscopic objective lens and a camera taking an image of a leaf (Fig. 4.11a). As one zooms in and out, the image of the leaf remains identical except for its scale. If we look at e.g., the mean value of the image in the vertical direction, one can see from the plots that the zooming only affects the scaling of the shape[19]. In contrast, as one scales non-natural images (Fig. 4.11b), the profiles do not scale by a constant λ^γ, but rather stay identical. They are in fact 0:th order scale invariant, i.e., $\gamma = 0$ in Eq. 4.24. This might seem obvious to the reader; however, it has far reaching consequences as will be seen in the next section.

[19] This is further corroborated by the applied fits, which are simply scaled versions of each other.

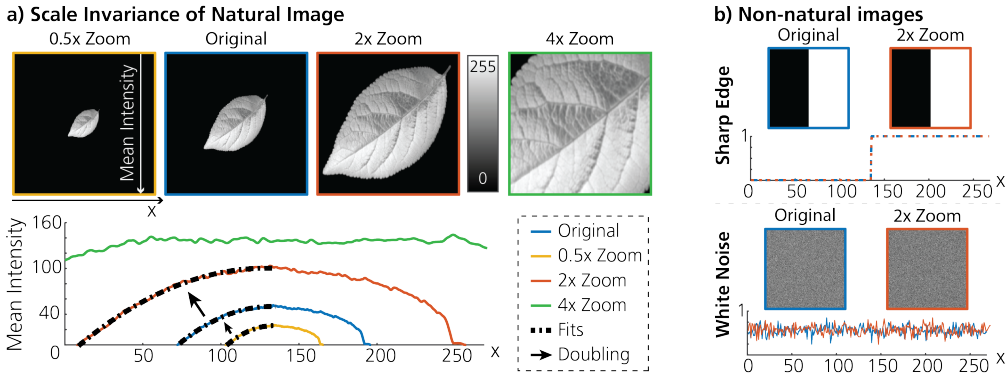


Figure 4.11: Behaviour of images at different scales.
a) Natural images scale accordingly when zooming in and out as seen by the horizontal intensity profiles of the leaf. **b)** Non-natural images such as white noise and sharp edges do not scale accordingly, i.e., their structure does not change as one zooms in or out. They can be said to be scale invariant to the zeroth degree.

Of course this is only completely true to a certain degree, as is exemplified by zooming in such that the entire leaf is no longer visible (green curve of Fig. 4.11a). This is the reason why we constrain this thought experiment to a single objective lens (with zoom constraints) and camera (with binning constraints)[20]. If we were to change the system to a microscope, we would see leaf cells which look completely different compared to the macroscopic case. However, within the constraints of the microscope a similar scale invariance will hold nonetheless.

The effect of scale invariance on the Fourier representation

Scale invariance has a far reaching consequence in terms of the amplitude spectrum, $G(f_x, f_y)$, of natural images [104]. If an image of the flamingos has the energy E within the frequency band $[f_0, nf_0]$, then, due to scale invariance, when zooming by a factor a, the energy E will be shifted to the frequency band $[af_0, anf_0]$. The trick is to understand that the energy within that spatial content of the image must stay constant irrespective of this shifting operation[21]. Formalising this concept: the energy at a given frequency, is proportional to the complex amplitude squared multiplied by the frequency (performed in one dimension for simplicity):

$$E(f) \propto |G|^2(f)f. \tag{4.26}$$

[20] This constraint will be less harsh if we work with images such as the cameraman or flamingos as they contain so much detail that the profiles will average out over larger regions within the zooms.

[21] Staying constant is obviously theoretical as pixel pitch, discrete sampling of a signal, quantum efficiency etc. will play a role.

Hence the energy within a small frequency band $[f_0, \epsilon f_0]$ (small band such that $|G|^2(f) \approx |G|^2$ is constant) is proportional to:

$$E \propto \int_{f_0}^{\epsilon f_0} |G|^2 df. \tag{4.27}$$

Upon zooming by a factor a, the band of interest will shift by said factor while the energy within must remain constant:

$$E \propto \int_{f_0}^{\epsilon f_0} |G|^2 df = \int_{af_0}^{a\epsilon f_0} |G|^2 df = A, \tag{4.28}$$

for a constant A. Performing this integral results in:

$$|G|^2 f = A \;\Rightarrow\; |G|^2 = \frac{A}{f}. \tag{4.29}$$

Hence the Fourier spectrum of scale invariant images is characterized by a $1/f$ dependence! Now we will turn the argument back on itself as we will look at various natural images and see that this is in fact the case.

3.3 1/f Fourier spectrum dependence of natural images

Fig. 4.12a presents the flamingo image along with its logged Fourier transform, $\log_{10}(|G|^2)$. G is then extracted for each frequency along the marked directions: horizontally, vertically, diagonally and in circular bands (denoted doughnuts, where the mean amplitude within the band-pass is extracted). They are then plotted in a log-log plot as a function of their spatial frequency. Performing fits, we indeed see that: $|G|^2 \propto f^{-(1+\eta)}$ for a small value of η. Furthermore, as we assumed at the beginning of the derivation, the inset shows how the energy, $|G|^2 f$, stays constant over the frequency bands. This behaviour remains even as the camera settings are varied. Fig. 4.12b illustrates this by zooming and binning the flamingo image. Plotting then fitting via doughnut integration (as in Fig. 4.12a), results in the equivalent $f^{-(1+\eta)}$ structure that we have come to expect for natural images.

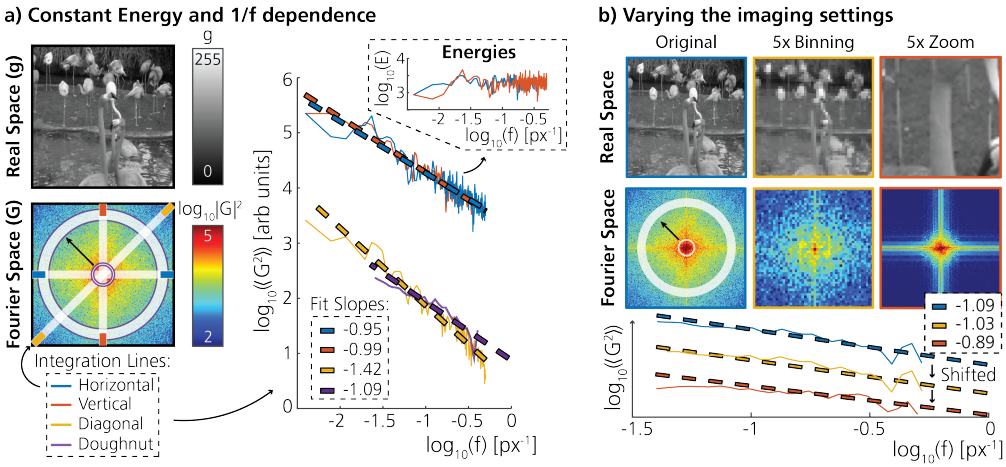


Figure 4.12: 1/f amplitude spectrum of natural images
a) Plotting the values of the Fourier coefficients as a function of their coordinates (i.e., their corresponding frequency) results in the derived $1/f$ structure. Note how the energy within the frequency bands does indeed stay constant as plotted in the inset. **b)** Varying the camera settings by zooming and binning does not drastically change this amplitude structure, further corroborating the image's scale invariance.

Concluding remarks on Fourier transforms and images

In this chapter we have seen that the subset of images that we are interested in measuring, the natural images, are characterised by an $f^{-(1+\eta)}$ structure of their Fourier spectrum (this is a general result where e.g., [109] shows this thoroughly). We have also seen that even though various types of objects and scenes have different amplitude structures, they are all characterised by a diamond-like shape around their central frequency. In conjunction, these two facts show that, given a high sampling rate, there are always areas within reciprocal space of low power, especially in the outskirts which can then be targeted for the storage of image information. In the next chapter we will use our newfound knowledge of image propagation through space, Fourier transforms of natural images and the introduced theorems in order to target that void within the Fourier domain as a means of storing image information for the purpose of ultrafast videography.

Chapter 5

Storing Images in the Fourier Domain

At this point, you the reader, have had to endure my rambling about three separate pieces of the puzzle. First we looked into the fact that photons carry multiple dimensions of information and that the job of optical diagnostic techniques is to perform operations within photon space (**C**), real measured space (**P**) and in between (**M**). Then we looked into the current state of technology, in particular the extreme developments in FPAs and pixel size. Furthermore, we looked into how various active and passive videography techniques differ in the way they categorise photons with respect to their temporal information. Finally, we discussed and looked into the formalism of the Fourier transform in both theory and in the laboratory, where an extra weight was put on image formation. These three concepts are key to really obtain a full understanding of how structured light can be utilised in order to achieve ultrafast videography and unlock some of the mysteries of nature.

In this chapter, I would like to offer the reader, and maybe a future spatial frequency division based videographer, some discussion points and terminology that need to be regarded when working within this niche field. The findings within this chapter are not entirely complete, however the aim is that they should provide jumping boards to bounce off of when applying structured light to ultrafast videography. With that, let us dive head first into putting all our newfound knowledge together into the self-contained ultrafast videography technique that is FRAME.

1 Manipulating Image Frequency Coordinates

Consider a natural image, $g(x,y)$. If we multiply it by a spatial intensity modulation, $M = 1/2\big[1+\cos(\mathbf{f}_0\cdot\mathbf{r})\big]$, of frequency $\mathbf{f}_0 = (f_{0\mathrm{X}}, f_{0\mathrm{Y}})$, the resulting image, g^{M}, will have the form:

$$g^{\mathrm{M}}(x,y) = g(x,y)M(x,y) = g(x,y)\left[\frac{1}{2}+\frac{1}{2}\cos\left[2\pi(f_{0\mathrm{X}}x+f_{0\mathrm{Y}}y)\right]\right]. \tag{5.1}$$

From the convolution theorem, we know that a multiplication in real space corresponds to a convolution in Fourier space, hence:

$$G^{\mathrm{M}}(f_{\mathrm{X}},f_{\mathrm{Y}}) = \frac{A}{2}G(f_{\mathrm{X}},f_{\mathrm{Y}}) + \frac{A}{4}G(f_{\mathrm{X}}+f_{0\mathrm{X}},f_{\mathrm{Y}}+f_{0\mathrm{Y}}) + \frac{A}{4}G(f_{\mathrm{X}}-f_{0\mathrm{X}},f_{\mathrm{Y}}-f_{0\mathrm{Y}}), \tag{5.2}$$

for a constant A. From this last equation, we see that, after multiplication by M, three weighted copies of G have been formed at positions $(0,0)$ and $\pm(f_{0\mathrm{X}},f_{0\mathrm{Y}})$ in the Fourier domain. This is a strong result, and the basis of FRAME, since, via this simple multiplication in image space, we have been able to shift image information to a unique position in Fourier space, separated from the central DC component. Since the extent of G is limited by a $1/f$ diamond like structure[1], by choosing $\mathbf{f}_0$ such that it falls within the spatial resolution of the camera, but is significantly larger than Fourier extent of the image under question, one can store g in the Fourier domain with minimal interference from its DC component!

This is quite simple to perform in the lab. Fig. 5.1a shows a star target (it is not a natural image however the discussion still holds) that has been illuminated with modulated light (see the zoom-ins of (i)). Upon performing the Fourier transform of this recorded image, (ii), we indeed find three Fourier clusters at the expected positions.

From the inverse theorem of the Fourier transform and the fact that the Fourier transform gives a full representation of the image (it forms a complete orthonormal basis), it turns out that it is possible to isolate the image information stored at $\pm(f_{0\mathrm{X}},f_{0\mathrm{Y}})$ i.e., $2\times(A/4)G$, and fully reconstruct g to a multiplicative factor. This can be achieved via a so-called *lock-in algorithm*. By multiplying g^{M} with $\exp\left[-i2\pi(f_{0\mathrm{Y}}x+f_{0\mathrm{Y}}y)\right]$[2] one can undo the initially applied modulation, M, and shift the entire manifold, G^{M}, such that the information which was initially situated at $\pm(f_{0\mathrm{X}},f_{0\mathrm{Y}})$ ends up in the DC component, (iii). Applying a low-pass filter (LPF) to this demodulated image results in the isolation of the information

[1] Under the assumption that the camera in use is modern enough. In the same way as physical existence is enough for a Fourier transform to exist, in this day and age a modern camera is a sufficient requirement for the limited extent of natural images. This is given that we are not using SPAD arrays or other exotic technology.

[2] This is the elementary function of the applied modulation, i.e., it has the same frequency as the one applied in the lab.

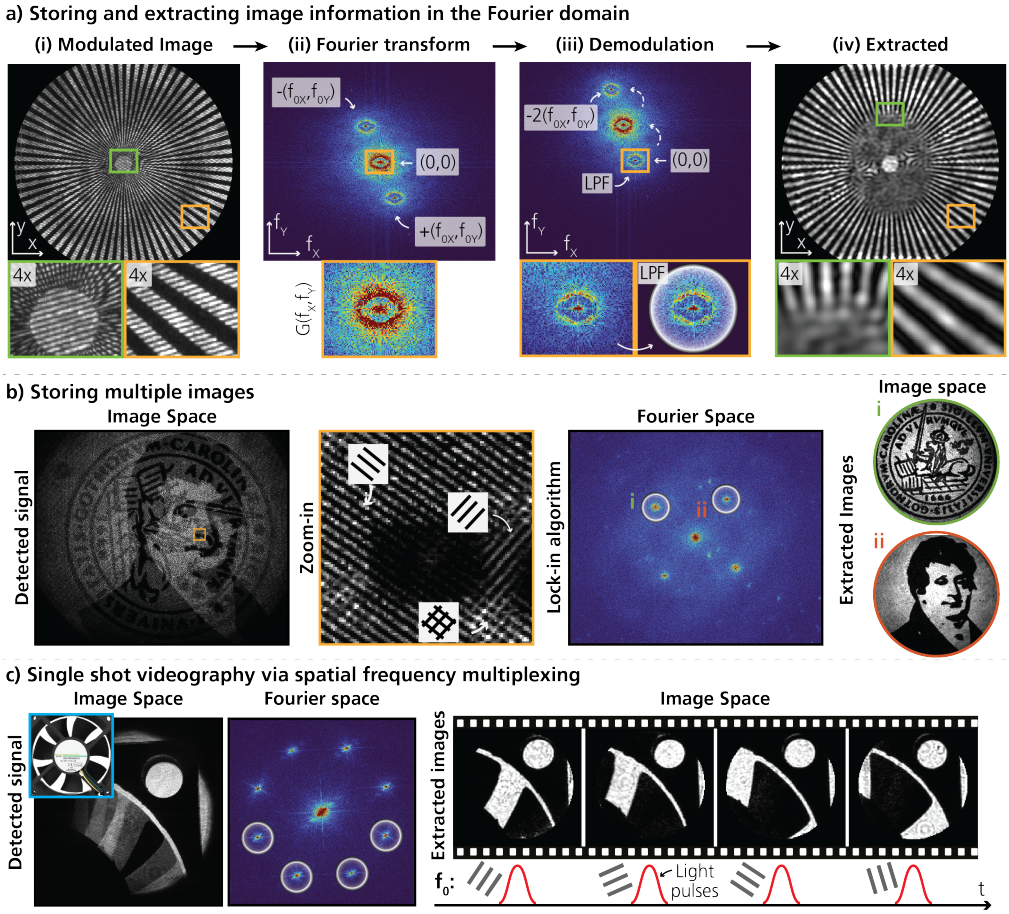


Figure 5.1: Storing and extracting modulated image information.
a) A sector star target, illuminated by modulated light, is recorded on a camera, (i). Performing the Fourier transform reveals Fourier clusters at the frequency positions of the applied modulations along with a higher intensity DC component, (ii). Multiplying by a complex exponential with identical frequency, demodulates the modulated signal, i.e., shifts the entire Fourier manifold such that the signal of interest is moved to the DC position, (iii). Applying a low-pass filter (LPF), isolates the signal of interest, thus, performing an inverse Fourier transform of the low-pass filtered, demodulated signal results in the extracted image information, (iv). Note the decrease in spatial resolution. **b)** Multiple images can be stored and extracted via the procedure of **a** with the caveat that the algorithm must be applied separately to each cluster. In this case two separate images, one of the Lund University logo and the other of Jean-Baptiste Fourier, are stored on and extracted from a single camera exposure. **c)** By illuminating a moving object, in this case a rotating computer fan, with spatially modulated light at different times and performing the procedures described in **a** and **b**, a video of said object can be extracted. This constitutes the essence of single shot videography via spatial frequency multiplexing, i.e., the basis of FRAME.

in the DC component, i.e., $2 \times (A/4)G$ (iii)[3]. Subsequently, by performing an inverse Fourier transform, g can be reconstructed to a multiplicative factor (iv). With the risk of repeating myself, note the lack of modulation in the extracted image, which is a direct consequence of having shifted the information located at $\pm(f_{0x}, f_{0y})$ into the DC position

[3]Note that (iii) is slightly misleading since the Fourier clusters have only been shifted upwards. This is for illustrative purposes.

during the demodulation step. A more in-depth mathematical description of the lock-in algorithm is included in Appendix A.

By repeating the entire aforementioned method, from modulating image information in the lab to performing the lock-in algorithm, multiple images can in fact be stored on a single FPA within a single exposure. In practice, this involves applying separate modulations, M_i, to separate images, $g_i(x, y)$, such that copies are placed at various locations within the Fourier domain. By then performing the lock-in algorithm on each set of $\mathbf{f}_i = (f_{ix}, f_{iY})$, it is possible to isolate and extract the image information from each unique frequency location. Fig. 5.1b illustrates this for two separate images, the Lund University logo and Mr. Fourier, while Fig. 5.1c illustrates this concept for temporally separated illumination such that upon extraction, a video of a rotating fan is revealed. The technical term for this strategy of storing of information in frequency space is *frequency multiplexing*. These ideas have existed for a long time within electronic communication technologies, however Weinstein et. al really applied the ideas in the realm of DFTs in 1971, concepts that are strongly relevant to this thesis [99]. Frequency multiplexing constitutes the basis of how AM radios work[4]. Following this analogy, if Radio Station Vasses Mixtape is at 550 kHz ($f_{0t} = 550$ kHz) and you want to listen to it, you resort to turning the seeking knob until you start hearing my bangers. This act of turning the knob is you forcing the electronics to demodulate a given frequency. The electronics contained within the radio then automatically low-pass filter around said frequency and the amplitude information embedded within gets passed on to the speakers. In essence, tuning the radio is you performing a one dimensional, time-based lock-in algorithm! So, next time you listen to the radio, you can flex and say this to the person sitting next to you in the car - you will definitely sound so smart.

2 Spatial Resolution vs Specificity

Applying a low-pass filter is central to the lock-in algorithm in order to isolate the modulated signal of interest. However, akin to placing an aperture in the optical Fourier domain, low-pass filtering in the post-processing step results in lower spatial resolutions (compare Fig. 5.1a (i) with (iv)). Hence there is always a trade-off between spatial resolution and *specificity*, i.e., how well the filter isolates the signal of interest from other modulated signals or the DC component. The type of filter and its corresponding size are critical variables when approaching this trade-off [110]. Fig. 5.2a illustrates this by applying variously shaped filters with different radii to the Fourier domain of a star target and extracting MTF curves. As expected, the application of filters with larger radii result in higher spatial resolutions. Less obvious is how the filter shape affects the MTF: the hexagonal and super-Gaussian

[4]FM radios also work in a similar manner however they are not exactly the same as one extracts the frequency shifts while in AM one extracts the amplitude envelope - more akin to what we do here.

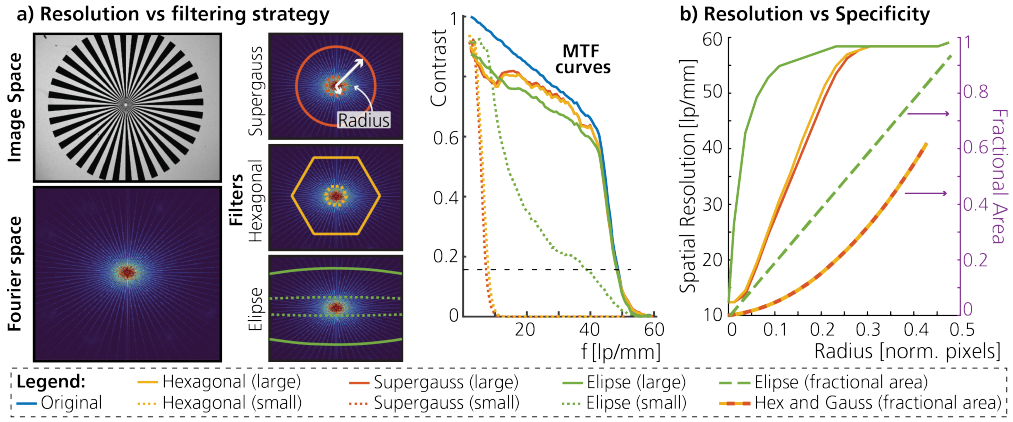


Figure 5.2: Low-pass filtering strategies and spatial resolution.
a) Applying filters of different shapes and sizes affects the extracted MTF in various ways, where, as one would expect, larger radii results in MTFs that approach the original. For the ellipse, the horizontal resolution stays constant while the vertical resolution changes with the size of the semi-minor axis. This is summarised by the MTF's irregular shape compared to those of the hexagonal and Gaussian shaped filters. **c)** Plotting the Rayleigh spatial resolution as a function of filter radius results in an expected positive correlation between the two. The fractional area of the given filters also increases with radius illustrating the trade-off between specificity and spatial resolution.

filters deliver similar results while the ellipse filter MTF follows an exponential type structure attributed to the fact that the horizontal resolution stays constant while the vertical resolution varies with semi-minor axis[5]. Quantifying this trade-off involves extracting the Rayleigh resolution as a function of normalised radius[6] (Fig. 5.2b). On the right axis, the fractional area that the filter covers in the Fourier domain is also plotted, illustrating the filters' specificity (a larger fractional area leaves less space for other unique images to be stored without interference). The way both these functions grow with radius clearly demonstrates the resolution vs specificity trade-off, unfortunately a fact of life within the world of multiplexed signals.

2.1 Choosing a filter size

So how does one choose a radius that is specific enough but simultaneously maximises the attainable spatial resolution? In theory, one could resort to a Nyquist-based approach, however oftentimes when multiplexing in practice, this radius lacks specificity. This section attempts to give the reader a "recipe" to be used upon application of a lock-in algorithm to natural images[7]. A rotating fan was illuminated four times by four temporally separated,

[5]This type of filter is paramount to the application of a technique called Periodic Shadowing on spectrometers and streak cameras, a technique applied in Papers v and vi

[6]The normalised radius is defined as the filter radius divided by half the size of the image or the maximum radius.

[7]As with food recipes online I must write about how the first time I tasted this I was taken aback by its simplicity while it was still so balanced. The year was 1734...

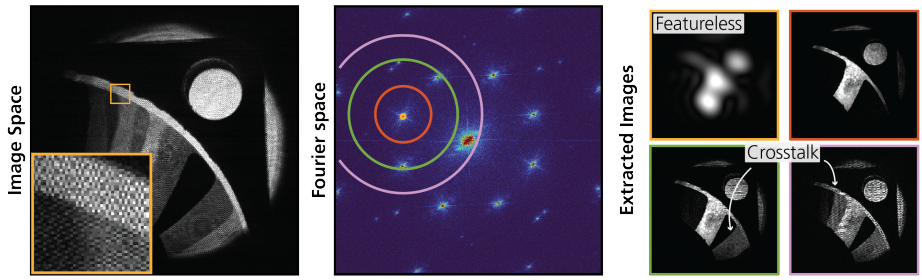


Figure 5.3: Varying the filter radius.
An extracted signal is sensitive to filter radius and results in images that suffer from featurelessness and blurriness for small filters. By applying larger and larger radii, the information in neighbouring clusters will start to fall within one's low-pass filtered image, resulting in a decrease in contrast and an appearance of modulated structures, an effect termed crosstalk.

uniquely modulated laser pulses. The resulting Fourier domain therefore consists of four Fourier clusters along with a DC component and some harmonics (Fig. 5.3). By applying filters of various sizes to the leftmost cluster, we see that at small radii the image is close to featureless or extremely blurred, while at larger radii information from other clusters become visible as interfering modulations and decreased contrast - effects that fall under the umbrella term of *crosstalk*.

Frequency cutoff according to noise floor

In order to choose a radius one can, as in the previous chapter, plot the intensities of the Fourier cluster of interest as a function of frequency. This was done along a horizontal integration path for the central (in blue), left- (yellow) and rightmost (green) Fourier clusters in Fig. 5.4a. One can notice how the plotted functions consist of two parts: (1) the characteristic $1/f$ structure that we have come to expect from natural images and (2) a flat noise floor, whose level and standard deviation is plotted as a shaded box. A choice of filter cutoff frequency would be at the point where the fitted $1/f$ line crosses the noise floor. This ensures that only information with SNR ≥ 1 is included in the extraction, omitting unnecessary power in the form of noise. The result is three different radii, where, as expected, higher SNR in the signal of interest results in larger cutoff frequencies.

Frequency cutoff according to neighbouring clusters

Another consideration to be had, is the interference of adjacent Fourier clusters. To illustrate this, Fig. 5.4b plots the intensities of the rightmost Fourier cluster (green of Fig. 5.4a) along a vertical integration line, along which there lies a neighbouring cluster. In this specific case the neighbour does not interfere with previous filter choices, however if it were closer, as illustrated by the left-shifted dark green curve, the cutoff frequency of choice

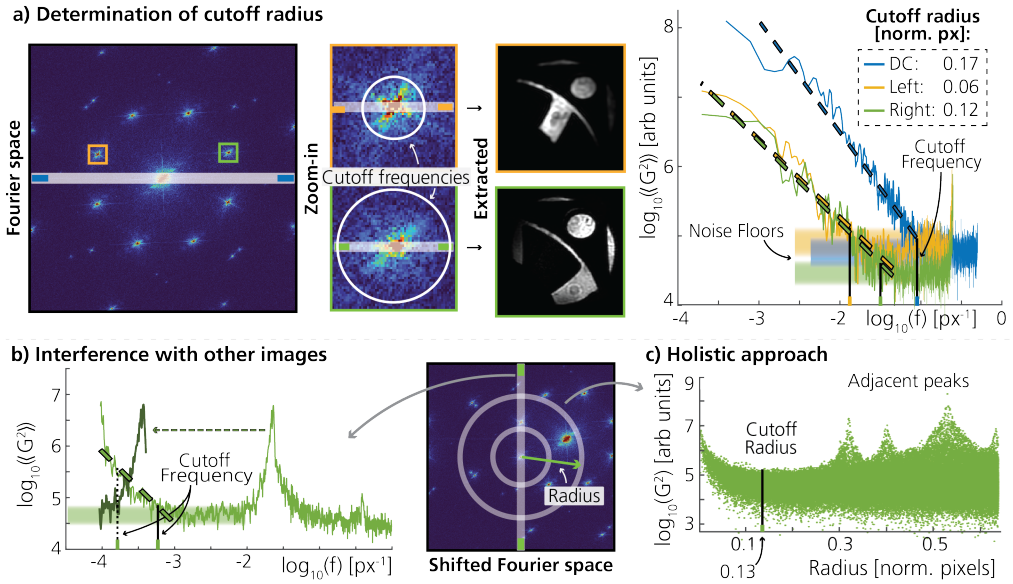


Figure 5.4: Choosing an appropriate filter radius.
a) Plotting the horizontally integrated Fourier intensity as a function of frequency we note the $1/f$ structure along with a noise floor for the various clusters. An appropriate cutoff frequency for ones LPF is where the signal crosses said noise floor. **b)** When a neighbouring cluster falls within the cutoff frequency defined in **a** (dark green curve), one can instead resort to a cutoff frequency where a lack of cross talk is ensured, that is when the neighbouring cluster falls under the noise floor. **c)** In practice (especially in the lab) one can choose a radius in a more holistic way. Here, the Fourier intensities are plotted as a function of their radius with respect to the rightmost cluster (in the green box of **a**). A cutoff frequency is chosen somewhere in between the peak of interest and the closest neighbour at 0.13 normalised pixels.

would be where the neighbouring signal falls under the noise floor, effectively guaranteeing no cross talk.

Holistic approach

In the majority of cases, especially when working in the lab, one can approach this choice of radius in a more holistic manner. Fig. 5.4c plots the logged intensity of the pixels in the Fourier domain as a function of their radial position with respect to the rightmost Fourier cluster (again, the green cluster of Fig. 5.4a). The choice of radius that I would go for would be where there is a flat structure in between the peak of interest and the next one. In this case, this analysis would yield a radius of ≈ 0.13 normalised pixel units, quite close to the more rigorous approaches but within three lines of simple code.

3 Discretising Spatial Modulations

We have seen, from Eq. 5.2, that upon the application of a sinusoidal spatial intensity modulation, half the image energy will remain within the DC component while the other half will be shifted into the modulated components. This is however only completely true for continuous signals. Upon detection with a camera, a sinusoidal spatial intensity profile will be discretised at a rate that can be expressed in number of pixels per period (or pixels per line-pair). Upon discretisation the sinusoidal will morph into a more square type shape as less pixels are used to represent a single period (Fig. 5.5a for a 1D case). This deformation in the spatial domain is evident in the Fourier domain, where the strength of the modulated component (its first harmonic) decreases, and its energy is redistributed to other, higher harmonics.

Investigating the effect this has on the image energy fraction that is transferred into the modulated component involves performing a simulation at varying discretisation rates and extracting the Fourier amplitude of the first modulated component (Fig. 5.5b). From this simulation we see that the ratio between the DC and modulated component begins at the theoretical 50% for a near continuously sampled signal, and decreases fast at a discretisation rate of 10 pixels per period (right purple axis and purple curves). This can be detrimental to FRAME, as one wants to place image information far out in the Fourier domain in order to gain as much spatial resolution as possible. The resulting trade-off, however, is a lower image energy which in turn affects the dynamic range of said stored image.

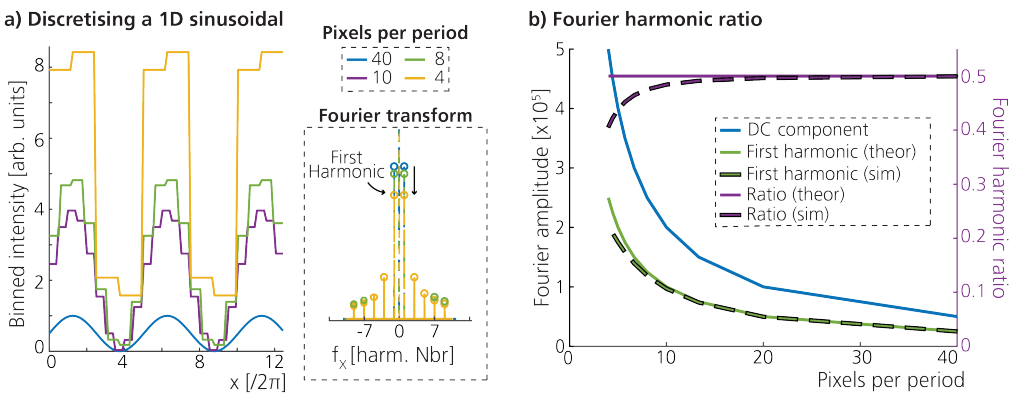


Figure 5.5: The effect of discretising a sinusoidal on Fourier energy distribution.
a) As a sinusoidal intensity function is discretised by the detection system it gets deformed towards a square-like structure. This results in a redistribution of energy in the Fourier domain, where higher harmonics appear at the cost of lowered first harmonic energy. **b)** Upon modulation with a sine function, half the image information should end up in the first harmonic (Eq. 5.2). However, as the function is discretised, said ratio between first harmonic and DC component energy will decrease fast from the continuous value of 0.5 (purple curves with respect to the right axis).

4 Storage Efficiency of Image Multiplexing

Is it possible for a spatial frequency division technique to reach the empty high-sequence depth areas on Fig. 3.4? Indeed, CUP, a compressive sensing approach has been able to push upwards in said plot, and spatial frequency multiplexing is just a type of image compression (compare JPEG or MPEG compression).

In this section I will discuss the storage efficiency of image information when using the spatial frequency multiplexing strategy which itself if a continuation of the analysis performed in Paper II. The analysis will concentrate on comparing this approach with the spatial division approach for an intuitive understanding of what is meant by the various parameters. In order to perform these measurements a Michelson interferometer with acousto-optic deflectors (AODs) in each arm is used to redirect a coherent laser beam such that a pulse train with unique spatial modulations ($\mathbf{f}_0$) is incident on the object plane (for the optical setup, see Appendix D). The camera in question is a 16 MP Imperx Bobcat B4822 with a pixel pitch (size) of 7.4×7.4 μm^2 and a 12-bit dynamic range, while the quantification of the storage efficiency is performed via the analysis of MTFs and dynamic ranges. At this point, I must reiterate my disclaimer: the analysis provided herein is by no means complete, nor is it set in stone. The aim is to point to certain trends and introduce terminology for further discussions about the frequency division approach, an approach that is central to the functioning of FRAME.

4.1 Packing strategy

The first step to attaining the high-sequence depth goal is finding a way to maximise the utilisation of the Fourier domain, where the interplay between the choice of spatial modulation frequencies and the corresponding filters is central. Multiple different combinations of said parameters have been tested (Fig. 5.6) and, as has been proven by bees, a hexagonal structure in the Fourier domain in conjunction with hexagonal low-pass filters results in the most efficient use of spatial bandwidth, or in more technical terms, the highest packing factor. There is a detail that needs to be addressed within the context of spatial frequency multiplexing approaches and that is the extent of the central DC component. Due to its existence, a packing factor of unity is unreachable. In fact, when applying this technique in the laboratory, it is necessary to measure the extent and shape of the DC component before choosing positions and filters that minimise the interference with said DC component.

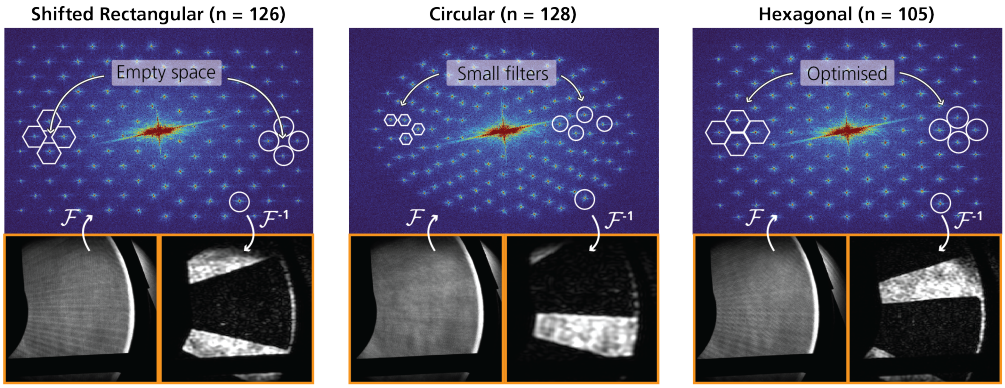


Figure 5.6: Packing strategies for spatial frequency multiplexing.
Images of a rotating fan illuminated >100 times by unique modulations are shown in the bottom left corners. Upon performing the Fourier transform, various structures are revealed: a rectangular, a circular and a hexagonal distribution. In conjunction with hexagonal filters, the most efficient packing strategy is, as expected from bees, the hexagonal structure.

4.2 Spatial efficiency

For a video to be of adequate quality in order to visualise natural events while simultaneously pushing the limits to sequence depth, an analysis of the stored image information as a function of sequence depth needs to be performed. The analysis will be compared to the more intuitive spatial division approach which itself is similar to frequency multiplexing in that it stores separated image information on a single chip exposure.

Spatial resolution as a function of sequence depth

As mentioned previously, in the case of spatial division each image is directed to a different area on the sensor, hence the number of pixels necessary for each individual image, N_{IMG}, is given by N/n for N the total number of pixels on the FPA and n the sequence depth. We assume ideal conditions with a packing factor, or fill factor, of unity, meaning that all images are perfectly distributed and every pixel contributes to the stored images. From this it can be shown that the linear spatial resolution of the spatial division strategy decreases by $1/\sqrt{n}$.

In order to investigate this in the context of spatial *frequency* division, images of star targets were acquired under illumination of pulse trains of varying lengths. The spatial modulations were chosen with a hexagonal packing structure approach such that hexagonal filters, with radii half the distance to the next Fourier cluster, are used for a maximised fill factor (Fig. 5.7a). MTF curves were then extracted and average values of the linear spatial resolution, or the Rayleigh resolution, are plotted against sequence depth (Fig. 5.7b). Here, the red dashed curve corresponds to the $1/\sqrt{n}$ resolution of a spatial division approach,

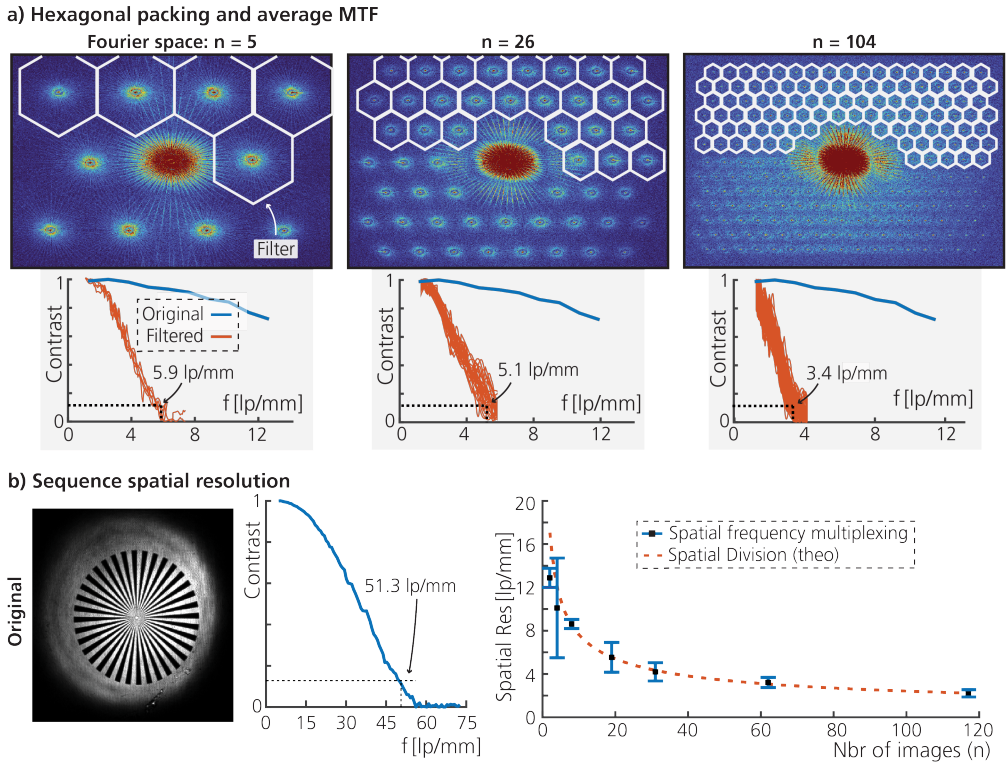


Figure 5.7: Spatial resolution as a function of sequence depth for the spatial frequency multiplexing strategy
a) Filling the Fourier domain with image information in a hexagonal packing structure results in lower spatial resolutions as a function of sequence depth, n. The extracted values are 5.9, 5.1 and 3.4 lp/mm (as compared to the original 51.3 lp/mm) for sequence depths of 5, 26 and 104 respectively. **b)** The Rayleigh spatial resolution of the extracted images as a function of sequence depth for the spatial frequency multiplexing strategy results in a $1/\sqrt{n}$ structure, akin to spatial division where the theoretical curve has been normalised to the original resolution.

normalised to the original resolution, while the black points are the average measured spatial resolutions of the frequency division approach with error bars corresponding to the standard deviation over the sequence in question. We see that, similarly to the spatial division strategy, in terms of spatial resolution the $1/\sqrt{n}$ trend is able to describe frequency multiplexed videos in an adequate manner.

Effect of oversampling

When performing the lock-in algorithm, one extracts sub-images that are of the same size (in terms of pixel number) as the original image. Naively, this would mean that the number of pixels that are a part of an entire sequence N_{SEQ} is equal to $n{\times}N$ (for n the sequence depth and N the number of pixels on the sensor). In contrast, the corresponding value for a spatial division approach would be N. This can be interpreted as the frequency division approach

being n times more efficient at storing data compared to the spatial division approach. However, due to the lower spatial resolution after low-pass filtering, in conjunction with the Nyquist sampling rate, the extracted images from a spatial frequency multiplexing method will be highly oversampled, hence this increased storage efficiency interpretation is simply not true.

We can define a coefficient of oversampling, α, used in order to downsample extracted images, such that an estimate of the number of *effective* pixels can be attained. For a star target of width 10 mm (this is the case for the images of Fig. 5.8a) the *minimum* number of pixels necessary to image it at a resolution of R (in lp/mm) would be:

$$N_{\text{min}} = (2 \times R \times 10)^2. \tag{5.3}$$

Hence, the coefficient of oversampling will be:

$$\alpha = \frac{N_{\text{min}}}{N}. \tag{5.4}$$

Fig. 5.8a, illustrates some of these downsampled star targets along with the coefficient of oversampling for the given sequence length. Comparing between the full resolution and the downsampled versions, it is visible by eye that the spokes are properly represented by their downsampled counterparts. Extracting the coefficient of oversampling shows that in a frequency division approach the extracted images are "smaller" than what can naively be expected. This is akin to the smaller individual images of a spatial division approach and results in a fairer comparison of the image compression abilities between the two.

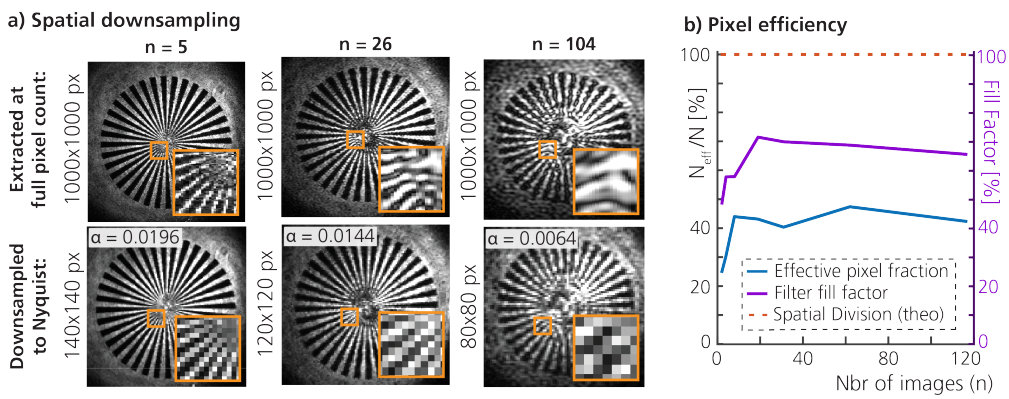


Figure 5.8: Pixel efficiency for the spatial frequency multiplexing strategy.
a) The extracted images are displayed at full pixel resolution. However, due to their decreased spatial resolution, these images are oversampled (see the zoom-ins). Therefore, they can be represented by images that have been downsampled by a factor α, a value that is directly related to the Nyquist sampling rate via Eq. 5.3 and 5.4. **b)** The sequence efficiency for the spatial frequency division strategy, $N_{\text{eff}}^{\text{sfd}}/N$, as a function of sequence depth, lies around 50% (left axis) for a fill factor of around 70% (right axis). Therefore, in its current state, this approach will not be able to reach 100% pixel efficiency. However, there are possible, novel approaches to address this issue.

Pixel efficiency for an entire image sequence

Each pixel of a camera is usually necessary in order to capture an entire movie. However, due to the strong spatial correlations of natural images and the existence of compression algorithms, it turns out that oftentimes the information content within every single pixel is not entirely necessary. Indeed, multiple pixels can be represented as one i.e., a pixel efficiency of unity is not always the case. In order to understand the pixel efficiency of the spatial frequency division approach as compared to the spatial division approach, one can investigate and compare the number of pixels necessary for an entire sequence, assuming that the spatial division approach is optimised for the event of interest, i.e., it cannot be compressed further. For a sequence, or video, of sequence depth n, the spatial frequency division strategy would necessitate $N_{\text{eff}}^{\text{sfd}} = N_{\text{min}} \times n$ pixels. In contrast the spatial division approach necessitates, in the theoretically best case, the entire sensor, i.e., $N_{\text{eff}}^{\text{sd}} = N$ pixels. Comparing the two values shows that the pixel use of the spatial frequency division approach is not fully optimised in the present measurements (Fig. 5.7b). Here, a sequence efficiency ($N_{\text{eff}}^{\text{sfd}}/N$) of around 50% (as compared to the spatial division approach) is achieved. Furthermore, this value is at a fill factor (the fractional area covered by all the filters compared to the sensor size) of around 70% showing that the *individual pixel efficiency* for the spatial frequency division strategy lies around $50/70 = 70\%$ in its current state. Furthermore, since the fill factor will never reach full coverage due to the extent of the DC component, this pixel efficiency will never quite reach 100% due to the resulting lower sequence depths. With that said, the radii of the filters have not been chosen according to the noise considerations discussed in Fig. 5.4, therefore, the fill factor is probably overestimated in these measurements. In fact, we have seen effective pixel fractions that are larger than the fill factor (Paper II) hinting towards single pixel efficiencies that are greater than 100%. However, as more images are superimposed upon each other, the amount of individual image signal above the noise floor decreases with sequence depth, resulting in a trade-off between spatial resolution and dynamic range. This trade-off is quite non-trivial and is to be considered on a single measurement basis as it really comes into play for these extreme sequence depths. Luckily, there might be a more novel trick that can be applied for increased pixel efficiencies.

Spatial phase manipulation for higher pixel efficiency

Within the lock-in algorithm, the demodulation step is performed with a complex exponential, hence a sine and a cosine. The use of these orthogonal functions as the references eliminates the unknown spatial phase of the applied modulation. Hence, if this phase is known, one could essentially store two individual images within a single spatial frequency by phase shifting their carriers by $\pi/2$ (given a modulation depth of one for both modulations). This would mean that at a fill factor of 70% (for the current measurements) one

could double the sequence depth, resulting in a $N_{\text{eff}}^{\text{sfd}}/N$ of nearly 100%. This trick has successfully been performed for two images [111], yet preliminary unsuccessful measurements were performed by me with this AOD-based setup (probably due to the lack of a reference measurement and the need of a modulation depth of one). Nonetheless, this is a promising approach for FRAME as it could push the technique up on the sequence depth ladder by a factor of two, for free.

4.3 Dynamic range

Since all images are stacked on top of each other the images extracted by the lock-in algorithm will be, in a similar manner to spatial oversampling, oversampled in dynamic range. The meaning of this is that, if a camera has a 12-bit sensor, all stored images must share these available 4095 counts. We know from Eq. 5.2 that half the individual image power ends up in the DC component while the other half ends up in the multiplexed components. Hence, theoretically, each multiplexed image can only have an effective dynamic range, DR_{eff}, of:

$$\text{DR}_{\text{eff}} = \frac{2^{\text{bit depth}} - 1}{2} \times \frac{1}{n}, \tag{5.5}$$

for a sequence depth of n images. An analysis of this trend on the same dataset used for the previous section indeed follows this modified $1/n$ structure closely (Fig. 5.9a). The fact that the measured data is slightly shifted downwards from the theoretical curve originates from the act of avoiding saturation in the experimental setting.

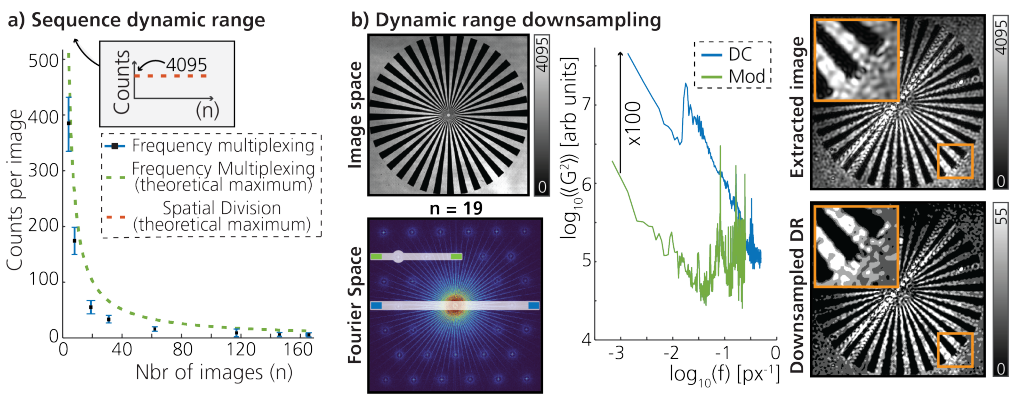


Figure 5.9: Dynamic range for the spatial frequency multiplexing strategy.
a) The number of available counts per image in the theoretical case (green dashed line) and for the experimental data (black dots with blue error bars) is presented. The lower than theory values for the experimental data is attributed to acquiring below the camera's saturation limit. The inset illustrates the constant dynamic range for spatial division approaches. **b)** Extracting the dynamic range from experimental data involves extracting the cluster SNR in relation to the DC component. The downsampled dynamic range for this case, $n = 19$, results in an intensity range from 0 to 55 counts.

In practice, downsampling the dynamic range can be performed by analysing the DC component strength in comparison to the modulated component strengths (Fig. 5.9b). Performing a horizontal integration results in recognisable curves where the maxima of the modulated components are around 100 times weaker compared to the DC component (for the case of $n = 19$). By realising that the DC component's maximum SNR corresponds to half of the total number of counts, then the number of counts available to the modulated components can be deduced from their respective maximum SNRs. Performing this analysis for the case of $n = 19$ results in the number of available counts per image being 55. The downsampled extracted star target (in terms of dynamic range) will hence consist of intensity values between 0 and 55 (Fig. 5.9b) instead of the extracted 0 to 4095.

In the case of transmission imaging of 2-bit information (the star target) the resulting low dynamic ranges do not affect the extracted image to a large degree. However, if one is to approach fluorescence imaging with a spatial frequency division approach, this fact can pose a problem, and could ultimately be the main limit to the number of images within a sequence. On the other hand, if the goal is to monitor natural events that move on an ultrafast timescale, this consideration could pose less of an issue. Indeed, as the sample moves across the sensor, it will fall upon new pixels, hence the dynamic range will not necessarily need to be shared between each modulated image.

4.4 Concluding remarks on multiplexed image storage efficiency

Although at first glance the spatial frequency division approach seems to outdo the spatial division approach, when introducing variables that allow for correct comparisons, the two approaches seem to be quite similar. For example, even though the effective pixel fraction for an entire sequence is measured as being below 50% for the frequency division approach, this is not too surprising as the lock-in algorithm in its current state applies two orthogonal functions in order to determine the unknown phase, essentially halving the bandwidth of the approach.

A significant pro to the spatial division approach is that each image will be able to use the entire dynamic range no matter the event under investigation. However, due to the optical design necessary for achieving spatial division, stereoscopic effects are usually unavoidable and complicate its use in e.g., microscopic applications. On the other hand, spatial frequency multiplexing has the advantage of pixel-to-pixel correlations between extracted images and a significantly enlarged field-of-view (the field-of-view is independent of sequence depth) allowing for the tracking of events over larger distances and extraction of quantities, such as accelerations, without the need for spatial calibrations (Paper III).

All in all, comparing ultrafast videography approaches in terms of these introduced parameters (other than only speed, sequence depth and image size) opens up for well-informed

discussions on the most suitable videography method for a given application. The papers accompanying this thesis constitutes of, what I think, good examples of applications where FRAME excels in terms of imaging, algorithmic applicability and experimental ease, all considerations that can be made with, among other things, the terminology introduced in this section.

5 Exotic Packing Strategies

During this thesis work, we have experimented with various, more exotic ways of packing information in the Fourier domain. These include creating sidebands to the already modulated data and encoding information about the fundamental light-matter interactions that occur at the sample via the use of structured light.

5.1 Doubly modulated packing strategy

Within the context of fluorescence imaging, we have applied a more exotic packing strategy, dubbed double modulation (Paper VIII). The underlying idea is to modulate already modulated information creating satellite or sideband modulations around each Fourier cluster. Mathematically speaking, if one applies a modulation, M_0 (of the usual form) to an object with light distribution $g(x,y)$ then the resulting detected signal, g^{M_0}, will have the, at this point recognisable, form:

$$g^{M_0} = g(x,y)M_0 = g(x,y)\left[\frac{1}{2} + \frac{1}{2}\cos\left(2\pi\mathbf{f}_0\mathbf{r}\right)\right], \tag{5.6}$$

for spatial frequency $\mathbf{f}_0 = (f_{0x}, f_{0y})$ and $\mathbf{r} = (x,y)$ spatial coordinates. If we were to further modulate g^{M_0} with a spatial intensity modulation M_1 of frequency $\mathbf{f}_1$, the resulting detected image, $g^{M_0 \cdot M_1}$ will have the form:

$$g^{M_0\cdot M_1} = g(x,y)\left[\frac{1}{2} + \frac{1}{2}\cos\left(2\pi\mathbf{f}_0\mathbf{r}\right)\right]\left[\frac{1}{2} + \frac{1}{2}\cos\left(2\pi\mathbf{f}_1\mathbf{r}\right)\right] \Rightarrow$$

$$g^{M_0\cdot M_1} = \frac{g(x,y)}{4}\left[\frac{1}{2}\cos\left(2\pi(\mathbf{f}_0+\mathbf{f}_1)\mathbf{r}\right) + \frac{1}{2}\cos\left(2\pi(\mathbf{f}_0-\mathbf{f}_1)\mathbf{r}\right) + \cos\left(2\pi\mathbf{f}_0\mathbf{r}\right) + \cos\left(2\pi\mathbf{f}_1\mathbf{r}\right)\right], \tag{5.7}$$

where the last step is performed via product-to-sum trigonometric identities. $g^{M_0\cdot M_1}$ will hence be built up of image copies of $g(x,y)$ at the expected frequency positions $\mathbf{f}_0$, $\mathbf{f}_1$ and at the sidebands $\mathbf{f}_0 \pm \mathbf{f}_1$.

Fig. 5.10 presents measurements of exactly this, performed by illuminating fluorophores with low spatial frequency magnitudes, $|\mathbf{f}_1|$, and detecting the fluorescence through Ronchi

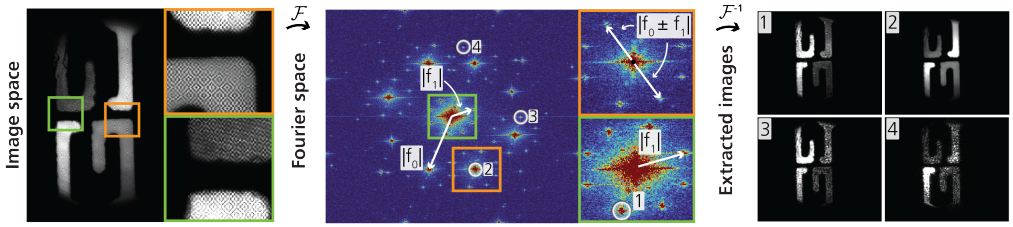


Figure 5.10: Double modulation as an exotic packing strategy
Fluorophores are illuminated with modulations of frequency magnitude $|\mathbf{f}_1|$ and detected via Ronchi gratings of frequency magnitude $|\mathbf{f}_0|$. This results in Fourier clusters at said frequencies along with sidebands of lower intensities at the cross-term positions. Extracting the images results in 24 unique views of the fluorescence in terms of illumination and fluorescence response (they are illuminated with different lasers and detected via different spectral filters), where four excerpts are labelled in the Fourier domain and depicted to the right.

gratings with much higher frequency magnitudes, $|\mathbf{f}_0|$. Hence, we were able to multiplex four different laser wavelengths with four different spectral filters resulting in 24 unique views of the spectral information embedded in the image (four excerpts are shown in Fig. 5.10). In laymen's terms, the real benefit of this packing strategy is that we can specifically identify the emission spectra of the individual fluorophores when excited at a specific wavelength, a fact that allowed us to separate fluorophores of highly overlapping spectral structure with high fidelity.

Other than the decreasing power within the sidebands, there is theoretically no limit to how far one wants to follow the chain of modulating modulated image information for applications to multi-dimensional imaging. For example, adding polarisation or temporal sensitivity in a third multiplexing step would open up for further views of the sample that are sensitive to different types of photon information.

5.2 Encoding light-matter interactions with structured light

When sinusoidally modulated light is incident on an object of interest, it essentially samples the light-matter response of said object over intensities varying from 0 to I_{max} for I_{max} the sinusoidal's peak intensity. In the case of single photon fluorescence, the response, or the fluorescence, to the structured light will be a one-to-one mapping with intensity, hence the detected fluorescence will be identical in structure to the input, i.e., a sin input function will result in sin structured fluorescence. However, the response can be much more complicated and during this thesis work we were able to test two different types of responses: two-photon response and saturated fluorescence response (Klara Lozani's bachelor thesis work [112]).

For the case of two-photon response (Fig. 5.11a), the output fluorescence intensity is dependent on the square of the input intensity. Therefore, an input sine intensity profile becomes a sine squared, visible as the less rounded peaks in the spatial intensity plot. Performing a Fourier transform we see multiple new peaks arise, a direct effect of said squared

response in conjunction with the existence of a detection threshold. This approach brings with it an unexpected advantage, namely signal amplification. Since the average value of a $\sin^2$ function is 50% higher than a sin, the average fluorescence from a two-photon response will be amplified upon illumination with a sinusoidal spatial intensity pattern as compared to a constant spatial intensity profile.

An identical analysis for saturated fluorescence is shown in Fig. 5.11b. For each intensity point on the input sinusoidal intensity pattern, the fluorophores will respond according to their saturation curves, resulting in a sinusoidal with a rounded structure and, as expected, an increased number of Fourier peaks. This approach brings with it the advantage that, instead of sampling the saturation curve with individual input intensity data points (requiring individual experiments), the entire function can be sampled within a single shot, enabling the identification of various fluorophores via the direct extraction of their saturation curves and their induced structure within the Fourier domain. This sampling aspect of sinusoidally structured light brings with it both pros and cons. On the positive side, light-matter responses that are dependent on intensity can be probed within a single shot opening up for the identification of various processes such as e.g., multiphoton excitation or laser-induced incandescence [113]. Furthermore, the application of sinusoidals in conjunction with a saturated fluorescence response creates overtones such that a theoretically

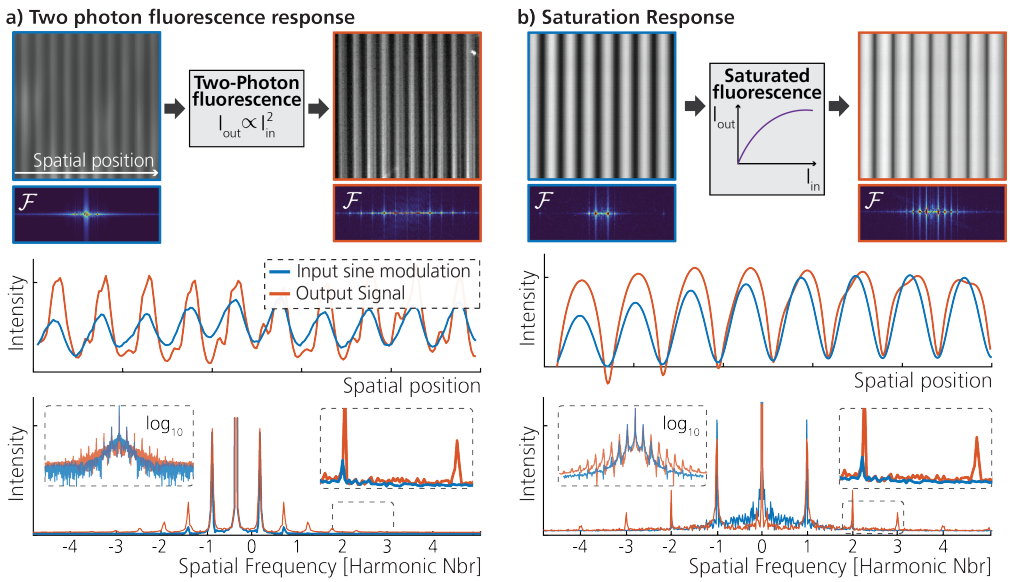


Figure 5.11: Probing light matter interactions with structured light.
a) Performing two photon PLIF with a sinusoidal intensity modulation across the excitation sheet results in a $\sin^2$ response. Multiple new peaks arise in the Fourier domain bringing with them multiple advantages such as signal amplification and a direct fingerprint of non-linear light matter interactions. **b)** Saturation fluorescence with a sinusoidal structured sheet will also result in the creation of overtones. This application has the advantage of being able to sample the fluorescence response over intensities and can lead to enhanced spatial resolutions.

infinite spatial resolution is achievable [114]. On the other hand, if one is to use structured light for other purposes that involve the use of the lock-in algorithm (such as 3 dimensional imaging) one must take care of the induced Fourier overtones in order to be able to reach quantifiable results.

These constitute two examples of how information about light-matter interactions can be stored within the Fourier domain. Indeed, by analysing the difference between the input and the output Fourier structures, a possible new optical diagnostic strategy could open up.

6 Pushing to the Limits

With all these considerations in mind an attempt at storing as much image information as possible within the Fourier domain is overdue. Therefore, a 100 kHz pulse train with 498 pulses of unique spatial modulations was created in order to illuminate a rotating fan. The packing structure is chosen to be hexagonal and the low-pass filtering is performed with a hexagonal filter of radius half the distance to the neighbouring Fourier clusters. The resulting image and its Fourier domain is depicted in Fig. 5.12. Note how, the stored images are placed such that they encircle the DC component at quite a large radius.

A selection of extracted images is shown in the rightmost panel, depicting the rotating fan. A couple of details representative to such high sequence depths need to be addressed:

- The contrast of the fan blades is not always 1, i.e., they are not completely black. This is due to the fact that the SNR of the Fourier components is extremely low, hence shot noise power, i.e., noise that exists over the entire Fourier domain, becomes a significant issue. Indeed, about 50% of images will have signal in the same spatial positions as where the blade lies in the other 50% of the images. Hence this shot noise will leak into all images resulting in reduced contrast. This is not crosstalk!

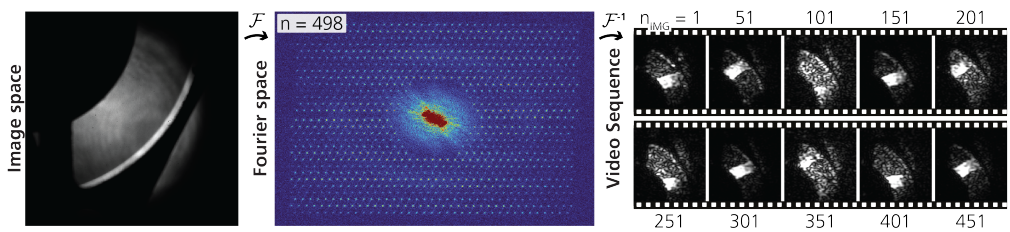


Figure 5.12: Pushing the limits of image storage - $\mathbf{n = 498}$ images.
A rotating fan is illuminated by a 100 kHz pulse train consisting of 498 pulses with modulation frequencies distributed over a hexagonal Fourier structure. Extracting the video sequence results in images where the fan blades are indeed identifiable and their progression can be followed.

- The images are characterised by a "squiggly" structure, something we internally call a bag of worms[8]. This is due to oversampled noise. In short, since the image size is larger than what is necessary according to the low-pass filtering, the noise will not be completely uncorrelated resulting in said "bag of worms" type structure. The solution is, as previously discussed, to downsample the image to its Nyquist sampling rate.

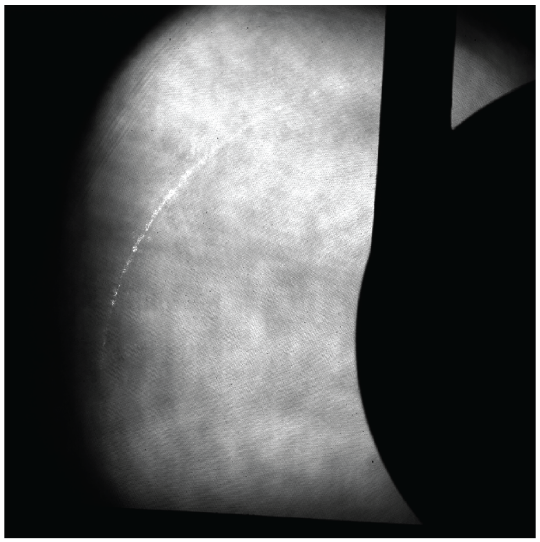
Even though the image quality is not ideal, the ability to track the fan blade throughout this video, showcases how frequency division multiplexing can indeed be pushed towards the higher reaches of sequence depth and the holy grail of ultrafast videography.

An ending worthy of FRAME

Thus, I would like to end this chapter on a high note. To your right you see an image of a rotating fan illuminated by 124 uniquely spatially modulated pulses at a speed of 10 kHz.

[8] TM David Andersson

Image space



Now turn the page...

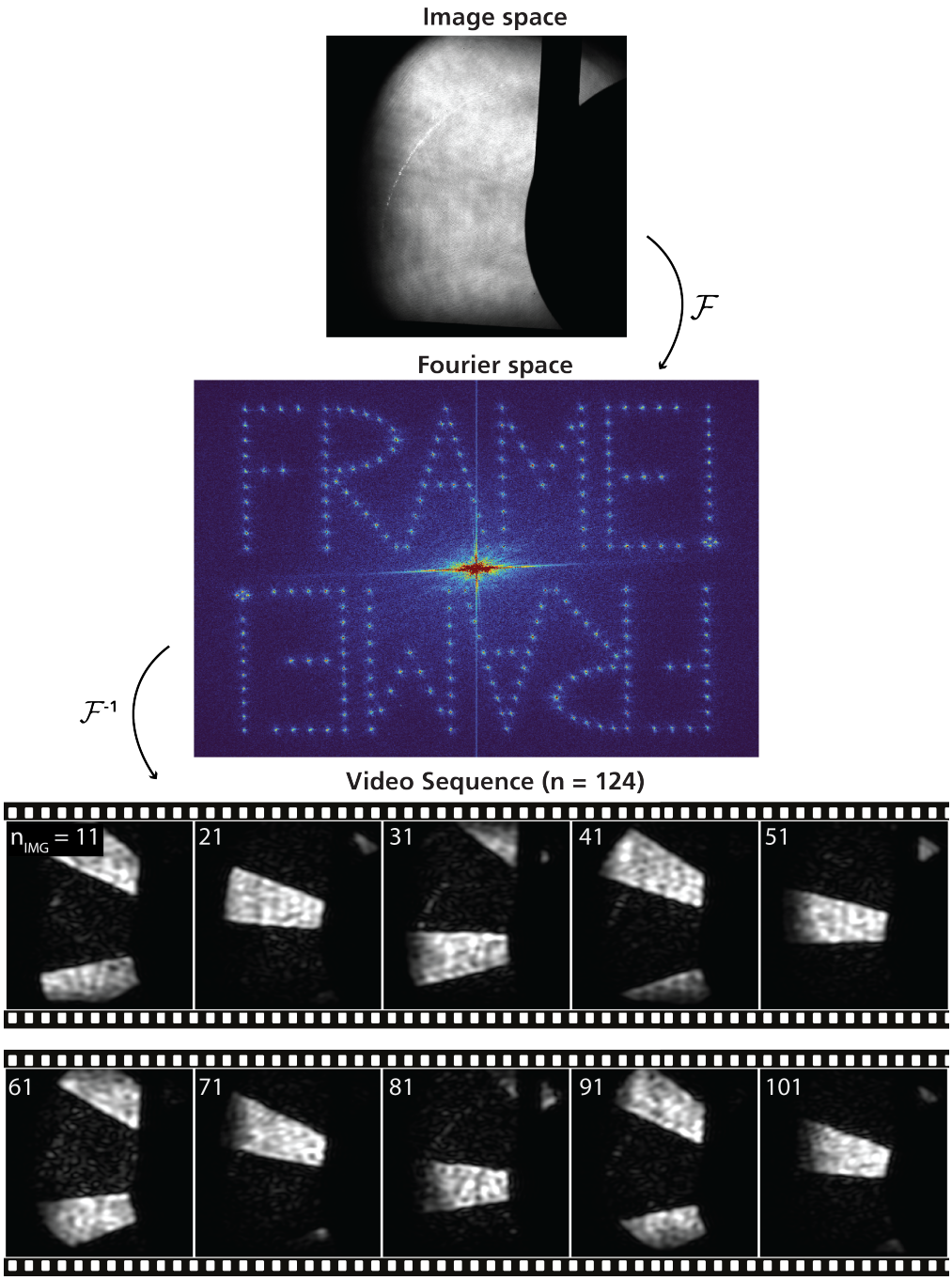


Image space
$\mathcal{F}$
Fourier space
$\mathcal{F}^{-1}$
Video Sequence (n = 124)
n_{IMG} = 11
21
31
41
51
61
71
81
91
101

Chapter 6

Outlook and Final Remarks

This thesis has been divided into four chapters, from optical diagnostics, through Fourier domain formalism to actually utilising all this in the context of ultrafast videography. In this chapter I will attempt to wrap all this information up and supply the reader with an explanation as to you why I have chosen these specific topics for discussion. But first we have one more piece of the puzzle to clear up and that is temporal resolution.

1 Temporal Resolution

There is one more central component to filming and videography: namely that the act of videography can essentially be boiled down to a spatiotemporal sampling of an event. We have previously defined spatial sampling in Chapter 4 and temporal sampling in Chapters 2 and 3. However, for the latter case, it was only trivially connected to the time-bin size definition of frame rate. Let me explain.

For passive approaches, when a single camera is used, the temporal resolution is oftentimes the inverse of the exposure time. Indeed, in most cases the pixels need to be emptied of electrons before a new frame is acquired. For active approaches, especially when pushed to their limits, the illuminating pulses will overlap to a certain degree. Hence one can sometimes see faster imaging rates than individual pulse lengths or, in other words, higher frame rates that what is allowed by the time-bin size. Therefore, if we are to hold the criterion that each image must be unique for a series of images to count as a video, we must propose a definition of uniqueness with respect to time-bin size.

For Gaussian illumination, i.e., a Gaussian shaped integration window, it is generally accepted that its full width at half maximum (FWHM) corresponds to the integration window size. In order to generalise this to variously shaped light pulses, we can turn to a discussion of what the FWHM of a Gaussian means in terms of energy. 76% of a Gaussian pulse's energy lies within the FWHM and this is the uniqueness value that we will define as the lower limit for the sampling of the time-axis of an event (Fig. 6.1a). Indeed, if a pulse train of Gaussian pulses are temporally placed within each other's FWHM, then we will define this as an oversampling of the time-axis, due to a lower rate of uniqueness than 76% between images. Fig. 6.1a illustrates this concept for different pulse shapes that can be encountered upon the use of various illumination technologies such as the type applied in Paper I.

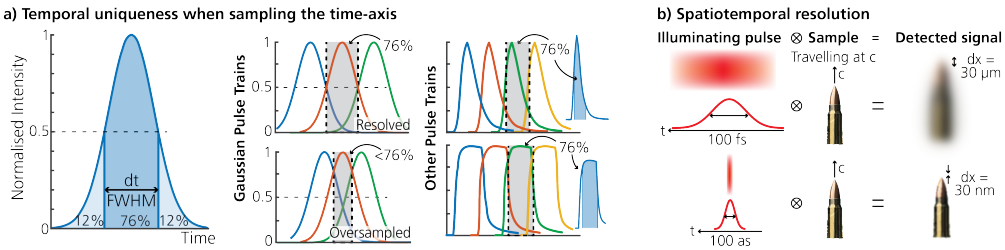


Figure 6.1: Spatiotemporal resolution for active videography.
a) 76% of a Gaussian's power lies in within its FWHM. This value is used as the lower integration window boundary for temporal uniqueness when sampling the time-axis of an event. This generalisation allows for the definition's use with all types of integration windows and pulse trains. **b)** Illuminating an object travelling at the speed of light with a pulse of length 100 fs will result in a motion blurred image if one's spatial resolution is higher than $\sim$30 µm. On the other hand, illuminating with a pulse of length 100 as would necessitate a spatial resolution of approximately 30 nm for such fast illumination to be useful.

Spatiotemporal resolution

In the realm of high-speed and ultrafast videography, the spatial and temporal resolution are intrinsically connected by the speed at which an event unfolds. For example, say that we are monitoring an explosion who's shockwave moves at a speed of $v = 9000$ km/h or $v = 2500$ m/s, a speed that is not too far-fetched within the field of plasma physics or hypersonics. We want to film it with a high-speed camera solution with an exposure time of $dt = 1$ µs (or a 1 MHz frame rate). Then the distance the shockwave moves within a single exposure will be $dx = v \times dt = 2.5$ mm. Therefore, a spatial resolution superior to 2.5 mm is unnecessary for such a measurement situation, as the speed of the event will cause a motion blur within the integration window.

In the limit where the dynamics of a sample unfold at the speed of light, the relation between spatial and temporal resolution will be fully characterised by c, the speed of light. For example, if the illuminating pulse (or integration window) is 100 fs long then the spatial resolution for an event unfolding at the speed of light does not need to exceed $dx \approx 30$ µm (Fig. 6.1b). This is an extremely useful fact for spatial frequency division techniques since, as we have thoroughly discussed, the spatial bandwidth is compromised in order to amplify temporal bandwidth. For the fastest type of events under femtosecond illumination, the loss of spatial resolution required by FRAME can in fact be counteracted by this argument as motion blur is closely related to spatiotemporal oversampling.

Turning this argument around, if one has a fixed spatial resolution it is not always necessary to push one's system to film faster. Indeed, if the illuminating light pulse is 100 as long, then the optics need to allow for a 30 nanometer spatial resolution in order to extract sharp edges, i.e., we would need to turn to super-resolution microscopy.

All in all, due to the convolution between illuminating pulse and event, the ultimate necessary spatial resolution for a given measurement will be directly connected to the integration window and the speed at which the event unfolds, a fact that is not usually discussed within the ultrafast videography community.

2 Outlook

Within this thesis I have defined structured light quite narrowly, namely in terms of sinusoidals. However, within this confined definition, work within this division and thesis has been undertaken to broaden their use into ultrafast videography (Paper I - IV), controlling overtones in the context of light-matter interactions, simultaneous encoding of multi-dimensional information (Papers VII - VIII) and for their use within scattering and noise reduction in detection apparatus (Papers V - VI).

However, I would like to offer the reader an expansion of the term structured light when it comes to its marriage to ultrafast videography or more generally multi-dimensional imaging. When googling structured light, one usually ends up in a rabbit hole of orbital angular momentum (OAM) beams, Gaussian hollow core (GHC) beams, exotic particle like topologies [115, 116] and plenty more. Similarly, when googling mathematical transforms one can easily fall into the rabbit hole of wavelet transforms, Hankel transforms and again, many, many more. By combining various ways of structuring light with corresponding transforms, one could open up for entirely new approaches to achieve Structured Light for Ultrafast Videography[1]. To digress: why not couple the photon dimension of time to e.g., Bessel beams and then extract said information via Hankel transforms[2]?

With that, the outlook I would like to propose, is allowing the reader to interpret the title of this thesis in a much broader sense than the approach described within. The sky's the limit (by sky I mean technological, physical and mathematical constraints).

3 Concluding Remarks

In the very first introductory chapter of this thesis we discussed how ultrafast videography is the natural step that science needs to take for the continued investigation of physical phenomena. Even though I had mentioned it, the fact that it is currently feasible to take this step was not entirely clear at that point. Hints were made by mentioning the multiple Nobel Prizes awarded to the relevant fields throughout modern history, however, the questions of accessibility and ease of use remained unanswered. Attaining these two criteria are key for granting non-experts access to these extreme technologies in everyday laboratory work, opening up for novel applications on phenomena that were hitherto undetectable at high spatiotemporal resolutions.

The second chapter approached these questions by discussing optical diagnostics on a broad scale. The fact that we delved directly into the field of optics is not a matter of chance; it is in alignment with our goal of achieving ultrafast videography. Indeed, due to their lack of mass, photons are the fastest moving particles in the universe, satisfying one third of the title: **Ultrafast**. Moreover, due to our advanced understanding of electromagnetism and materials, there exists a plethora of optical elements, technologies and strategies to measure and manipulate photon information across its dimensions.

In order to fully understand the technological challenges that ultrafast videography is up against, in the third chapter we looked into the current state of technology with respect

[1] Although I got dibs on the title.

[2] I will hedge this statement by saying that I have no idea if this is possible, the 'why not' part is merely a rhetorical tool.

to the three criteria of ultrafast videography. Here we saw that passive approaches are extremely useful and accessible for monitoring a majority of phenomena, however they fall just short of the ultrafast regime. In order to reach the ultrafast, we observed that active approaches are necessary and, due to the increased accessibility of femtosecond laser technology and innovative materials processing, this is a field that has seen a surge of developments since the mid-2010s. Furthermore, today's FPA technology with their extreme spatial bandwidths and dynamic ranges (also a new development) open up for the efficient conversion of photon information to the spatial dimension, a key for the majority of active videography approaches, thus fulfilling another third of the title: **Videography**.

The fourth chapter constituted of, what I would call, the most technical part of the thesis. Here I tried to introduce Fourier optics in a manner such that it is directly applicable to laboratory work. These concepts were introduced alongside a parallel thread of how Fourier optics can be turned into an intuitive sport. The most important result of this section however, was the diamond-like extent of the class of natural images within reciprocal space. In conjunction with current high spatial bandwidth FPA technology, this simple property of natural images creates large voids in the Fourier domain where spatial frequency division approaches can be leveraged, fulfilling the final third of the title: **Structured Light**. Putting this into a larger context, this technological opening can be directly compared to earlier developments within heterodyne detection and frequency modulation in the field of spectroscopy. To go even further, similarities with frequency multiplexing within radio communication can be found – a technology that is central to where society is today.

In the last chapter of this thesis I displayed work that was mainly focused on trying to fill the long-sequence depth hole of ultrafast videography (the upper right corner of the large scale overview of technologies). Here I tried to discuss and illustrate the advantages and pitfalls of spatial frequency multiplexing. Since the star target constituted the sample of choice, the measurements within chapter five should be interpreted as a lower boundary on what is possible to accomplish with the FRAME strategy. For example, monitoring moving targets will increase the upper dynamic range bound of the stored images while monitoring natural images will allow for more leeway with respect to the low-pass filtering part of the lock-in algorithm. However, concepts such as the discretisation of modulations or the non-trivial coupling between spatial resolution and dynamic range are all concepts that I hope to have introduced such that further discussions are made possible.

All in all, due to the ease at which we can perform photon operations, the current state of technology, linear systems theory and an innate curiosity for nature, I hope I have convinced you, the reader, that the time is ripe for the field of ultrafast videography to start answering the difficult questions of nature. Hopefully this thesis, **Structured Light for Ultrafast Videography**, can make a dent in the push for the understanding nature at its limits.

Chapter 7

Appendix

If you are not already tired of me, here comes the best part! The supplementary information to this thesis.

Appendix A: Lock-in Detection

Lock-in detection is a technique most commonly used for increasing the signal to noise ratio of weak signals within e.g., spectroscopy. By first heterodyning said weak signal followed by detection at the applied carrier frequency, $1/f$ noise can essentially be eliminated. In essence, lock-in detection at a given frequency amounts to *detecting the amplitude of the signal at said frequency*. Here I will try to carry the reader through the mathematics of lock-in detection (in 1D) with the assistance of a purely graphical explanation in Fig. 7.1.

Consider a signal under investigation, $S(x)$, which varies along an axis, x, and is sampled at 1000 discrete points (Fig. 7.1a). In this case it is a simple Gaussian with its Fourier transform, (also a Gaussian), depicted to the right. We apply a modulation, M, of the form:

$$M(x) = \frac{1}{2}\Big[1 + \sin(\omega x + \psi_{\text{mod}})\Big], \tag{7.1}$$

where ω is the frequency of the sinusoidal given in this case as $2\pi f_{\text{mod}}/L$ with $f_{\text{mod}} = 100$ and $L = 1000$ and ψ_{mod} is some unknown phase of the applied modulation (Fig. 7.1b). The detected signal, D, is given as:

$$D(x) = M(x)S(x) = \frac{S(x)}{2}\Big[1 + \sin(\omega x + \psi_{\text{mod}})\Big]. \tag{7.2}$$

Hence, the detected signal, shown in Fig. 7.1c, is now modulated and the peaks corresponding to the applied modulation will appear in the Fourier domain at the frequency of M i.e., $\omega = f_{\text{ref}}/L$. The lock-in post-processing algorithm is then applied in two steps: firstly the signal is multiplied by two reference signals, R_1 and R_2, of the same frequency, ω, but with a $\pi/2$ phase shift between them (R_1 is a sine and R_2 is a cosine, Fig. 7.1d). This results in two locked signals, L_1 and L_2, depicted in Fig. 7.1e. Below I perform some calculation steps between R_1 and L_1, however an exactly equivalent calculation can be made for R_2 and L_2:

$$L_1 = R_1 D = \sin(\omega x)\frac{1}{2}S(x)\Big[1 + \sin(\omega x + \psi_{\text{mod}})\Big],$$

$$\Rightarrow L_1 = \frac{S(x)}{2}\left(\sin(\omega x) + \frac{\cos(\psi_{\text{mod}})\cos(2\omega x)}{2} + \frac{\cos(\psi_{\text{mod}})}{2} + \frac{\sin(\psi_{\text{mod}})\sin(2\omega x)}{2}\right), \tag{7.3}$$

where simple trigonometric identities are used in between the two steps. We see from the resulting representation of L_1 that there are three different frequency components in the signal, at 0, ω and 2ω, also clearly visible in the locked signals as four peaks in the Fourier domain of Fig. 7.1e. A more intuitive explanation to the emergence of four Fourier peaks is

that when modulating the signal, peaks will appear to each side. Hence, when modulating already modulated information, if the applied modulation frequency is equal to the existing modulation frequency, one of the new peaks will end up at the DC position of the Fourier domain, i.e., at 0ω and the other will show up at 2ω. A detail that is worth mentioning is that the information from the opposite phase, i.e., the peak at $-\omega$, will also undergo the same transformation, conserving all the phase information in the DC component of panel e. Low-pass filtering with a filter of width $< \omega$ (in practice around $\omega/2$ since the peaks have a certain width and due to the Nyquist sampling criteria) results in the low pass filtered locked signals $\tilde{L}_1$ and $\tilde{L}_2$ (Fig. 7.1f):

$$\tilde{L}_1 = \frac{S(x)}{2}\frac{\cos(\psi_{\text{mod}})}{2}, \tag{7.4}$$

$$\tilde{L}_2 = \frac{S(x)}{2}\frac{\sin(\psi_{\text{mod}})}{2}. \tag{7.5}$$

Then one can realise that:

$$L_1^2 + L_2^2 = \left(\frac{S^2(x)}{4}\right)\big(\underbrace{\sin^2(\psi_{\text{mod}}) + \cos^2(\psi_{\text{mod}})}_{=1}\big),$$

meaning that the original signal, S, is equal to:

$$S(x) = 4\sqrt{\tilde{L}_1^2 + \tilde{L}_2^2}, \tag{7.6}$$

or the rescaled norm of the locked and low-pass filtered modulated signals (Fig. 7.1g). This algorithm can be directly performed in two dimensions, as has been done throughout this thesis, by noting that:

$$S = S(x, y) \tag{7.7}$$

$$M = \frac{1}{2}\Big(1 + \sin(\omega_x x + \omega_y y + \psi_{\text{mod}})\Big) \tag{7.8}$$

$$R = \exp\big(-i(\omega_x x + \omega_y y)\big), \tag{7.9}$$

where R is given in Euler's form as this includes both the sine and cosine term.

There are four important things to note when performing lock-in:

1. After completing the algorithm, the detected modulation, which was applied to the signal in panel b, is eliminated. This since the reference signals "moved" the modulated part of the signal into the DC position of the Fourier domain before low-pass filtering. This algorithm would be equivalent to band-pass filtering at ω and shifting into the centre right after panel c, however here care would have to be taken in order to include the imaginary parts and signal of opposite phase for correct reconstruction.

2. The application of two, $\pi/2$ shifted reference waves eliminates the unknown phase of the applied modulation. Indeed, this process corresponds to projecting the applied modulation phasor onto an orthogonal basis and then taking their norm.

3. If a background signal B_g is applied to the signal in the detection phase, this contribution will not be modulated. Hence when multiplying with a reference, the background will attain an ω component essentially shifting it to a position outside of the low-pass filter. This is the key to eliminating $1/f$ noise from the signal.

4. It is important to note that the width of the peaks in the Fourier domain versus the applied modulation need to be such that the low-pass filter is able to include the entire width of the signal peak. This is of course never the case and a trade off in the resolution of the extracted signal compared to the specificity of the lock-in technique is always present.

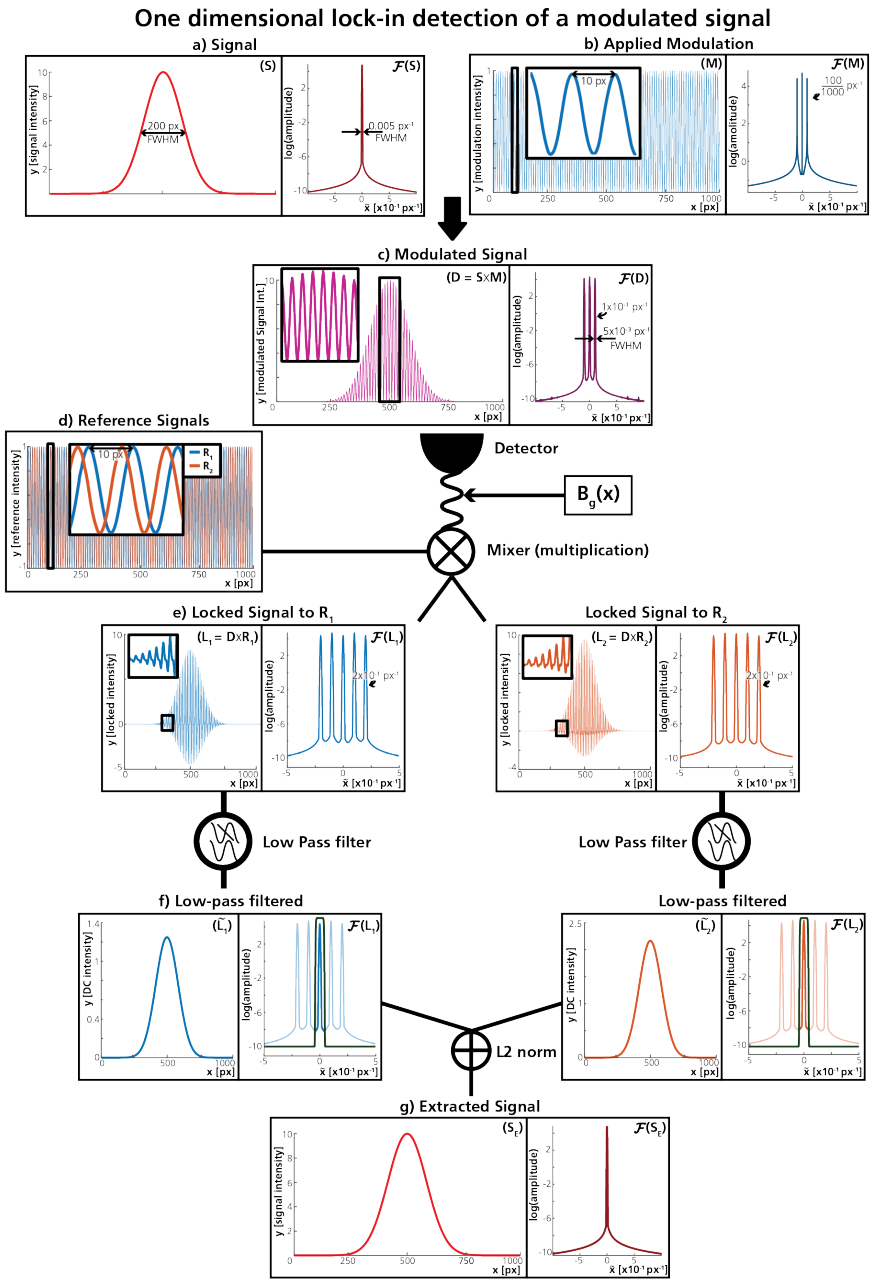
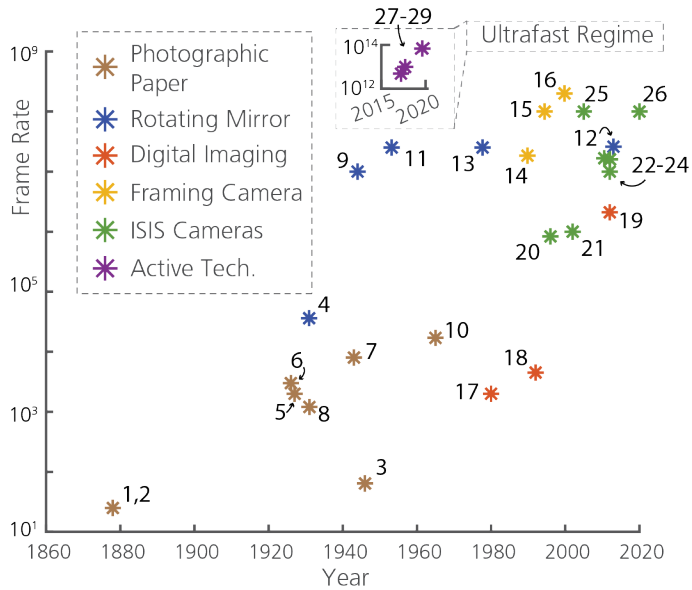


Figure 7.1: Graphical representation of the lock-in algorithm.
a) A signal and **b)** its applied modulation result in a modulated signal with two new Fourier peaks **c)**. Upon detection an unknown background B_g is added. **d)** Two orthogonal reference signals are mixed to the detected signal such that the two so-called locked signals L_1 and L_2 are obtained, with opposite phase. **f)** Low-pass filtering said locked signals results in an extracted signal that is simply a rescaled version of the original without the background contribution of B_g **g)**. This can be easily expanded to a two-dimensional approach by the application of multi-dimensional analysis.

Appendix B: References to "History of high-speed videography" plot

The history of high-speed videography begins at Muybridge's movie of Sallie Gardner and Gilbert Domm. Since then extreme technological advances have ensued, meaning that it becomes extremely difficult to track the development. Therefore, I took the liberty of creating a historical plot in Chapter 3 and here I have included all the references used in the building of said plot. The numbers placed next to the points are the reference numbers corresponding to the reference list below.



[1] J D Stillman and Eadweard Muybridge. *The horse in motion as shown by instantaneous photography, with a study on animal mechanics founded on anatomy and the revelations of the camera, in which is demonstrated the theory of quadrupedal locomotion. / By.* 1882.

[2] Eadweard Muybridge. *Animal locomotion*, volume 534. Da Capo Press Cambridge, MA, 1887.

[3] Cine-Kodak Magazine 8 camera | Science Museum Group Collection.

[4] *Popular Science*, 119(2):24, 1931.

[5] Edward G. Beardsley. The N.A.C.A. Photographic Apparatus for Studying Fuel Sprays from Oil Engine Injection Valves and Test Results from Several Researches, January 1928. NTRS Author Affiliations: NTRS Report/Patent Number: NACA-TR-274 NTRS Document ID: 19930091342 NTRS Research Center: Legacy CDMS (CDMS).

[6] W. H. Connell. The Heape and Grylls machine for high speed photography. *Journal of Scientific Instruments*, 4(3):82, December 1926.

[7] An Ultra-High-Speed Motion-Picture Camera. *Nature*, 153(3872):77–77, January 1944. Number: 3872 Publisher: Nature Publishing Group.

[8] E. Newton Harvey and Alfred L. Loomis. HIGH SPEED PHOTOMICROGRAPHY OF LIVING CELLS SUBJECTED TO SUPERSONIC VIBRATIONS. *Journal of General Physiology*, 15(2):147–153, November 1931.

[9] Berlyn Brixner | Atomic Photographers Guild.

[10] Red Lake Labs Inc. *HYCAM, high speed motion picture camera. 1965:.* 1965.

[11] C4 Rotating Mirror High Speed Camera bolted to rectangular wheeled steel frame - wheels missing | Science Museum Group Collection.

[12] Cordin Rotating Mirror Framing Camera Systems.

[13] Cordin Rotating Mirror Framing Camera Systems.

[14] Brett R. Lawrence. Review of ULTRANAC high-speed camera: applications, results, and techniques. In *22nd International Congress on High-Speed Photography and Photonics*, volume 2869, pages 882–887. SPIE, May 1997.

[15] Imacon468 user guide.

[16] Drs technologies - imacon200.

[17] The sp2000 motion analysis system.

[18] Goji Etoh. 1. high speed camera. *Television Society Journal*, 46(5):543–545, 1992.

[19] Fastcam sa-z hardware manual.

[20] W.F. Kosonocky, Guang Yang, Chao Ye, R.K. Kabra, Liansheng Xie, J.L. Lawrence, V. Mastrocolla, F.V. Shallcross, and V. Patel. 360/spl times/360-element very-high-frame-rate burst image sensor. In *1996 IEEE International Solid-State Circuits Conference. Digest of TEchnical Papers, ISSCC*, pages 182–183, 1996.

[21] T. Goji Etoh, D. Poggemann, A. Ruckelshausen, A. Theuwissen, G. Kreider, H.-O. Folkerts, H. Mutoh, Y. Kondo, H. Maruno, K. Takubo, H. Soya, K. Takehara, T. Okinaka, Y. Takano, T. Reisinger, and C. Lohmann. A ccd image sensor of 1 mframes/s for continuous image capturing 103 frames. In *2002 IEEE International Solid-State Circuits Conference. Digest of Technical Papers (Cat. No.02CH37315)*, volume 1, pages 46–443 vol.1, 2002.

[22] Takeharu G. Etoh, Dung H. Nguyen, Son V. T. Dao, Cuong L. Vo, Masatoshi Tanaka, Kohsei Takehara, Tomoo Okinaka, Harry van Kuijk, Wilco Klaassens, Jan Bosiers, Michael Lesser, David Ouellette, Hirotaka Maruyama, Tetsuya Hayashida, and Toshiki Arai. A 16 mfps 165kpixel backside-illuminated ccd. In *2011 IEEE International Solid-State Circuits Conference*, pages 406–408, 2011.

[23] T. Arai, J. Yonai, T. Hayashida, H. Ohtake, H. van Kuijk, and T. Goji Etoh. Back-side-illuminated image sensor with burst capturing speed of 5.2 Tpixel per second. In Ralf Widenhorn and Antoine Dupret, editors, *Sensors, Cameras, and Systems for Industrial and Scientific Applications XIV*, volume 8659, page 865904. International Society for Optics and Photonics, SPIE, 2013.

[24] Yasuhisa Tochigi, Katsuhiko Hanzawa, Yuri Kato, Rihito Kuroda, Hideki Mutoh, Ryuta Hirose, Hideki Tominaga, Kenji Takubo, Yasushi Kondo, and Shigetoshi Sugawa. A global-shutter cmos image sensor with readout speed of 1tpixel/s burst and 780mpixel/s continuous. In *2012 IEEE International Solid-State Circuits Conference*, pages 382–384, 2012.

[25] Leonid Lazovsky, Daniel Cismas, Gary Allan, and Doug Given. CCD sensor and camera for 100 Mfps burst frame rate image capture. In *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications II*, volume 5787, pages 184–190. SPIE, May 2005.

[26] Manabu Suzuki, Yuki Sugama, Rihito Kuroda, and Shigetoshi Sugawa. Over 100 Million Frames per Second 368 Frames Global Shutter Burst CMOS Image Sensor with Pixel-wise Trench Capacitor Memory Array. *Sensors*, 20(4):1086, January 2020. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[27] 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, 03 2017.

[28] Jinyang Liang, Liren Zhu, and Lihong Wang. Single-shot real-time femtosecond imaging of temporal focusing. *Light: Science and Applications*, 7, 08 2018.

[29] Lihong Wang, Peng Wang, and Jinyang Liang. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nature Communications*, 11, 04 2020.

Appendix C: Notation and abbreviations for Chapter 4

Below I have included all the notations and abbreviations that are necessary to follow the discussions of Chapter 4.

Variable	Short description
$h(\mathbf{r})$	The elementary function as a function of spatial variables **r**
$g(x,y)$	A spatial distribution (usually used as a spatial distribution of light) in the spatial domain. Used as the image in the majority of cases.
$G(f_x,f_y)$	A spatial distribution (usually used as a spatial distribution of light) in the spatial Fourier domain. Also called the Fourier coefficients
$\mathcal{F}/\mathcal{F}^{-1}$	The Fourier/inverse Fourier transform operation.
$g_s(x,y)$	The sampled version of $g(x,y)$
$G_s(f_x,f_y)$	The Fourier transformed sampled light distribution, i.e., $\mathcal{F}\{g_s\}$
$U(\xi,\eta)$	An arbitrary light distribution, usually used as an object light distribution, i.e., equivalent to g. This is the light distribution that we usually use as the starting point in the Fresnel diffraction integral.
$U(x,y)$	The resulting light distribution after having followed $U(\xi,\eta)$ for a given distance/through optical components in the Fresnel diffraction integral.
t_l	The phase transformation function of a thin lens
$U_l(x,y)$	The light distribution immediately after a light distribution has passed through a lens, i.e., it has been transformed by t_l

Appendix D: The optical system used for the results presented in Chapter 5 - the Fourier machine

In order to create arbitrary spatial intensity modulations at the object plane, an acousto-optic based Michelson interferometer was built during my thesis work (Fig. 7.2). By sending RF signals to the acousto-optic deflectors (AODs) the incoming laser light will obtain an angle depending on the frequency of the acoustic wave and a deflected power dependent on the RF power. By varying said frequency one can impart arbitrary angles on the laser beam, resulting in, after an imaging stage, arbitrary spatial intensity modulations at the image plane. Note that in order to create stable intensity modulations, the same RF signal must be sent to each AOD, hence RF amplification and splitting is required. Furthermore, one of the AODs is simply turned upside down along with an extra mirror in within its laser path in order to double the deflection angle between the two arms. This setup has been able to perform Schlieren imaging, transmission imaging, reflectance imaging and has the potential to perform fluorescence imaging, a project that is under design.

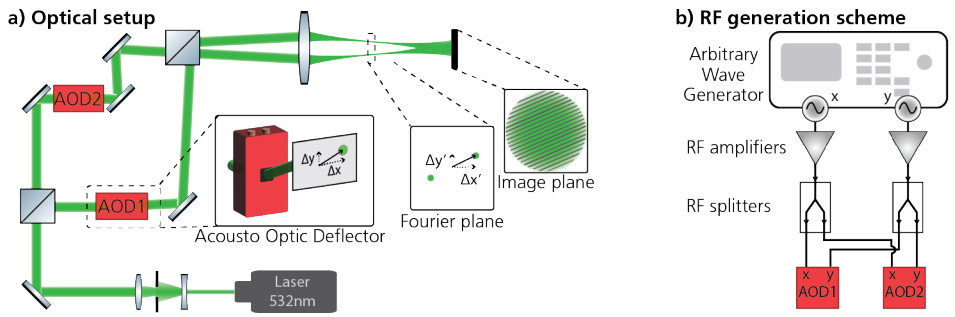
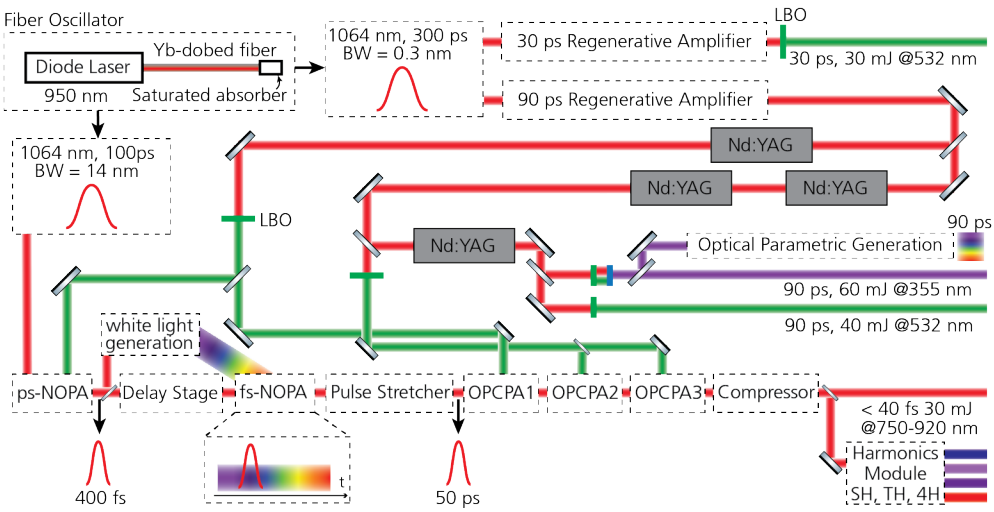


Figure 7.2: An acousto-optic based Michelson interferometer for spatial frequency multiplexing
a) The optical setup where two acousto-optic deflectors are placed in each arm of a Michelson interferometer. Upon interaction, their acquired angles in x and y result in spatial intensity modulations in the image plane. **b)** The RF generation and setup used in order to perform these measurements and create arbitrary spatial intensity modulations in the image plane. The setup is capable of reaching 1 MHz frame rates and close to arbitrary sequence lengths.

Appendix E: Laser System used for the work presented in this thesis

Below is a diagram of the Ekspla FFL002 laser system used for the work within this thesis. Due to its size and complexity, I have included it to assist in discussions about FRAME, ultrafast videography in general and what is possible to achieve with today's OPCPA based laser technology. An important note is that all the pulses are optically synchronised, therefore the jitter between them is negligible.



Appendix F: Coherence Lifetime Imaging

I think you got a J-number problem

When a femtosecond laser pulse is incident on an ensemble of molecules, it will accelerate the charge clouds of said molecules resulting in isotropically emitted Rayleigh and spontaneous Raman light[1]. There will however exist a subset of molecules within this isotropic distribution that are aligned with the polarisation of the incident pump laser. These molecules, or induced dipoles, will emit light according to the stimulated Raman process. The reactionary effect on the incident femtosecond pulse will in practice be measured as

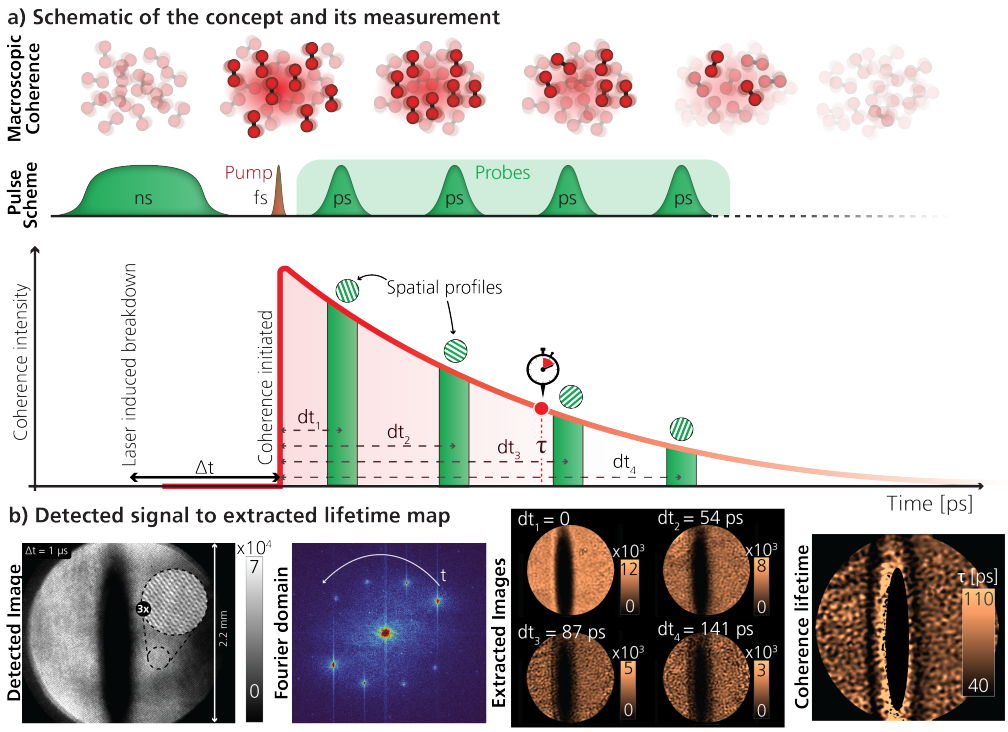


Figure 7.3: Coherence lifetime imaging of nitrogen molecules.
a) A general schematic of the measurement where a nanosecond laser is incident onto a flow of nitrogen such that a laser induced breakdown event occurs. After a time, Δt, a broadband femtosecond pulse initiates the macroscopic coherence over the rotational manifold of nitrogen. Four spatially multiplexed, 30 ps pulses then probe the coherence decay via a stimulated Raman process. **b)** The four anti-stokes images of the generated coherence are incident onto a camera such that four Fourier clusters are visible. The extracted images, clearly decrease in intensity as the coherence decays. A weighted non-linear least squares fit is performed on the extracted images such that a lifetime map is compiled. Here we see the shockwave clearly as shorter lifetimes.

[1]The observed wavelength shifts will correspond to the ro-vibrational (at standard conditions only rotation) energies populated within the ensemble.

gain since the coherent redistribution of energies within the rotational manifold mostly fall within the large spectral bandwidth of the femtosecond pulse[2].

However, if the pump laser pulse is intense enough, the AC Stark effect will come into play breaking the degeneracy of the rotational J levels such that the ensemble of molecules with redistributed energies becomes larger[3]. On a microscopic level, the affected molecules (1) align according to the polarisation of the incident laser source and (2) their rotational quantum state is reset to a well-defined phase (Fig. 7.3a, upper panel).

Resetting the phases of the individual molecules' quantum states results, on a macroscopic level, in an ensemble of coherently redistributed molecules, lying within the propagation volume of the pump pulse. This can be described as a wave packet where the constituent electric dipoles oscillate in phase at integer multiple frequencies. However, once the incident pump laser pulse has passed, inter-molecular collisions will cause random phase jumps within the single molecules. Such phase jumps will result in a decrease in the ensemble size that contributes to the macroscopic coherence. Given enough time, all molecules will have returned to their natural states with random rotational phases and the macroscopic coherent structure will have dissipated. This decay can be summarised by a multi-exponential decay curve akin to e.g., the much slower process of fluorescence decay where the so called "coherence intensity" of the lower panel of Fig. 7.3a, constitutes a measure of the ensemble size.

Hence, the incident femtosecond pump pulse has the ability to generate a short-lived ($\approx$ 100 ps), delicate macroscopic coherent state whose rate of decay is dictated by the parameters of the local environment, primarily inter-molecular collision cross sections, molecular composition, temperature and pressure. Therefore, quantifying the decay of this extremely short-lived coherence in 2D would open up for new diagnostic means to measure these fundamental physical drivers of chemical reactions on a sub-nanosecond timescale. From an experimental standpoint, measuring the *decay* rather than the magnitude at a single point in time, has the advantage of being, akin to fluorescence lifetime imaging, quantitative as well as insensitive to number density variations, laser power fluctuations and detection yield. In addition, contra the slower process of fluorescence, this coherent state can generate signal from dark (non-fluorescing) species with rotational manifolds, such as ground state N_2, and is characterised by a directional, laser-like coherent signal.

However, such a measurement would necessitate hardware that is able to reach picosecond temporal resolutions and furthermore, in order to see chaotic local environmental variations, a two dimensional technique is a necessity. We approach this problem by probing

[2]I have no better spot to place this but here is a reference on the first article about quantised rotational spectra written by Eva von Bahr in 1914 [117]. Note that she has extremely few citations even though she laid the groundwork for the quantum theory of rotations with her molecules of choice being CO and H_2O.

[3]In short, this is due to an extra potential induced by the applied torque from the laser light.

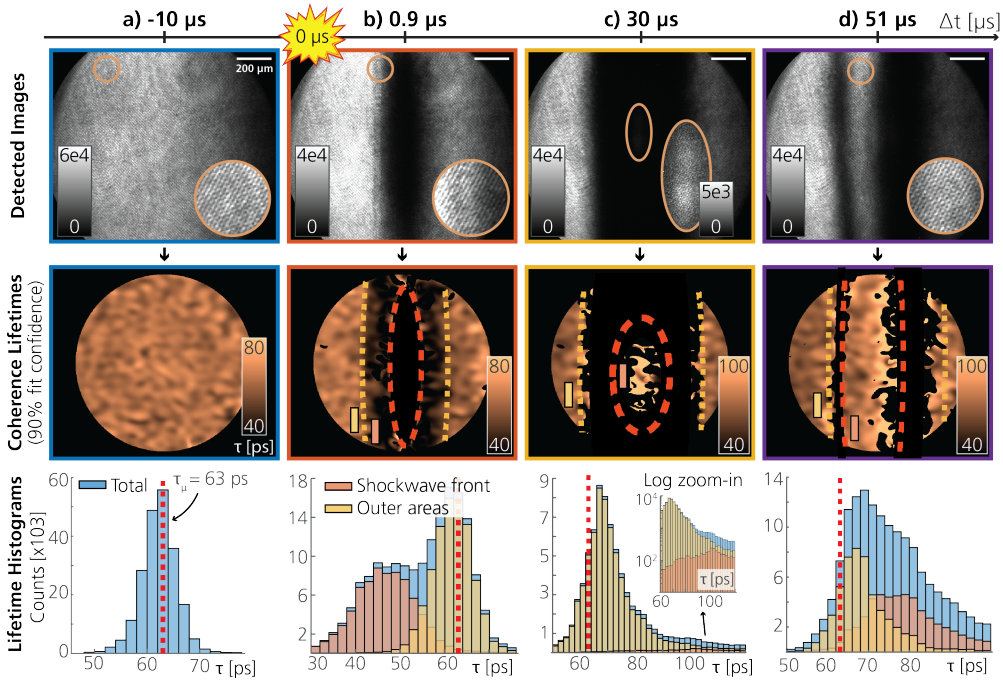


Figure 7.4: Lifecycle of laser induced breakdown.
Multiplexed Coherent Anti-stokes Raman signals are detected on the camera at different Δt with respect to the laser induced breakdown. Lifetime maps are then extracted for each pixel followed by the histogramming of said lifetimes within set boundaries (pink and yellow dashed lines). Here, four typical stages are included for further analysis. **a)** Before the breakdown is initiated, a flat field of room temperature nitrogen is recorded. A mean lifetime of 63 ps is recorded (red dashed line in the histogram). **b)** After some time, a shockwave is created and pushes outwards, inducing lower lifetimes at the front. **c)** After enough time signal starts to reappear in the centre of the hole with a longer lifetime. **d)** As the hole fills up, longer lifetimes are recorded for the nitrogen that arises in the centre.

the decay with four individual, two dimensional, probe pulses, separated by $\sim$ 50 ps in time (Fig. 7.3a). As these probe pulses scatter on the macroscopic coherence, in a way that can be summarised by a stimulated Raman process, the resulting anti-stokes signal will decay in intensity as a function of the coherent ensemble size. By following this intensity decay as a function of probe pulse, the decay time can be extracted into a two-dimensional coherence lifetime map. As expected, we do this by employing the ultrafast videography technique that is FRAME in order to probe said coherence decay where a detailed schematic of the optical setup is included below (Fig. 7.5).

This technique is applied to probe the chaotic dynamics of a laser induced breakdown process, initiated by an 8 nanosecond 532 nm laser pulse where an example image of the detected signal, 1 µs after breakdown, is depicted in Fig. 7.3b. As expected, the Fourier transform exhibits four separated clusters which decay in intensity in an exponential manner. The four extracted images are then used in order to perform a pixel-by-pixel weighted non-linear least squares fitting routine resulting in a lifetime map.

Lifecycle of a laser induced breakdown event

Typical measurements during the lifecycle of the laser induced plasma breakdown process are shown in Fig. 7.4. Approximately 10 µs before the breakdown process, a "whitefield" image is acquired, i.e., a coherence lifetime map of room temperature nitrogen, resulting in the expected mean lifetime of $\tau_\mu = 63$ ps. Upon breakdown ($\Delta t = 0$), the molecules are ripped apart and this sudden dumping of energy results in a shockwave that propagates outwards from the centre at a speed of $\sim$ 9000 km/h. This shockwave heats, dissociates and excites the nitrogen molecules in its path, resulting in the disappearance of a coherence signal within the shockwave volume (the eye like shape is due to the geometry of the setup and is described by the cross-sectional cut of a sphere). At the shockwave front however, the increased collisional rate significantly decreases the coherence lifetime to $\sim$ 45 ps while on the outskirts, i.e., where the shockwave has not yet reached, the lifetime stays close to the room-temperature value of 63 ps (red dashed line at $\Delta t = 0.9$ µs). As time goes on, the eye becomes bigger however around the 30 µs signal starts to appear in the centre of the hole. Its low intensity is attributed to either low number density of molecules or a rotational excitation (a high rotational temperature) that is beyond the bandwidth of the measurement. The resulting lifetimes from said signals are also longer, peaking at around 100 ps. Again, this could be due to lower number density or higher rotational excitation. As the eye fills with molecular nitrogen, longer lifetimes remain (the double peak on the histogram of 51 µs) until room temperature gas starts mixing and cooling the gas within the probe volume. The entire process finally relaxes back to room temperature and pressure after $\sim$ 1 ms.

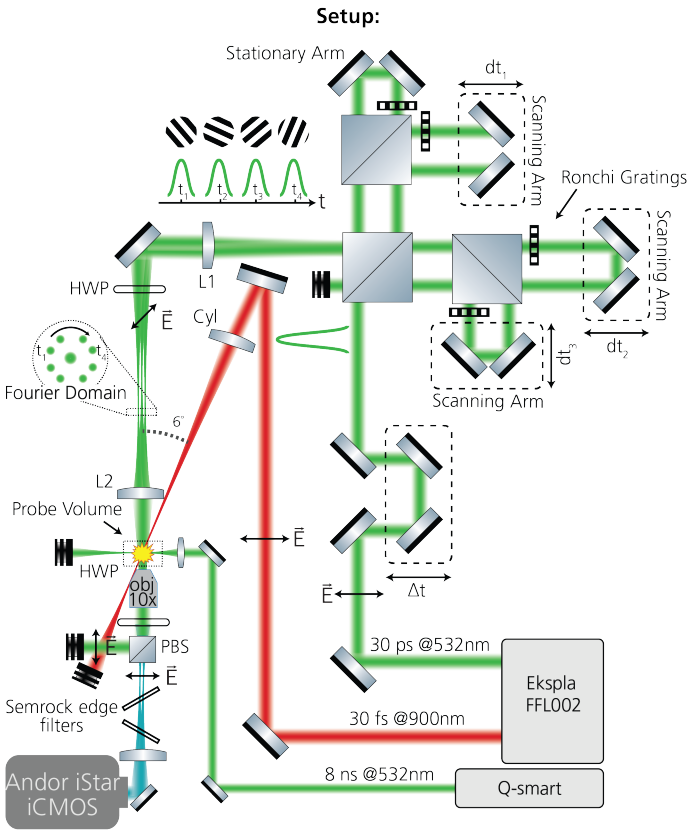


Figure 7.5: Coherence lifetime imaging optical setup.
A 30 fs transform limited (at the probe volume) pulse centred at 900 nm is focused by a cylindrical lens into a sheet on the probe volume where an 8 ns 532 nm pulse is focused such that a laser induced breakdown event is underway. An optically synchronised 30 ps 532 nm pulse is globally delayed before it enters a FRAME pulse shaping stage. Here 2" beamsplitters divide the pulse into four arms where unique spatial intensity modulations and delays are imparted onto the individual pulses before being recombined into a spatially multiplexed pulse train. These pass a 4f imaging setup (L1 and L2) before being incident on the probe volume. Here a CARS signal is generated and imaged through a microscope setup where the Raman shifted light is isolated via polarisation dumping and spectral filtering. The resulting coherence signal is incident on an intensified CMOS camera.

References

[1] The Nobel Prize in Physics 1964.

[2] The Nobel Prize in Physics 2018.

[3] The Nobel Prize in Physics 2023.

[4] Nobel Prize Outreach AB. Illustrated information. `https://www.nobelprize.org/prizes/physics/2009/illustrated-information`, 2009.

[5] Nils Abramson. Light-in-flight recording by holography. *Opt. Lett.*, 3(4):121–123, Oct 1978.

[6] *UNCTAD Digital Economy Report 2019*. Digital Economy Report. September 2019.

[7] Liang Gao and Lihong V. Wang. A review of snapshot multidimensional optical imaging: Measuring photon tags in parallel. *Physics Reports*, 616:1–37, February 2016.

[8] A. A. Jørgensen, D. Kong, M. R. Henriksen, F. Klejs, Z. Ye, Ò B. Helgason, H. E. Hansen, H. Hu, M. Yankov, S. Forchhammer, P. Andrekson, A. Larsson, M. Karlsson, J. Schröder, Y. Sasaki, K. Aikawa, J. W. Thomsen, T. Morioka, M. Galili, V. Torres-Company, and L. K. Oxenløwe. Petabit-per-second data transmission using a chip-scale microcomb ring resonator source. *Nature Photonics*, 16(11):798–802, November 2022. Number: 11 Publisher: Nature Publishing Group.

[9] Benjamin J. Puttnam, Georg Rademacher, and Ruben S. Luís. Space-division multiplexing for optical fiber communications. *Optica*, 8(9):1186–1203, September 2021. Publisher: Optica Publishing Group.

[10] J. M. Senior and S. D. Cusworth. Wavelength division multiplexing in optical fibre sensor systems and networks: a review. *Optics & Laser Technology*, 22(2):113–126, April 1990.

[11] Jonathan N. Tinsley, Maxim I. Molodtsov, Robert Prevedel, David Wartmann, Jofre Espigulé-Pons, Mattias Lauwers, and Alipasha Vaziri. Direct detection of a single

photon by humans. *Nature Communications*, 7(1):12172, July 2016. Number: 1 Publisher: Nature Publishing Group.

[12] Hasselburg: Ein Dorf geht online, February 2012.

[13] Mark A. Linne, Megan Paciaroni, Edouard Berrocal, and David Sedarsky. Ballistic imaging of liquid breakup processes in dense sprays. *Proceedings of the Combustion Institute*, 32(2):2147–2161, 2009.

[14] Takakazu Suzuki, Ryohei Hida, Yuki Yamaguchi, Keiichi Nakagawa, Toshiharu Saiki, and Fumihiko Kannari. Single-shot 25-frame burst imaging of ultrafast phase transition of Ge2Sb2Te5 with a sub-picosecond resolution. *Applied Physics Express*, 10(9):092502, August 2017. Publisher: IOP Publishing.

[15] 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.

[16] Liang Gao, Jinyang Liang, Li Lin, and Lihong Wang. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature*, 516:74–7, 12 2014.

[17] P.C.D. Hobbs. *Building Electro-Optical Systems: Making It All Work*. Wiley Series in Pure and Applied Optics. Wiley, 2022.

[18] Michael Francis Tompsett. Charge transfer imaging devices. `https://patents.google.com/patent/US4085456/en`, April 1978.

[19] Fairchild Semiconductors. Fairchild semiconductors - ccd201. `https://www.tvcameramuseum.org/earlyccd/misc/pix/ccd201adc-datasheet.pdf`, 1974.

[20] Samsung. Samsung Introduces the 200-Megapixel Image Sensor for the Ultimate High Resolution Experience in Flagship Smartphones. `https://news.samsung.com/global/`.

[21] H. I. Bjelkhagen. HOLOGRAPHY, APPLICATIONS | Holographic Recording Materials and Their Processing. In Robert D. Guenther, editor, *Encyclopedia of Modern Optics*, page 50. Elsevier, Oxford, January 2005.

[22] Takeharu Goji Etoh, Anh Quang Nguyen, Yoshinari Kamakura, Kazuhiro Shimonomura, Thi Yen Le, and Nobuya Mori. The theoretical highest frame rate of silicon image sensors. *Sensors*, 17(3), 2017.

[23] Image Sensors Market Size, Share | CAGR of 8.1%. `https://market.us/report/image-sensors-market/`.

[24] Son Dao, T.G. Etoh, Kazuhiro Shimonomura, Anh Quang Nguyen, N. Hayashi, Y. Kamakura, Chao Zhang, E. Charbon, P. Goetschalckx, L. Haspeslagh, and P.De Moor. Toward 10 gfps: Factors limiting the frame rate of the bsi mcg image sensor. 06 2015.

[25] Futa Mochizuki, Keiichiro Kagawa, Shin-ichiro Okihara, Min-Woong Seo, Bo Zhang, Taishi Takasawa, Keita Yasutomi, and Shoji Kawahito. Single-shot 200mfps 5×3-aperture compressive cmos imager. *Digest of Technical Papers - IEEE International Solid-State Circuits Conference*, 58:116–117, 03 2015.

[26] Takeharu G. Etoh and Quang A. Nguyen. Evolution of High-Speed Image Sensors. In Kinko Tsuji, editor, *The Micro-World Observed by Ultra High-Speed Cameras: We See What You Don't See*, pages 81–101. Springer International Publishing, Cham, 2018.

[27] P. Hofmann. *Solid State Physics: An Introduction*. Physics textbook. Wiley, 2015.

[28] Eadweard Muybridge. *Muybridge's Complete human and animal locomotion: all 781 plates from theAnimal locomotion / / by Eadweard Muybridge ; introd. to the Dover edition by Anita Ventura Mozley.* New York: Dover Publications, 1979.

[29] NASA Technical Reports Server (NTRS). *NASA Technical Reports Server (NTRS) 19930091503: The NACA Apparatus for Studying the Formation and Combustion of Fuel Sprays and the Results from Preliminary Tests*. 1933.

[30] H. E. Edgerton, J. K. Germeshausen, and H. E. Grier. High Speed Photographic Methods of Measurement. *Journal of Applied Physics*, 8(1):2–9, January 1937.

[31] C4 Rotating Mirror High Speed Camera bolted to rectangular wheeled steel frame - wheels missing | Science Museum Group Collection. `https://collection.sciencemuseumgroup.org.uk/objects/co8358474`.

[32] Andreas Ehn, Joakim Bood, Zheming Li, Edouard Berrocal, Marcus Aldén, and Elias Kristensson. Frame: femtosecond videography for atomic and molecular dynamics. *Light: Science & Applications*, 6:42, 03 2017.

[33] Jinyang Liang, Liren Zhu, and Lihong V. Wang. Single-shot real-time femtosecond imaging of temporal focusing. *Light: Science & Applications*, 7(1):42, August 2018. Number: 1 Publisher: Nature Publishing Group.

[34] Phillip Reu and Timothy Miller. The application of high-speed digital image correlation. *Proposed for publication in Journal of Strain Analysis.*, 43, 02 2008.

[35] Julien Manin, Scott A. Skeen, and Lyle M. Pickett. Performance comparison of state-of-the-art high-speed video cameras for scientific applications. *Optical Engineering*, 57(12):124105, 2018.

[36] T. Etoh, Son Dao, Kazuhiro Shimonomura, E. Charbon, C. Zang, Yoshinari Kamakura, and Toshimasa Matsuoka. Toward 1gfps: Evolution of ultra-high-speed image sensors -isis, bsi, multi-collection gates, and 3d-stacking-. volume 2015, 12 2014.

[37] Hideharu Mikami, Cheng Lei, Nao Nitta, Takeaki Sugimura, Takuro Ito, Yasuyuki Ozeki, and Keisuke Goda. High-speed imaging meets single-cell analysis. *Chem*, 4(10):2278–2300, 2018.

[38] Christopher K. Haluska, Karin A. Riske, Valérie Marchi-Artzner, Jean-Marie Lehn, Reinhard Lipowsky, and Rumiana Dimova. Time scales of membrane fusion revealed by direct imaging of vesicle fusion with high temporal resolution. *Proceedings of the National Academy of Sciences*, 103(43):15841–15846, 2006.

[39] Lianhao Yin, Marcus Lundgren, Zhenkan Wang, Panagiota Stamatoglou, Mattias Richter, Öivind Andersson, and Per Tunestål. High efficient internal combustion engine using partially premixed combustion with multiple injections. *Applied Energy*, 233-234:516–523, 2019.

[40] Takafumi Hirata and Zen Miyazaki. High-speed camera imaging for laser ablation process: For further reliable elemental analysis using inductively coupled plasma-mass spectrometry. *Analytical chemistry*, 79:147–52, 01 2007.

[41] W.F. Kosonocky, Guang Yang, Chao Ye, R.K. Kabra, Liansheng Xie, J.L. Lawrence, V. Mastrocolla, F.V. Shallcross, and V. Patel. 360/spl times/360-element very-high-frame-rate burst image sensor. In *1996 IEEE International Solid-State Circuits Conference. Digest of TEchnical Papers, ISSCC*, pages 182–183, February 1996. ISSN: 0193-6530.

[42] T. G. Etoh, D. Poggemann, G. Kreider, H. Mutoh, A. J. P. Theuwissen, A. Ruckelshausen, Y. Kondo, H. Maruno, K. Takubo, H. Soya, K. Takehara, T. Okinaka, and Y. Takano. An image sensor which captures 100 consecutive frames at 1000000 frames/s. *IEEE Transactions on Electron Devices*, 50(1):144–151, 2003.

[43] Takeharu G. Etoh, Dung H. Nguyen, Son V. T. Dao, Cuong L. Vo, Masatoshi Tanaka, Kohsei Takehara, Tomoo Okinaka, Harry van Kuijk, Wilco Klaassens, Jan Bosiers, Michael Lesser, David Ouellette, Hirotaka Maruyama, Tetsuya Hayashida, and Toshiki Arai. A 16 Mfps 165kpixel backside-illuminated CCD. In *2011 IEEE International Solid-State Circuits Conference*, pages 406–408, February 2011. ISSN: 2376-8606.

[44] L. Wu, D. San Segundo Bello, P. Coppejans, J. Craninckx, P. Wambacq, and J. Borremans. A 20 Mfps high frame-depth CMOS burst-mode imager with low power

in-pixel NMOS-only passive amplifier. In T. Goji Etoh and Hiroyuki Shiraga, editors, *Selected Papers from the 31st International Congress on High-Speed Imaging and Photonics*, volume 10328, page 1032803. International Society for Optics and Photonics, SPIE, 2017.

[45] T. Goji Etoh, D. Poggemann, A. Ruckelshausen, A. Theuwissen, G. Kreider, H.-O. Folkerts, H. Mutoh, Y. Kondo, H. Maruno, K. Takubo, H. Soya, K. Takehara, T. Okinaka, Y. Takano, T. Reisinger, and C. Lohmann. A ccd image sensor of 1 mframes/s for continuous image capturing 103 frames. In *2002 IEEE International Solid-State Circuits Conference. Digest of Technical Papers (Cat. No.02CH37315)*, volume 1, pages 46–443 vol.1, 2002.

[46] Toshiki Arai, J. Yonai, Tetsuya Hayashida, H. Ohtake, H. Kuijk, and T. Etoh. Back-side-illuminated image sensor with burst capturing speed of 5.2 tpixel per second. *Proceedings of SPIE - The International Society for Optical Engineering*, 8659:04–, 02 2013.

[47] Rihito Kuroda, Shigetoshi Sugawa, and Manabu Suzuki. Over 100 million frames per second high speed global shutter cmos image sensor. page 12, 01 2019.

[48] Vu Truong Son Dao, Nguyen Ngo, Anh Quang Nguyen, Kazuhiro Morimoto, Kazuhiro Shimonomura, Paul Goetschalckx, Luc Haspeslagh, Piet De Moor, Kohsei Takehara, and Takeharu Goji Etoh. An image signal accumulation multi-collection-gate image sensor operating at 25 mfps with 32 × 32 pixels and 1220 in-pixel frame memory. *Sensors*, 18(9), 2018.

[49] Manabu Suzuki, Yuki Sugama, Rihito Kuroda, and Shigetoshi Sugawa. Over 100 million frames per second 368 frames global shutter burst cmos image sensor with pixel-wise trench capacitor memory array. *Sensors*, 20(4), 2020.

[50] Er Qiang Li, Kenneth R. Langley, Yuan Si Tian, Peter D. Hicks, and Sigurdur T. Thoroddsen. Double contact during drop impact on a solid under reduced air pressure. *Phys. Rev. Lett.*, 119:214502, Nov 2017.

[51] Jaka Mur, Fabian Reuter, Jernej Jan Kočica, Žiga Lokar, Jaka Petelin, Vid Agrež, Claus-Dieter Ohl, and Rok Petkovšek. Multi-frame multi-exposure shock wave imaging and pressure measurements. *Opt. Express*, 30(21):37664–37674, Oct 2022.

[52] Gavin Lukasik, Jacob Rogers, Thomas Lacy, and Waruna Kulatilaka. High-speed imaging diagnostics of hypervelocity projectile-water droplet interactions. page LTu5B.5, 01 2022.

[53] Chien Ting, Charles Lancé, Jerome Borsboom, Frits Mastik, Martijn Frijlink, Nico Jong, Michel Versluis, and D. Lohse. Brandaris 128: A digital 25 million frames per

second camera with 128 highly sensitive frames. *Review of Scientific Instruments - REV SCI INSTR*, 74, 12 2003.

[54] Xucai Chen, Jianjun Wang, Michel Versluis, Nico de Jong, and Flordeliza S. Villanueva. Ultra-fast bright field and fluorescence imaging of the dynamics of micrometer-sized objects. *Review of Scientific Instruments*, 84(6):063701, 2013.

[55] Chien Ting Chin, Charles Lancée, Jerome Borsboom, Frits Mastik, Martijn E. Frijlink, Nico de Jong, Michel Versluis, and Detlef Lohse. Brandaris 128: A digital 25 million frames per second camera with 128 highly sensitive frames. *Review of Scientific Instruments*, 74(12):5026–5034, 2003.

[56] Joseph Honour. New high-speed electronic multiframing camera. In Ung Kim, Joon-Sung Chang, and Seung-Han Park, editors, *21st International Congress on: High-Speed Photography and Photonics*, volume 2513, pages 28 – 34. International Society for Optics and Photonics, SPIE, 1995.

[57] Brett R. Lawrence. Review of ULTRANAC high-speed camera: applications, results, and techniques. In Dennis L. Paisley and ALan M. Frank, editors, *22nd International Congress on High-Speed Photography and Photonics*, volume 2869, pages 882 – 887. International Society for Optics and Photonics, SPIE, 1997.

[58] Rainer Hain, Christian Kähler, and Cameron Tropea. Comparison of ccd, cmos and intensified cameras. *Experiments in Fluids*, 42:403–411, 03 2007.

[59] V. Tiwari, Michael Sutton, and S. McNeill. Assessment of high speed imaging systems for 2d and 3d deformation measurements: Methodology development and validation. *Experimental Mechanics*, 47:561–579, 07 2007.

[60] Ryosuke Kodama, P Norreys, Kunioki Mima, A Dangor, Roger Evans, H Fujita, Y Kitagawa, Karl Krushelnick, T. Takeshi Matsuoka, Noriaki Miyanaga, T Norimatsu, S Rose, T Shozaki, Keisuke Shigemori, Atsushi Sunahara, M. Tampo, Kazuo Tanaka, Yusuke Toyama, T Yamanaka, and Matt Zepf. Fast heating of ultrahigh-density plasma as a step towards laser fusion ignition. *Nature*, 412:798–802, 09 2001.

[61] J. Olofsson, M. Richter, M. Aldén, and M. Augé. Development of high temporally and spatially (three-dimensional) resolved formaldehyde measurements in combustion environments. *Review of Scientific Instruments*, 77(1):013104, January 2006. _eprint: https://pubs.aip.org/aip/rsi/article-pdf/doi/10.1063/1.2165569/15917638/013104_1_online.pdf.

[62] Xuanke Zeng, Shuiqin Zheng, Yi Cai, Qinggang Lin, Jinyang Liang, Xiaowei Lu, Jingzhen Li, Weixin Xie, and Shixiang Xu. High-spatial-resolution ultrafast framing imaging at 15 trillion frames per second by optical parametric amplification. *Advanced Photonics*, 2(5):056002, 2020.

[63] Hamamatsu. Guide to streak cameras. https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/sys/SHSS0006E_STREAK.pdf, 2008.

[64] JS Courtney-Pratt. A new method for the photographic study of fast transient phenomena. *Research; a journal of science and its applications*, 2(6):287–294, 1949.

[65] Universal streak camera C16910-01 | Hamamatsu Photonics. https://www.hamamatsu.com/eu/en/product/photometry-systems/streak-camera/universal-streak-camera/C16910-01.html.

[66] Ramanujan Venkata Krishnan, Atsushi Masuda, VE Centonze, and B Herman. Quantitative imaging of protein–protein interactions by multiphoton fluorescence lifetime imaging microscopy using a streak camera. *Journal of biomedical optics*, 8(3):362–367, 2003.

[67] Ali Hosseinnia, Meena Raveesh, Armand Dominguez, Maria Ruchkina, Mark Linne, and Joakim Bood. Single-shot coherent control of molecular rotation by fs/ns rotational coherent anti-stokes raman spectroscopy. *Optics Express*, 30(18):32204–32214, 2022.

[68] Erik Adli, A Ahuja, O Apsimon, Robert Apsimon, A-M Bachmann, D Barrientos, MM Barros, J Batkiewicz, Fabian Batsch, Jeremie Bauche, et al. Experimental observation of proton bunch modulation in a plasma at varying plasma densities. *Physical review letters*, 122(5):054802, 2019.

[69] Andreas Velten, Di Wu, Adrian Jarabo, Belen Masia, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Moungi Bawendi, Diego Gutierrez, and Ramesh Raskar. Femto-photography: capturing and visualizing the propagation of light. *ACM Transactions on Graphics (ToG)*, 32(4):1–8, 2013.

[70] Xiaofang Wang, Lihe Yan, Jinhai Si, Shigeki Matsuo, Huailiang Xu, and Xun Hou. High-frame-rate observation of single femtosecond laser pulse propagation in fused silica using an echelon and optical polarigraphy technique. *Appl. Opt.*, 53(36):8395–8399, Dec 2014.

[71] K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirosawa, H. Liao, T. Ushida, K. Goda, F. Kannari, and I. Sakuma. Sequentially timed all-optical mapping photography (STAMP). *Nature Photonics*, 8(9):695–700, September 2014. Number: 9 Publisher: Nature Publishing Group.

[72] Takakazu Suzuki, Fumihiro Isa, Leo Fujii, Kenichi Hirosawa, Keiichi Nakagawa, Keisuke Goda, Ichiro Sakuma, and Fumihiko Kannari. Sequentially timed all-optical mapping photography (stamp) utilizing spectral filtering. *Opt. Express*, 23(23):30512–30522, Nov 2015.

[73] Takao Saiki, Takuya Hosobata, Yukihiro Kono, Masahiro Takeda, Ayumu Ishijima, Ayumu Ishijima, Miu Tamamitsu, Yutaro Kitagawa, Keisuke Goda, Keisuke Goda, Keisuke Goda, Shin-ya Morita, Shinobu Ozaki, Kentaro Motohara, Yutaka Yamagata, Keiichi Nakagawa, Keiichi Nakagawa, Keiichi Nakagawa, Ichiro Sakuma, Ichiro Sakuma, and Ichiro Sakuma. Sequentially timed all-optical mapping photography boosted by a branched 4f system with a slicing mirror. *Optics Express*, 28(21):31914–31922, October 2020. Publisher: Optical Society of America.

[74] Hirofumi Nemoto, Takakazu Suzuki, and Fumihiko Kannari. Single-shot ultrafast burst imaging using an integral field spectroscope with a microlens array. *Optics Letters*, 45(18):5004–5007, September 2020. Publisher: Optical Society of America.

[75] Mohamed Touil, Saïd Idlahcen, Rezki Becheker, Denis Lebrun, Claude Rozé, Ammar Hideur, and Thomas Godin. Acousto-optically driven lensless single-shot ultrafast optical imaging. *Light: Science & Applications*, 11:66, 03 2022.

[76] Tomoyoshi Inoue, Takashi Kakue, Kenzo Nishio, Toshihiro Kubota, Osamu Matoba, and Yasuhiro Awatsuji. Recent advances in imaging of light propagation with light-in-flight recording by holography. *Ultrafast Science*, 3:0043, 2023.

[77] Aya Komatsu, Yasuhiro Awatsuji, and Toshihiro Kubota. Dependence of reconstructed image characteristics on the observation condition in light-in-flight recording by holography. *J. Opt. Soc. Am. A*, 22(8):1678–1682, Aug 2005.

[78] Takashi Kakue, Kazuya Tosa, Junpei Yuasa, Tatsuki Tahara, Yasuhiro Awatsuji, Kenzo Nishio, Shogo Ura, and Toshihiro Kubota. Digital light-in-flight recording by holography by use of a femtosecond pulsed laser. *IEEE Journal of Selected Topics in Quantum Electronics*, 18(1):479–485, 2012.

[79] Peng Wang, Jinyang Liang, and Lihong V. Wang. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nature Communications*, 11(1):2091, April 2020.

[80] Dalong Qi, Fengyan Cao, Shuwu Xu, Yunhua Yao, Yilin He, Jiali Yao, Pengpeng Ding, Chengzhi Jin, Lianzhong Deng, Tianqing Jia, Jinyang Liang, Zhenrong Sun, and Shian Zhang. 100-Trillion-Frame-per-Second Single-Shot Compressed Ultrafast Photography via Molecular Alignment. *Physical Review Applied*, 15(2):024051, February 2021. Publisher: American Physical Society.

[81] Anna Olofsson, Emma Rose Simpson, Neven Ibrakovic, Samuel Bengtsson, and Johan Mauritsson. Spatial control of extreme ultraviolet light with opto-optical phase modulation. *Opt. Lett.*, 46(10):2356–2359, May 2021.

[82] Xianglei Liu, Jingdan Liu, Cheng Jiang, Fiorenzo Vetrone, and Jinyang Liang. Single-shot compressed optical-streaking ultra-high-speed photography. *Optics Letters*, 44(6):1387–1390, March 2019. Publisher: Optical Society of America.

[83] Xianglei Liu, João Monteiro, Isabela Albuquerque, Yingming Lai, Cheng Jiang, Shian Zhang, Tiago H. Falk, Tiago H. Falk, Jinyang Liang, and Jinyang Liang. Single-shot real-time compressed ultrahigh-speed imaging enabled by a snapshot-to-video autoencoder. *Photonics Research*, 9(12):2464–2474, December 2021. Publisher: Optica Publishing Group.

[84] Bahadir K. Gunturk and Martin Feldman. Frequency division multiplexed imaging. In Nitin Sampat and Sebastiano Battiato, editors, *Digital Photography IX*, volume 8660, page 86600P. International Society for Optics and Photonics, SPIE, 2013.

[85] Shah Rez Khan, Martin Feldman, and Bahadir K. Gunturk. Extracting sub-exposure images from a single capture through fourier-based optical modulation. *Signal Processing: Image Communication*, 60:107–115, 2018.

[86] 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.

[87] Yupan Bao, Karolina Dorozynska, Panagiota Stamatoglou, Chengdong Kong, Tomas Hurtig, Sebastian Pfaff, Johan Zetterberg, Mattias Richter, Elias Kristensson, and Andreas Ehn. Single-shot 3d imaging of hydroxyl radicals in the vicinity of a gliding arc discharge. *Plasma Sources Science and Technology*, 30, 04 2021.

[88] Mattias Richter, Elias Kristensson, Jesper Borggren, Edouard Berrocal, Andreas Ehn, and Marcus Aldén. Simultaneous multispectral imaging of flame species using frequency recognition algorithm for multiple exposures (frame). *Combustion and Flame*, 192:160–169, 12 2018.

[89] Jungho Moon, Seokchan Yoon, Yong-Sik Lim, and Wonshik Choi. Single-shot imaging of microscopic dynamic scenes at 5 thz frame rates by time and spatial frequency multiplexing. *Opt. Express*, 28(4):4463–4474, Feb 2020.

[90] Mark Gragston, Cary Smith, and Zhili Zhang. High-speed flame chemiluminescence imaging using time-multiplexed structured detection. *Applied Optics*, 57:2923, 04 2018.

[91] Mark Gragston, Cary Smith, D. Kartashov, M.N. Shneider, and Zhili Zhang. Single-shot nanosecond-resolution multiframe passive imaging by multiplexed structured image capture. *Optics Express*, 26:28441, 10 2018.

[92] Mark Gragston, Cary Smith, Jacob Harrold, and Zhili Zhang. Multiplexed structured image capture to increase the field of view for a single exposure. *OSA Continuum*, 2, 01 2019.

[93] Hang Li, YaHui Li, BoNan Sun, Kai He, GuiLong Gao, Ping Chen, WenYan Song, Xing Wang, and JinShou Tian. Resolution enhancement via guided filtering for spatial-frequency multiplexing single-shot high-speed imaging. *Optics Express*, 31(21):34074–34087, 2023.

[94] Tomoaki Shibata and Junko Omachi. Single-shot videography with multiplex structured illumination using interferometer. *Optics Express*, 31, 07 2023.

[95] Daniel Richardson. Arrays of glass wedges for multi-dimensional optical diagnostics. *Applied Optics*, 62, 09 2023.

[96] Daniel Richardson. Emission spectroscopy with coded apertures for enhanced dimensionality. *Journal of Applied Physics*, 130:093105, 09 2021.

[97] Joseph W. Goodman. *Introduction to Fourier Optics - Third Edition*. Roberts and Company Publishers, 4950 Yosemite street, Greenwood Village, 2005.

[98] R. N. Bracewell. *The Fourier transform and its applications, second revised edition*. McGraw-Hill Book Company, New York, 1965.

[99] S. Weinstein and P. Ebert. Data transmission by frequency-division multiplexing using the discrete fourier transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, 1971.

[100] Edouard Berrocal, Pettersson Sven-Göran Pettersson, and Elias Kristensson. High-contrast imaging through scattering media using structured illumination and fourier filtering. *Optics Letters*, 41:5612, 12 2016.

[101] Sam Taylor. High contrast imaging through turbid media using Fourier filtering and phase retrieval algorithm , 2022. Student Paper - Master Thesis.

[102] D L Ruderman. The statistics of natural images. *Network: Computation in Neural Systems*, 5(4):517, nov 1994.

[103] Antonio Torralba and Aude Oliva. Statistics of natural images categories. *Network (Bristol, England)*, 14:391–412, 09 2003.

[104] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394, Dec 1987.

[105] G. J. Burton and Ian R. Moorhead. Color and spatial structure in natural scenes. *Appl. Opt.*, 26(1):157–170, Jan 1987.

[106] David J. Field. What Is the Goal of Sensory Coding? *Neural Computation*, 6(4):559–601, 07 1994.

[107] D. J. Field. Wavelets, vision and the statistics of natural scenes. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 357(1760):2527–2542, 1999.

[108] Keju Peng, Xin Chen, Dongxiang Zhou, and Yunhui Liu. 3d reconstruction based on sift and harris feature points. In *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 960–964, 2009.

[109] D. J. Tolhurst, Y. Tadmor, and Tang Chao. Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12(2):229–232, 1992.

[110] Elias Kristensson and Edouard Berrocal. Crossed patterned structured illumination for the analysis and velocimetry of transient turbid media. *Scientific Reports*, 8, 08 2018.

[111] Chengyang Hu, Sigang Yang, Minghua Chen, and Hongwei Chen. Quadrature multiplexed structured illumination imaging. *IEEE Photonics Journal*, 12:1–1, 02 2020.

[112] Lozani Gerdhem, Klara. Saturated planar laser-induced fluorescence using sinusoidal intensity modulation, 2023. Student Paper - Bachelor Thesis.

[113] Daniel R. Richardson. Arrays of glass wedges for multi-dimensional optical diagnostics. *Appl. Opt.*, 62(30):8034–8041, Oct 2023.

[114] Mats Gustafsson. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102:13081–6, 10 2005.

[115] Danica Sugic, Ramon Droop, Eileen Otte, Daniel Ehrmanntraut, Franco Nori, Janne Ruostekoski, Cornelia Denz, and Mark Dennis. Particle-like topologies in light. *Nature Communications*, 12, 11 2021.

[116] Daniel Ehrmanntraut, Ramon Droop, Danica Sugic, Eileen Otte, Mark R. Dennis, and Cornelia Denz. Optical second-order skyrmionic hopfion. *Optica*, 10(6):725–731, Jun 2023.

[117] Eva von Bahr. X. on the quantum-theory and the rotation-energy of molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 28(163):71–83, 1914.

Summary of Scientific publications

Paper I: Beyond MHz image recordings using LEDs and the FRAME concept

The idea of this publication was to characterize and "stress test" a prototype FRAME setup based on LED (incoherent) illumination where we, among other things, showed that FRAME can indeed be performed with other light sources other than pulsed laser illumination. We also attempted to define spatial resolution such that it incorporates a definition of the field-of-view for snapshot imaging. Furthermore, we also defined the temporal resolution for a system with non-Gaussian integration windows.

I performed all experiments, analysed all the data and was responsible for the writing of the manuscript. All authors reviewed the manuscript.

Paper II : Long sequence single-exposure videography using spatially modulated illumination

The idea of this publication was to push the limits of information storage in the context of the FRAME technique. Furthermore, testing and working with a new way of creating pulse trains out of two optical elements: a DOE and a DMD (the presented results would have needed up to 2000 components with conventional designs) was performed. We were able to store up to 1024 images and were able to quantify how the stored images behaved after extraction with respect to spatial resolution and general quality.

I assisted with the experimental hardware control systems. E. Kristensson and I performed the spatial resolution and data storage capabilities analysis and I assisted in the writing of said section. All authors reviewed the manuscript.

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

The idea of this publication was to marry conventional high-speed videography technology, in the form of high-speed cameras, with FRAME based illumination and perform imaging at kHz and MHz rates simultaneously. We were then able to create an in-house algorithm for the extraction of 2D acceleration vectors out of each high-speed camera image such that we could follow the forces acting upon a spray injection over its entire lifetime at kHz rates. This was also the first time FRAME was conducted in microscopy mode.

I designed, built and performed all the experiments. I was responsible for writing the manuscript. All authors reviewed the manuscript.

Paper IV : High-speed videography of transparent media using illumination-based multiplexed schlieren

The idea of this paper was to perform Schlieren imaging by detecting the phase shift of the spatial modulation applied to the illuminating light, a paradigm shift for us considering we always looked at the absolute value of the Fourier information. This, in conjunction with FRAME, opened up for the use of ultrafast videography of refractive index changes within the path of the illumination. Furthermore, experiments showed that this was possible with both coherent and incoherent light and that the most sensitive spatial frequency is three pixels per line pair.

I assisted in the building of the optical setups that were used in the article. I reviewed the manuscript along with all the other authors.

Paper V : Improved temporal contrast of streak camera measurements with periodic shadowing

The idea of this paper was to use periodic shadowing in the same way that it was used previously for spectroscopy, but in conjunction with streak cameras. We did this with two methods, by mounting a Ronchi grating onto the slit and by using a fiber bundle. We showed in this proof of concept that we were able to resolve temporally separated pulses much more efficiently when using periodic shadowing and furthermore we were able to measure lifetimes of Rhodamine such that they fit previous experimental data much more closely.

I helped Y. Bao perform the single and double pulse experiments. I assisted in the analysis and the writing of the manuscript. All authors reviewed the manuscript.

Paper VI : The space-charge problem in ultrafast diagnostics: an all-optical solution for streak cameras

The idea of this paper was to push periodic shadowing in conjunction with streak cameras to its extremes. We performed single pulse and pulse train characterisation measurements with a femtosecond laser source aimed directly at the streak camera slit where a Ronchi grating was either mounted or imaged. We were able to specifically pinpoint that it is the space-charge effect that we are addressing when performing the periodic shadowing extraction routine. The results included increased dynamic range and higher temporal resolutions. With this work, we introduced a new, fully optical approach, for the alleviation of a purely electro-optic problem.

Y. Bao and I performed all the measurements and I performed the analysis. I was responsible for writing the manuscript with technical input from all authors. All authors reviewed the manuscript.

Paper VII : A versatile, low-cost, snapshot multidimensional imaging approach based on structured light

The idea of this paper was to work with FRAME on the passive side, i.e., impart the spatial modulations on the light that has been emitted by the object. By doing this in four individual channels and then choosing the type of filtering to be performed within each channel, colour-, polarisation-, depth- and temporal sensitivity were attained.

I specifically built the colour illumination setup for the temporally resolved measurements (section 3.3). I then performed those measurements, analysed that data and wrote that paragraph of text. All authors reviewed the manuscript.

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

The idea of this paper was to doubly multiplex the acquired images by imparting a spatial modulation on the illumination and then on the received image, i.e., FRAME was performed both actively and passively simultaneously. This was performed in the context of

fluorescence imaging. The result was a new type of structure in the Fourier domain with a total of 24 unique images stored in a snapshot (four illuminations and four spectral filters and their cross terms). This allowed for the high-fidelity separation of nine fluorophores with extremely closely spaced spectra, something that we could not find anybody to have done within a snapshot. We even showed that the fidelity of this separation stays high at low SNRs.

I assisted in the building of the experimental setup and the experiments. I performed the analysis of classification probability with respect to SNR and assisted in the writing of the manuscript. All authors reviewed the manuscript.

Paper I

SCIENTIFIC REPORTS
nature research

OPEN

Beyond MHz image recordings using LEDs and the FRAME concept

Vassily Kornienko[1]✉, Elias Kristensson[1,3], Andreas Ehn[1,3], Antoine Fourriere[2,3] & Edouard Berrocal[1,3]

Many important scientific questions in physics, chemistry and biology rely on high-speed optical imaging techniques for their investigations. These techniques are either passive, relying on the rapid readout of photoactive elements, or active, relying on the illumination properties of specially designed pulse trains. Currently, MHz imaging speeds are difficult to realize; passive methods, being dictated by electronics, cause the unification of high spatial resolution with high frame rates to be very challenging, while active methods rely on expensive and complex hardware such as femto- and picosecond laser sources. Here we present an accessible temporally resolved imaging system for shadowgraphy based on multiplexed LED illumination that is capable of producing four images at MHz frame rates. Furthermore as the LEDs are independent of each other, any light burst configuration can be obtained, allowing for instance the simultaneous determination of low- and high speed events in parallel. To the best of the authors' knowledge, this is the fastest high speed imaging system that does not rely on pulsed lasers or fast detectors, in this case reaching up to 4.56 MHz.

High-speed imaging and videography is important for a plethora of measurement situations such as for investigating chemical reactions[1], fuel injection[2,3], combustion[4,5], plasma formation[6,7] and laser induced material damage[8–10]. The first instance of kHz imaging (2 kHz) can actually be traced back to 1927 when Edward G. Beardsley illuminated a fuel spray from an oil engine injection with 20 individual sparks in the Langley Research Center of the National Advisory Committee for Aeronautics (NACA, the precursor of NASA)[11]. This being top secret work, 10 years later, Edgerton and Germeshausen were able to achieve a similar feat at a speed of 6 kHz, applying their work to filming both sprays and aircraft propellers[12]. During the early 1980s, before the dawn of digital storage, high-speed video cameras used either analog recording in magnetic tape (Kodak Spin Physics SP2000) or rotating mirrors imaging the event on photographic film. In 1991 Takeharu Etoh and Kodak invented the Ektapro line of cameras, of which the first one was capable of imaging at a speed of 40×10^3 frames per second (fps) for 64×64 pixels (a partial readout scheme on 256×256 pixels)[13]. This has opened the doors for the development of fast digital passive imaging technology, i.e. imaging where light emitted from an event is recorded on a 2D array of photoactive elements[14]. Applications of this camera include high-speed imaging of diesel fuel sprays[15] and obtaining particle image velocimetry (PIV) maps of the flow within an engine cylinder[16].

The next major improvement within passive high-speed imaging came with the in situ storage image sensors (ISIS), where the images are stored on the CCD chip before reading out. The temporal resolution of such sensors is thus limited by the electron transfer speed between CCD elements instead of the bottleneck readout speed. However, this also means that the video length cannot be longer than the number of storage elements that are reserved for a given pixel. The first burst image sensor of this kind had a frame count of 30 frames, a pixel count of 360×360 and a speed of 1 Mfps[17]. The frame rate culmination of ISIS technology came in 2005, upon the creation of a CCD with a readout rate of 100 Mfps and a sequence depth of 16 frames[18]. However, to achieve this performance the image size was reduced to 64×64 pixels. Currently available state-of-the-art high-speed cameras can provide ~ 1 Mpixel pixel resolution at Mfps speeds[19,20] and their applicability is essentially only limited by their cost.

In general for passive high-speed digital imaging, the necessity of in situ storage makes it difficult to attain high pixel numbers. This necessity leads to a trade-off between frame rate, spatial resolution and number of images (sequence depth). Silicon sensors also have an intrinsic theoretical speed limit of 90 Gfps caused by the mixing of electrons originating from the varying penetration depths of different wavelengths[21]. However, this theoretical limit has thus far not been achieved, partially due to another challenge associated with high-speed

[1]Department of Combustion Physics, Lund University, Lund, Sweden. [2]Federal Institute for Geosciences and Natural Resources (BGR), Hanover, Germany. [3]These authors contributed equally: Elias Kristensson, Andreas Ehn, Antoine Fourriere and Edouard Berrocal. ✉email: vassily.kornienko@forbrf.lth.se

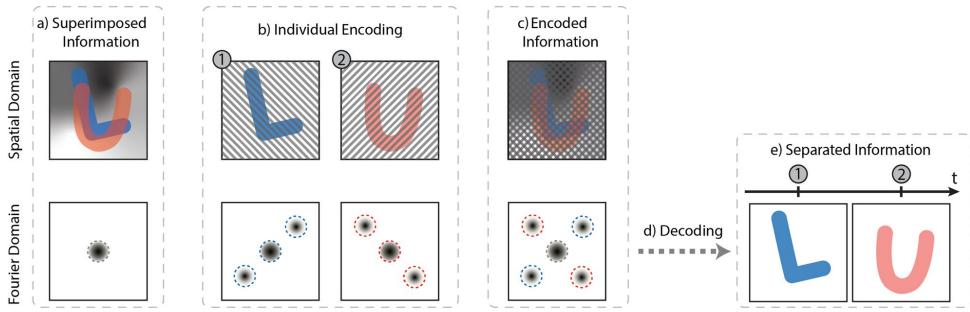


Figure 1. The working principle of FRAME. With no applied spatial modulation, all the image information (the L and U along with an independent background light field) is superimposed and centered in the Fourier domain (**a**). Upon application of unique spatial modulations, the individual image information is shifted into empty areas of the Fourier domain (**b**), allowing for the separation of the timestamped snapshots taken within one exposure (**c**). Applying a spatial lock-in algorithm followed by a low pass filter (**d**), the individual images can then be extracted, with a suppressed background light field (**e**).

imaging: as imaging speeds increase, passive imaging techniques fall victim to an insufficient amount of light incident on the sensors during the short exposure time. An increased sensor sensitivity could compensate for the reduction in the signal level yet with currently available technology this would imply increasing the photoactive area, yielding lower frame rates. A possible way to circumvent these problems and still attain high imaging speeds is to apply computational imaging approaches to regular cameras. Using an off-the-shelf cMOS camera in conjunction with a DMD and galvoscanner, the COSUP technique has demonstrated just this, with a 1.5 Mfps video comprising of 500, 0.5 Mpixel frames[22]. However, due to mechanical constraints on the galvo, higher speeds have not been demonstrated without incorporating another level of complexity in the form of a streak camera[23]. Thus, the development of faster imaging systems has steered into laser illumination-based (*active*) imaging concepts[24–28]. In contrast to passive high-speed imaging, active techniques illuminate 2D transient events with specially designed pulse trains in order to capture the dynamics. This shift to the illumination side allows for a boost in signal level and the ability to surpass the detector speed limit. However, all current demonstrations of these ultrafast methods require either pico- or femtosecond laser systems and thus cannot be scaled down to the MHz level without significantly altering their optical design. Hence the technology currently available for high speed imaging in the MHz regime mainly relies on systems—passive or active—that use advanced and/or expensive hardware.

In this paper we present an accessible LED-based fast illumination system, which, in combination with off-the-shelf camera technology, can image at a speed of 4.56 Mfps (limited by the pulse duration of current LED technology) while still maintaining both high spatial resolution and sensitivity.

Method

The illumination system presented here is a back-lit (transillumination) configuration based on the FRAME (Frequency Recognition Algorithm for Multiple Exposures) technique[27]. A time-resolved back-lit imaging system captures a snapshot of the light-obstruction (shadow) caused by the probed object, which casts a silhouette image upon a camera sensor. The presented FRAME-based configuration allows for a rapid succession of four such images to be acquired within a single camera exposure. By applying unique spatial modulations to the individual pulses of the pulse train, the shadow of the object at each instance in time becomes tagged during the light-matter interaction. These timestamped shadow images are then extracted using a spatial lock-in algorithm on the acquired data.

Frequency recognition algorithm for multiple exposures. Natural images are characteristically sparse in their spatial frequency (Fourier) domain, only utilizing a small fraction of the reciprocal space available to high resolution cameras (Fig. 1a). By applying a spatial modulation to an image, its information can be shifted into these otherwise empty areas of the Fourier domain (Fig. 1b), forming what is referred to as a Fourier cluster. When a set of uniquely spatially modulated images are recorded on the same sensor, the Fourier transform of the recorded image will thus exhibit several Fourier clusters (Fig. 1c). Since the spread of each cluster is small, the respective image information they carry does not overlap. A single photograph can thus contain a full set of images—a multiplexed image series. Using a lock-in algorithm followed by a low-pass filter[27,29,30] (Fig. 1d), the individual clusters can be isolated and the image information contained therein extracted (Fig. 1e). Thus, by implementing this spatial coding strategy, one can store multiple temporally separated observations of an event in a single exposure and extract them into an image sequence.

When FRAME is implemented for high-speed imaging, the coded information corresponds to a pulse train of uniquely spatially modulated light pulses probing an event. The event is then recorded in its entirety on a camera within a single exposure, i.e. all the coded image information is superimposed on the sensor. Since each pulse

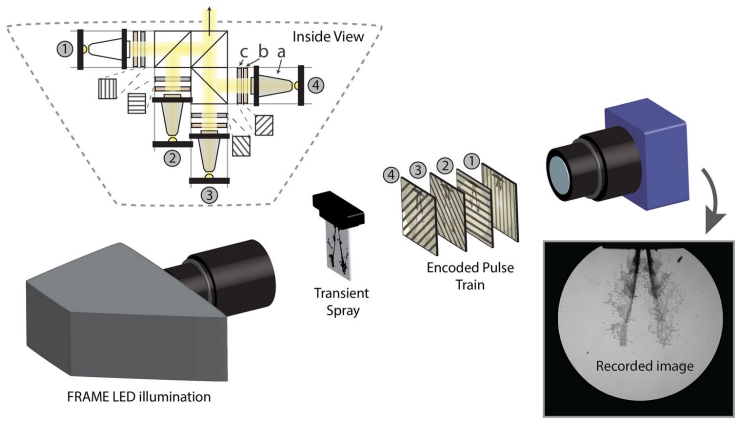


Figure 2. Recording an injection event with FRAME. The light from four pulsed white-light LEDs are led through individual optical arms consisting of a tapered light pipe with 3× magnification (**a**), a diffusor (**b**) and a Ronchi grating (**c**). Beamsplitters recombine these into a collinear pulse train. After passing through an objective lens, they illuminate the sample with a spatially modulated pulse train (indicated with numbers). The pulse train is subsequently recorded on a camera of choice within a single exposure.

interacted with the sample at different times (the "L" at time 1 and the "U" at time 2 w.r.t. Fig. 1), their respective Fourier clusters will correspond to a timestamped snapshot of the event. The attainable temporal resolution of the imaging system has thus been shifted from the electronics of the camera to the illumination characteristics of the pulse train. As a result, high speed imaging without the constraint on camera pixel number or sensitivity is made possible. In conjunction with a femtosecond laser source, this technique has demonstrated a 5 Tfps image sequence of a light pulse traveling through a Kerr medium[27]. Although Tfps speeds can be attained, ultrafast laser technology is not a fundamental necessity for FRAME to function. Indeed, the application of spatial modulations, paramount to the functioning of FRAME, is independent of light source. This choice, be it e.g. narrowband lasers or broadband LEDs, can be completely determined by the type of event to be probed. Furthermore the tagging is not solely limited to the temporal domain but also for the tagging of any image dimension, opening up for snapshot multispectral-[29], polarization-[29] and 3D-imaging[29,30].

This method of tagging images also carries with it another useful property, namely background light suppression. If the background grayscale light field of Fig. 1a is independent of the image information (independent of the L and U, as in Fig. 1b), it will remain unmodulated. The low-pass filter will hence effectively eliminate it, allowing for a higher signal to background ratio in the final extracted images (Fig. 1e and Supplementary Information 6).

Experimental setup. The aim of the presented work is to demonstrate the possibility of achieving high-resolution MHz imaging, without relying on advanced and expensive hardware such as time-gated detectors, streak cameras and/or pulsed lasers. Instead, the illumination system presented herein (a prototype of the FRAME Illumination Unit developed in collaboration with LaVision GmbH) consists of four Ce:YAG white light LEDs (Cree XHP35) that can be switched on and off on a sub-microsecond time scale (rise time of 205 ns, from 10 to 90% of full power, and temporal jitter FWHM of 34.5 ns—see Supplementary Information 7). Each LED is controlled individually, which not only allows for adjustable frame rates from 10^0 to 10^7 fps but also for more complex non-linear triggering sequences and variable brightness. Furthermore, by implementing the FRAME image-coding methodology, there is no need for fast image readout, which allows us to capture the transillumination frames using slow, yet sensitive high-resolution detectors. This is demonstrated here using a standard 10-bit CMOS 5 MP camera (LaVision, Imager M-lite) mounted with a telecentric objective lens (Edmund Optics Gold TL 1X or 0.5X).

The optical configuration of the system is schematically illustrated in Fig. 2. The light from each LED is guided along an optical arm consisting of a tapered light pipe with 3× magnification (a), a diffusor (b), and a Ronchi grating (50 lp/mm) rotated to a unique angle (c). The four spatially encoded white light pulses are thereafter combined into a collinear pulse train using beam splitters. A Nikon 50–300 mm camera objective is used to image the gratings onto the sample, permitting the magnification of the illuminated area—the viewed area—to be readily modified. In our study a commercially available fuel port injector, Bosch EV1 4-holes nozzle[31], with orifice size of 280 µm, running at 4.8 bar water injection pressure, was probed at different times after start of injection (ASOI). Its complex transient structures allows us to demonstrate the ability of our system in obtaining images with both high spatial- and temporal resolution.

During an acquisition, the spray is illuminated with the four intensity-modulated pulses and the transmitted light from all pulses is summed on the camera sensor within a single exposure (Fig. 3a). The magnifications in

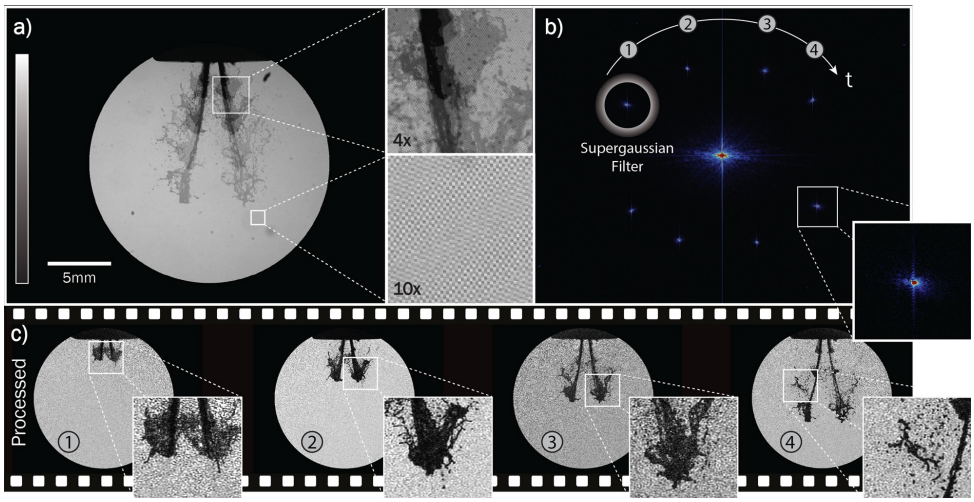


Figure 3. The extraction process of a 20 kHz image series. (**a**) The single exposure (raw) image with the four uniquely modulated pulses summed onto the camera sensor. The crosshatch pattern of the superimposed modulations is highlighted in the 10× zoom. (**b**) The Fourier transform of (**a**) where each Fourier cluster holds the image information from a single pulse. (**c**) The extracted, background-corrected image series, where the zoom-ins highlight the visibility of fine structures.

Fig. 3a show the resulting crosshatch pattern of the overlapping modulated structures (10× zoom) and the various grayscale shadows of the spray at different times as the pulse train illuminates it (4× zoom). It can be seen that the dynamic range of the sensor has been divided among the four illuminating pulses, resulting in gradually lighter shadows in the pixel positions the spray covers as it evolves. Here, in the spatial domain, the image information for the individual events overlap and cannot be accessed. However, the angle and frequency of the Ronchi gratings have been chosen such that their corresponding Fourier clusters are placed as far away from adjacent clusters and recurrent/static structures in reciprocal space as possible (Fig. 3b). The lock-in algorithm subsequently shifts each cluster to the center of the Fourier domain, where the data is low-pass filtered (2D-Gaussian to the power of 8) and finally inverse Fourier transformed back to the spatial domain. The images are then background corrected, resulting in the final temporally resolved image series (Fig. 3c).

Spatial resolution. The spatial resolution of an imaging system can be deduced by measuring its modulation transfer function (MTF). In practice, this involves imaging a Sector Star target (Siemens target), from which the system's cut-off spatial frequency is extracted, i.e. the number of line-pairs per millimeter (lp/mm) the system can *spatially resolve* in an arbitrary direction. According to the Rayleigh criteria, a MTF contrast of 15.3% is required to resolve two neighboring points sources and is often the value used to define the spatial resolution of an imaging system. Specifying the system's spatial resolution alone is, however, not enough to assess its imaging capabilities; it is equally meaningful to specify the viewing area or field-of-view (FOV) over which the measured spatial resolution is maintained. Here we define two parameters that take into account the relevant information concerning spatial resolution: (1) the maximum number of resolvable line-pairs per unit of length (*RLP*) and (2) the total number of resolvable line-pairs (*NLP*) across the FOV where the resolution, *RLP*, is maintained. In the text we will combine the two parameters, *RLP* and *NLP*, in a condensed form specified accordingly: $(RLP{:}NLP)_X$, where X indicates that the numbers refer to the X dimension. Using this nomenclature, an image with a spatial resolution of 60 lp/mm and a FOV of 1 cm in the y-direction would thus be written as $(60\ \text{lp/mm}{:}600)_Y$.

Temporal resolution. In contrast to most passive high speed imaging techniques, image series acquired using illumination-based systems can have frames that overlap in time, i.e. an acquisition rate of frames exceeding the inverse of the frame duration (integration time). The parameter governing the system's temporal resolution—interframe time or frame duration—may therefore vary depending on the type of system[32] and, hence, it becomes useful to clarify the term "temporally resolved" for high-speed imaging, in particular for illumination-based systems.

The purpose of all video systems is to extract a sequence of images that shows the temporal evolution of an event. Most often this means sampling the time axis with image frames at discrete points, where each frame has a set integration time (Δ). Even if illumination-based imaging systems have the ability to overlap image frames in time, according to the Nyqvist samling rate theorem[33] this approach does not provide additional information about the event's temporal evolution. To avoid such oversampling of data, the frames must instead be arranged

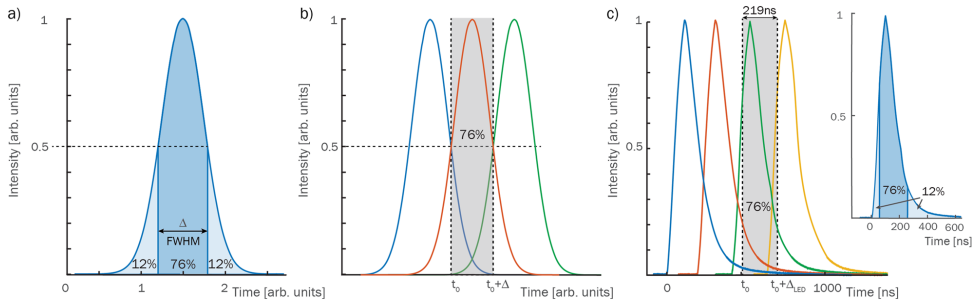


Figure 4. Temporal resolution of an imaging system. (**a**) A Gaussian gate profile has 76% of its power concentrated within its FWHM. Placing individual Gaussian pulses in a pulse train such that their half maxima limits coincide, (**b**), assures that 76% of the information generated by, for example, the red pulse is temporally unique to the time interval $[t_0, t_0 + \Delta]$, where $\Delta = FWHM$. Using this Δ as the discretizing step size on the time-axis, we consider the resulting pulse train *temporally resolved*. (**c**) Generalization of this concept to the temporal gate profiles of the LED pulses produced by the system presented herein leads to temporal limits within which 76% of the pulse's power resides. Discretizing the time-axis with these limits (corresponding to a step size of $\Delta_{\mathrm{LED}} = 219$ ns) results in a maximum imaging speed, or maximum sampling rate, of 4.56 MHz.

so that they are temporally separated. In order to determine when two or more frames are considered temporally separated, we propose an analysis based on image frames with Gaussian temporal gate functions. 76% of the image information contained within such a frame can be traced to the time-slot bounded by its FWHM (the Δ of Fig. 4a). Thus, by arranging these such that their FWHMs coincide (Fig. 4b), the time axis becomes discretized such that each sampled time-slot is associated with a given image to a degree of 76%. This strategy of sampling the time-axis—discretizing it with a step size of Δ—guarantees a 76% degree of "temporal uniqueness" to each image, it is generalizable to any temporal gate profile and avoids oversampling of the event. The inverse of the step size, $1/\Delta$, is then the maximum attainable sampling rate of the event's temporal evolution, i.e. the maximum imaging speed of the system.

The temporal profiles of the frames acquired using illumination-based imaging systems can be determined from the illuminating pulse train. For the present system, we measure the shortest attainable such profile of the LED pulses using a photodiode and 6 GHz oscilloscope (inset of Fig. 4c)—this is what will set maximum the sampling rate of the system. Here, equivalently to the Gaussian profiles, an upper and lower limit has been extracted within which 76% of the pulse power is concentrated. This Δ_{LED} of 219 ns is used as the discretizer of the temporal evolution of the event, resulting in a maximum temporal imaging speed of 1/219 ns = 4.56 MHz associated with the system presented herein (see Supplementary Information 4 for more measurement examples).

As for the description of spatial properties, the description of the temporal properties of an acquired image series requires, at a minimum, (1) the sequence length (*SL*), specified in time and (2) the sequence depth (*SD*), i.e. number of *temporally resolved* images over the time *SL*. To specify these in a condensed form, the same notation as earlier will be used for temporal properties: $(SL : SD)_T$ (where T indicates that the numbers refer to the time dimension). Using this nomenclature, an image series consisting of 13 temporally resolved frames acquired within 1500 µs would thus be written $(1500\ \mu s{:}13)_T$. For the system presented herein, the sequence depth is constant at four frames and the shortest presented sequence length is 876 ns, i.e. $(876\mathrm{ns}{:}4)_T$, corresponding to 4.56 Mfps (Fig. 4c).

Results

Unlike most high speed imaging systems, the current system can achieve frame rates in the MHz-regime without compromising FOV or sensitivity. Furthermore, since the illumination dictates the properties of the time axis, the image series can be tailored from the Hz to the MHz regime or a mixture of both. Variable magnification is also a possibility and dependent solely on the optics chosen. Here we present measurements that illustrate these capabilities.

Temporally resolved imaging at kHz to MHz frame rates. A set of single-exposure images was acquired at different speeds and FOVs using the FRAME illumination system. Upon extraction, these resulted in multiple image series of the event, with speeds ranging from 10 kHz to 4 MHz (Fig. 5).

Since the pulse duration for the LEDs is variable and each is controlled individually, the illumination can be optimized with regards to the sample under study. The atomization process studied herein required ~ 1 µs in order to be temporally frozen. Increasing the duration of the illumination boosts the signal level but generates motion blur (for the current sample), whereas imaging at frame rates above 1 Mfps requires a reduction of the illumination time.

For a FRAME sequence consisting of frames with no temporally overlapping information, the total dynamic range of the sensor will be divided equally among each unique time-slot. In the current case, this would yield $2^{10}/4 = 256$ temporally unique grayscale levels per image. However, according to the definition stated above,

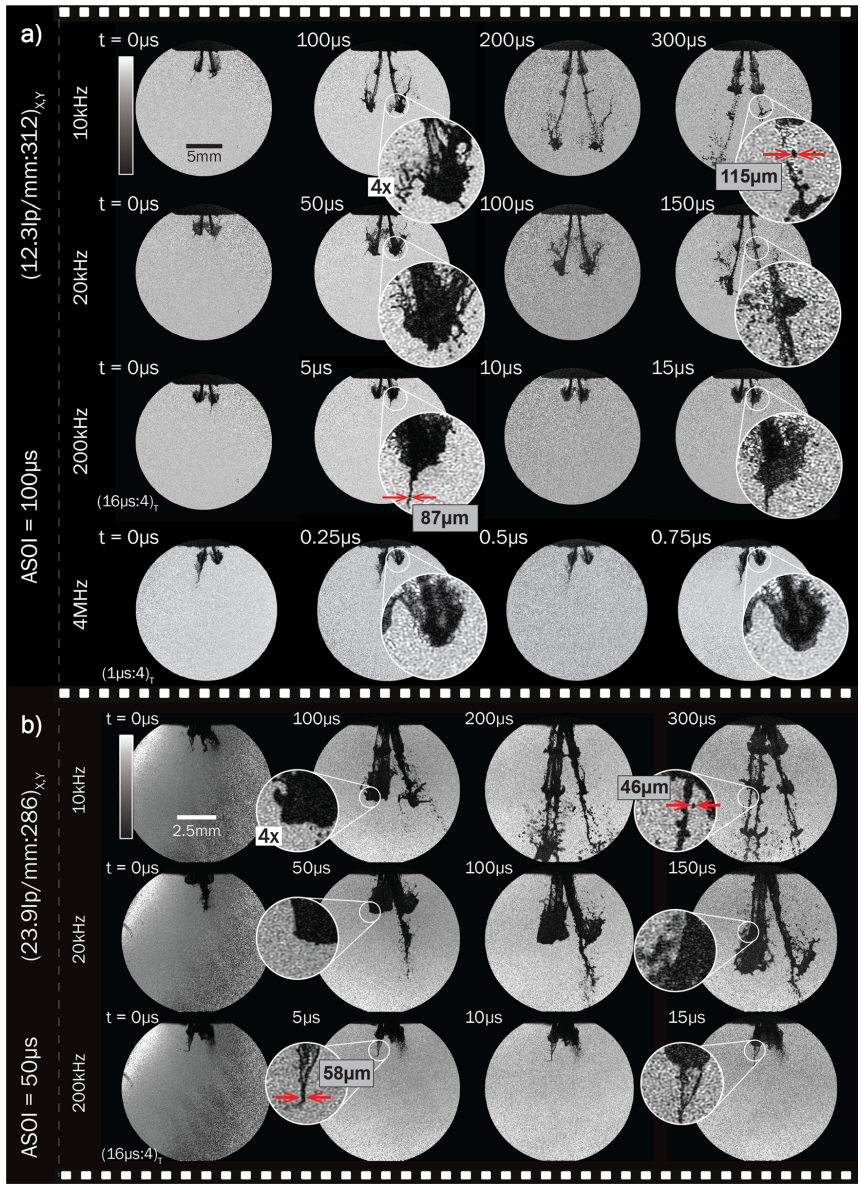


Figure 5. The versatility of speeds and magnifications allowed by the system. (**a**) shows a collection of series taken 100 µs after start of injection (ASOI) with a large FOV at speeds ranging from 10 kHz to 4 MHz. (**b**) is a collection of series taken at 50 µs ASOI and small FOV. The zoom-ins depict identical positions at different times, emphasizing the movement of the injection front at the given timestamps.

a sequence can still be temporally resolved even if the frames partially overlap in time—at the temporal resolution limit there will be a 24% information overlap within each time-slot. In the case of the 4 MHz image series

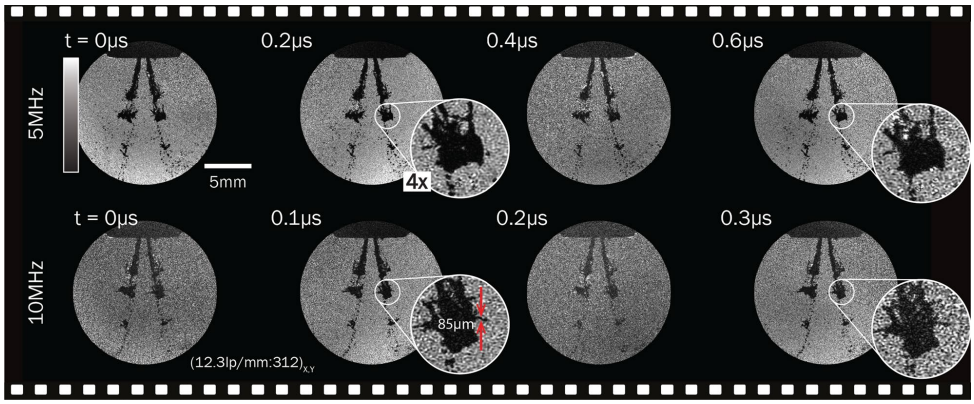


Figure 6. The preservation of FOV at high speeds. A set of image series of the injection event taken at speeds of 5 and 10 MHz, 2.3 ms ASOI. Fine structures are still equally visible for these higher speeds. Measurements on a Sector Star target show no degradation of the spatial resolution at these frame rates, i.e. $(12.3\ \text{lp/mm:}312)_{X,Y}$ is preserved.

of Fig. 5, this fact results in a decrease to 194 temporally unique grayscale levels per image, while the complete lack of overlap between the frames in the 10kHz series, results in no decrease of dynamic range per image (see Supplementary Information 8).

The presented set of measurements shows the possibility of probing a given event at various frame rates, allowing for easy access to many different types of dynamic samples. Furthermore, the ease with which the frame rate is interchangeable, by simply changing LED trigger times, allows the temporal sampling of the time-axis to be tailored to fit the object under study.

Preservation of the field of view. For the measurements presented herein, an average spatial resolution for the individual extracted frames was measured under the Rayleigh criteria (contrast of 15.3%), resulting in 12.3 lp/mm and 23.9 lp/mm for the large and small FOVs respectively, i.e, $(12.3\ \text{lp/mm:}312)_{X,Y}$ and $(23.9\ \text{lp/mm:}286)_{X,Y}$ (see Supplementary Information 2 for the measurements).

As has been mentioned earlier, the fixed pixel readout rate causes many passive techniques to suffer from a trade-off between speed and field of view. Since the videography speed of FRAME is completely determined by the illumination properties, the spatial resolution and FOV can be preserved independently of frame-rate. As a proof of concept two separate pulse trains consisting of four 200 ns pulses and four 100 ns pulses forming *non-temporally* resolved 5 and 10 MHz image series respectively, show the visibility of fine structures (Fig. 6). Indeed, measurements performed on a Sector Star target confirm that the system's FOV and spatial resolution does not degrade with speed, even when exceeding the limit of the system's maximum frame rate (see Supplementary Information 3 and 5).

Various timestamp configurations. Accurately tracking the evolution of a transient event requires the ability to visualize both the fast- as well as the slow moving parts of the sample. This information can be attained through either long image sequences or a non-linear interframe time. The system's versatility in the time domain allows for arbitrary frame timestamps, granting parallel access to events that occur in the regimes ranging from 10^0 to 10^{-7} s. This versatility is illustrated with a series of images with a constant inter-frame time (Fig. 7a), a PIV-inspired sequence (Fig. 7b.) and a series with an accelerating frame rate (Fig. 7c).

The red lines of Fig. 7a show the progression of the injection front as a function of time, superimposed on each other as the image series proceeds. The insets in Fig 7b depicts the absolute difference between the two adjacent images where the green shading emphasizes the evolution of the injection. Using this PIV arrangement, the speed on the injection front was approximated to about 40 m/s at both of the indicated times. The insets labeled "difference" in Fig. 7c depict the evolution of the front for two different inter-frame times, where the blue regions temporally precedes the red ones, highlighting the direction of the flow. In the second inset, these regions are reduced to the pixel or sub-pixel level, thus validating the ~ 1 μs time needed to temporally freeze the flow for this injector. Given the images' spatial resolutions of $(12.3\text{lp/mm:}312)_{X,Y}$, and a high shadowgraphy contrast, this approach of varying the temporal properties of the pulse train allows for a thorough quantitative investigation into the flow dynamics of the injection.

Discussion

In summary, we have presented a FRAME illumination system based on fast pulsed LED technology, allowing for high spatial resolution transmission imaging on the MHz timescale. Each individual LED illuminates a high frequency transmission grating with different orientations, in order to spatially code the individual light pulses.

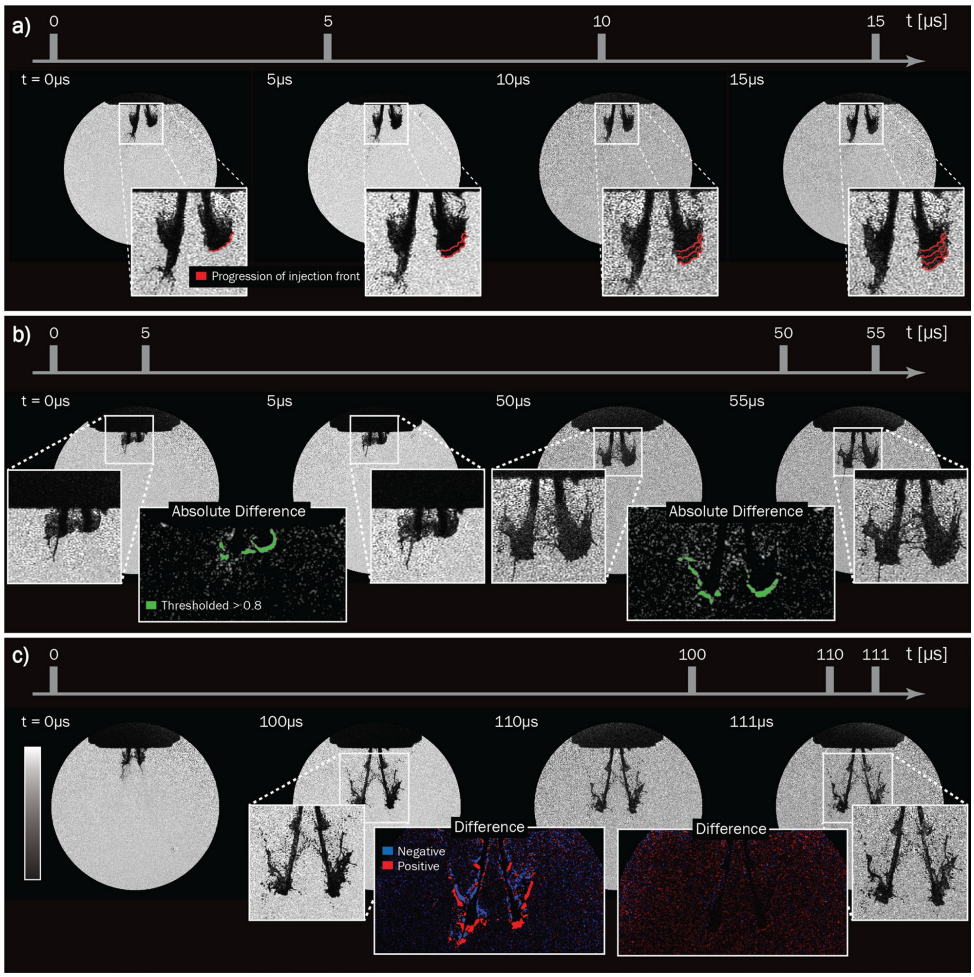


Figure 7. Illustration of the versatility in the time domain. (**a**) Shows a series depicting the leading edge of the injection (red) tracked at a constant inter-frame time. (**b**) Depicts a PIV-inspired event where two fast images were taken within 5 µs followed by a delay of 45 µs before another double trigger event was acquired. One can see that the green, representing the thresholded absolute difference between the images to its left and right, is approximately the same size for the two pairs of light pulses, corresponding to an injection front speed of about 40 m/s. (**c**) Shows an accelerating frame-rate image series where the time difference between the first two images is 100 µs followed by a difference of 10 µs and then 1 µs. The blue regions in the insets precedes the red ones in time, making the direction of the flow clearly visible. One can also observe that in the last inset the difference is reduced to a pixel or subpixel level, indicating that the atomization process indeed occurs on the ~ 1 µs timescale.

These periodic patterns are then optically relayed onto the sample and the transmitted light recorded in parallel by the sensor. Because all pulses are recorded in one exposure, there is no need for fast readout, allowing the illumination system to be combined with, in principle, any camera. To the best of the author's knowledge, this is the fastest system capable of MHz imaging without relying on either pulsed lasers or fast detectors.

The optical configuration has several technical advantages over high speed imaging systems based on pulsed lasers or fast detectors. First, since the LEDs are controlled individually, the pulse train can be arranged non-linearly in time and thereby adapt the interframe time to the temporal evolution of the sample. This possibility

can compensate for the currently relatively low sequence depth. Second, contrary to pulsed lasers otherwise employed with illumination-based high-speed imaging systems, LED illumination is harmless, less expensive and compact. Third, FRAME naturally provides suppression of background light and multiple light scattering, allowing for shadowgraphy imaging of objects that emit light or are hidden in turbid media. Moreover, the use of incoherent light is beneficial for transmission imaging as speckle patterns and diffraction effects are avoided. Finally, unlike most passive-based high-speed imaging techniques, the high pixel density and field of view is maintained even at the highest frame rates.

Currently the maximum temporal performance of the presented system has been shown to be $(876\text{ns}{:}4)_T$, i.e. 4.56Mfps. Image sequences acquired at even faster rates—5 and 10 MHz—were shown to maintain the spatial performance of either $(12.3\text{lp/mm}{:}312)_{X,Y}$ or $(23.9\text{lp/mm}{:}286)_{X,Y}$. The illumination system will hence directly benefit from any technological advancement towards faster and/or brighter LEDs, opening up for even shorter sequence lengths while still retaining the original FOVs. Brighter LED sources may also allow for longer sequence depths, given that the beam-splitter arrangement is upgraded to facilitate more pulses. Such technological advancement is not unrealistic, as the brightness vs cost of LED technology has been observed to progress in an exponential manner akin to Moore's law for transistor size—the LED counterpart being dubbed Haitz law[34].

For the sake of clarity we have chosen to illustrate the capabilities of the system with a fuel port injector, as it combines fast events with fine spatial details. However, as long as the spatial modulations can be imaged effectively, this system can be applied in other transmission imaging situations, such as flow visualization[35], shockwave imaging[36] and laser ablation[37].

The development towards high-speed video cameras that are faster while still providing high spatial resolution is indicative of the growing need to visually track the motion of objects on a sub-microsecond timescale. However, at such (or faster) imaging speeds, few objects have a sufficient spontaneous luminosity needed for visualization, with the exception of high-speed imaging of laser pulses[25,27,38,39]. The presented work demonstrates a new opportunity for videography on the sub-microsecond timescale, uniting slow, high resolution cameras with illumination-aided imaging concepts, allowing for the use of accessible imaging hardware to facilitate investigations of high-speed events.

Received: 18 June 2020; Accepted: 22 September 2020
Published online: 06 October 2020

References

1. Zewail, A. H. Femtochemistry: atomic-scale dynamics of the chemical bond. *J. Phys. Chem. A* **104**, 5660–5694. https://doi.org/10.1021/jp001460h (2000).
2. Crua, C., Shoba, T., Heikal, M., Gold, M. & Higham, C. High-speed microscopic imaging of the initial stage of diesel spray formation and primary breakup. In *SAE 2010 Powertrains Fuels and Lubricants Meeting*. https://doi.org/10.4271/2010-01-2247 (2010).
3. Crua, C., Heikal, M. R. & Gold, M. R. Microscopic imaging of the initial stage of diesel spray formation. *Fuel* **157**, 140–150. https://doi.org/10.1016/j.fuel.2015.04.041 (2015).
4. Sjöholm, J. *et al.* Ultra-high-speed pumping of an optical parametric oscillator (OPO) for high-speed laser-induced fluorescence measurements. *Meas. Sci. Technol.* **20**, 025306. https://doi.org/10.1088/0957-0233/20/2/025306 (2009).
5. Kaminski, C. F., Hult, J. & Aldén, M. High repetition rate planar laser induced fluorescence of oh in a turbulent non-premixed flame. *Appl. Phys. B* **68**, 757–760. https://doi.org/10.1007/s003400050700 (1999).
6. Kawahara, N., Tsuboi, K. & Tomita, E. Laser-induced plasma generation and evolution in a transient spray. *Opt. Exp.* **22**, A44–A52. https://doi.org/10.1364/OE.22.000A44 (2014).
7. Freitag, C. *Observation of Laser Materials Processing by Means of High-Speed Imaging* 222 (Springer, Cham, 2018).
8. Kaplan, A. F. H. & Powell, J. Spatter in laser welding. *J. Laser Appl.* **23**, 032005. https://doi.org/10.2351/1.3597830 (2011).
9. Berger, P. Reduced velocity of sound in the weld pool in the presence of bubbles. *Int. Cong. Appl. Lasers Electro-Opt.* **910–915**, 2014. https://doi.org/10.2351/1.5063141 (2014).
10. Heider, A., Weber, R., Herrmann, D., Herzog, P. & Graf, T. Power modulation to stabilize laser welding of copper. *J. Laser Appl.* **27**, 022003. https://doi.org/10.2351/1.4906127 (2015).
11. Beardsley, E. G. The NACA photographic apparatus for studying fuel sprays from oil engine injection valves and test results from several researches. *NACA Technical Report* (1928).
12. Edgerton, H. E., Germeshausen, J. K. & Grier, H. E. High speed photographic methods of measurement. *J. Appl. Phys.* **8**, 2–9. https://doi.org/10.1063/1.1710232 (1937).
13. Eoh, H. A high-speed video camera operating at 4500 fps. *J. Inst. Telev. Eng. Jpn.* **46**, 543–545. https://doi.org/10.3169/itej1978.46.543 (1992).
14. Liang, J. & Wang, L. V. Single-shot ultrafast optical imaging. *Optica* **5**, 1113–1127. https://doi.org/10.1364/OPTICA.5.001113 (2018).
15. Klein-Douwel, R., Frijters, P., Somers, L., de Boer, W. & Baert, R. Macroscopic diesel fuel spray shadowgraphy using high speed digital imaging in a high pressure cell. *Fuel* **86**, 1994–2007. https://doi.org/10.1016/j.fuel.2006.11.039 (2007).
16. Reeves, M., Towers, D., Tavender, B. & Buckberry, C. A high-speed all-digital technique for cycle-resolved 2-d flow measurement and flow visualisation within si engine cylinders. *Opt. Lasers Eng.* **31**, 247–261. https://doi.org/10.1016/S0143-8166(99)00026-3 (1999).
17. Kosonocky, W. *et al.* 360360-element very-high-frame-rate burst image sensor. 182–183. https://doi.org/10.1109/ISSCC.1996.488562 (1996).
18. Lazovsky, L., Cismas, D., Allan, G. & Given, D. CCD sensor and camera for 100 Mfps burst frame rate image capture. **5787**, 184–190. https://doi.org/10.1117/12.604523 (2005).
19. Specialised Imaging - Kirana. *The ultimate high speed video camera* (2020).
20. Olofsson, J., Richter, M., Aldén, M. & Auge, M. Development of high temporally and spatially (three-dimensional) resolved formaldehyde measurements in combustion environments. *Rev. Sci. Instrum.* **77**, 013104 (2006).
21. Etoh, T. G. *et al.* The theoretical highest frame rate of silicon image sensors. *Sensors* **17**, 483. https://doi.org/10.3390/s17030483 (2017).
22. Liu, X., Liu, J., Jiang, C., Vetrone, F. & Liang, J. Single-shot compressed optical-streaking ultra-high-speed photography. *Opt. Lett.* **44**, 1387–1390. https://doi.org/10.1364/OL.44.001387 (2019).

23. Liang, J., Zhu, L. & Wang, L. V. Single-shot real-time femtosecond imaging of temporal focusing. *Light Sci. Appl.* **7**, 2047–7538. https://doi.org/10.1038/s41377-018-0044-7 (2018).
24. Nakagawa, K. *et al.* Sequentially timed all-optical mapping photography (stamp). *Nat. Photonics* **8**, 695–700. https://doi.org/10.1038/nphoton.2014.163 (2014).
25. Takasawa, K. *et al.* Single-shot 2-d burst ultrafast thz imaging utilizing sf-stamp. In *2019 Conference on Lasers and Electro-Optics Europe and European Quantum Electronics Conference* cc_{15} (2019).
26. Wang, X. *et al.* High-frame-rate observation of single femtosecond laser pulse propagation in fused silica using an echelon and optical polarigraphy technique. *Appl. Opt.* **53**, 8395–8399. https://doi.org/10.1364/AO.53.008395 (2014).
27. Ehn, A. *et al.* Frame: femtosecond videography for atomic and molecular dynamics. *Light Sci. Appl.* **6**, e17045. https://doi.org/10.1038/lsa.2017.45 (2017).
28. Maassen, K., Poursadegh, F. & Genzale, C. Spectral microscopy imaging system for high-resolution and high-speed imaging of fuel sprays. *J. Eng. Gas Turbines Power* **142**, 091004. https://doi.org/10.1115/1.4048057 (2020).
29. Dorozynska, K., Kornienko, V., Aldén, M. & Kristensson, E. A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Opt. Exp.* **28**, 9572–9586. https://doi.org/10.1364/OE.384535 (2020).
30. Kristensson, E., Li, Z., Berrocal, E., Richter, M. & Aldén, M. Instantaneous 3d imaging of flame species using coded laser illumination. *Proc. Combust. Inst.* **36**, 4585–4591. https://doi.org/10.1016/j.proci.2016.08.040 (2017).
31. Guénot, D. *et al.* Simultaneous laser-driven x-ray and two-photon fluorescence imaging of atomizing sprays. *Optica* **7**, 131–134. https://doi.org/10.1364/OPTICA.378063 (2020).
32. Versluis, M. High-speed imaging in fluids. *Exp. Fluids* https://doi.org/10.1007/s00348-013-1458-x (2013).
33. Nyquist, H. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Electr. Eng.* **47**, 617–644 (1928).
34. Haitz, R. & Tsao, J. Y. Solid-state lighting: the case 10 years after and future prospects. *Phys. Status Solidi (A)* **208**, 17–29. https://doi.org/10.1002/pssa.201026349 (2011).
35. Mishra, C. & Peles, Y. Flow visualization of cavitating flows through a rectangular slot micro-orifice ingrained in a microchannel. *Phys. Fluids* **17**, 113602. https://doi.org/10.1063/1.2132289 (2005).
36. Gregorčič, P. & Možina, J. High-speed two-frame shadowgraphy for velocity measurements of laser-induced plasma and shock-wave evolution. *Opt. Lett.* **36**, 2782–2784. https://doi.org/10.1364/OL.36.002782 (2011).
37. Hirata, T. & Miyazaki, Z. High-speed camera imaging for laser ablation process: for further reliable elemental analysis using inductively coupled plasma-mass spectrometry. *Anal. Chem.* **79**, 147–152. https://doi.org/10.1021/ac0612473 (2007).
38. Wang, P., Liang, J. & Wang, L. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nat. Commun.* https://doi.org/10.1038/s41467-020-15745-4 (2020).
39. Liang, J., Zhu, L. & Wang, L. Single-shot real-time femtosecond imaging of temporal focusing. *Light Sci. Appl.* https://doi.org/10.1038/s41377-018-0044-7 (2018).

Acknowledgements
European Research Council (ERC) (692987). European Research Council (ERC) (803634). The authors would also like to thank the company LaVision GmbH, for their contribution in the development of the prototype instrument presented in this article.

Author contributions
E.K., A.E. and E.B. designed the system and A.F. built it. V.K. conducted the experimental work, analyzed the data and was responsible for the writing. All authors reviewed the manuscript.

Funding
Open Access funding provided by Lund University.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-73738-1.

Correspondence and requests for materials should be addressed to V.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper II

scientific reports

Check for updates

OPEN

Long sequence single-exposure videography using spatially modulated illumination

Simon Ek, Vassily Kornienko & Elias Kristensson✉

Frequency recognition algorithm for multiple exposures (FRAME) is a single-exposure imaging technique that can be used for ultrafast videography, achieved through rapid illumination with spatially modulated laser pulses. To date, both the limit in sequence length as well as the relation between sequence length and image quality are unknown for FRAME imaging. Investigating these questions requires a flexible optical arrangement that has the capability of reaching significantly longer image sequences than currently available solutions. In this paper we present a new type of FRAME setup that fulfills this criteria. The setup relies only on (i) a diffractive optical element, (ii) an imaging lens and (iii) a digital micromirror device to generate a modulated pulse train with sequence lengths ranging from 2 to 1024 image frames. To the best of the authors' knowledge, this is the highest number of temporally resolved frames imaged in a single-exposure.

In a wide range of scientific research fields the ability to record fast transient events in real time is crucial for understanding of the subject at hand, e.g. the delay in photoemission[1], molecular motions[2], and photosynthetic systems[3]. High speed cameras have been developed to meet this demand, but due to practical—and eventually even theoretical—limitations in how fast a detector can be read out and made ready to capture the next frame, there is an upper limit on the attainable speed of a high speed camera[4]. To achieve even higher frame rates, an approach that does not rely on fast detectors is needed. One example is the pump-probe methodology, where a transient event is initiated by a pump pulse and probed by another pulse after a controllable time delay. By repeating the process multiple times with varying time delays, the course of the event can be captured with a temporal resolution set by the probe's pulse length[5,6]. Events that can not be repeated, e.g. due to being stochastic, can, however, not be studied using this method. As a response to this technological gap, a variety of single-exposure techniques, based on illuminating the sample with a train of short laser pulses, have been developed[7]. Since the laser pulses are separated in time, they will each be encoded with image information of temporally distinct parts of the studied event. To retrieve the image information of the individual pulses each pulse must be unique with regard to some characteristic, which can e.g. be angle[8], spatial position[9,10] or wavelength[11,12]. Single-shot femtosecond time-resolved optical polarimetry (SS-FTOP)[9] and light in-flight recording by holography (LIF-DH)[10] both rely on a spatial separation of the pulses, albeit in different ways. SS-FTOP uses a glass echelon of stepwise increasing thickness, while LIF-DH uses an obliquely sweeping reference pulse to achieve the space division. Compressed ultrafast photography (CUP)[13] and trillion CUP (T-CUP)[7] also rely on space division, through the use of a streak camera. However, they are distinct from the other two space division techniques, in that they are indirect imaging techniques, using compressive sensing to reconstruct the sequences. The CUP techniques have recently been further refined into compressed ultrafast spectral photography (CUSP)[14]. Sequentially time all-optical mapping photography (STAMP)[12] and spectrally filtered STAMP (SF-STAMP)[11] divert pulses into different parts of the detector, with respect to wavelength. STAMP does so using a transmission grating, while SF-STAMP uses a diffractive optical element (DOE) and a bandpass filter. The above techniques are able to reach picosecond- or, in some cases, femtosecond scale temporal resolution.

In 2017 our research group in Lund developed a new single-exposure filming method, called Frequency Recognition Algorithm for Multiple Exposures (FRAME), which relies on encoding unique spatial modulations into each pulse of the pulse train[15–17]. Although all the pulses reach the same part of the detector, the image information carried by the individual pulses are separated in Fourier space, according to the superimposed modulation of the pulses. This allows the sequence of frames, carried within the pulse train's constituents, to be reconstructed. Since the size of Fourier space is constant for a given sensor, if more frames are added each frame has to be reconstructed using fewer of the Fourier components, to avoid introducing neighbouring frames (crosstalk) in the reconstructed frame. Therefore the average image quality of the sequence will decrease as the

Department of Combustion Physics, Lund University, Lund, Sweden. ✉email: elias.kristensson@forbrf.lth.se

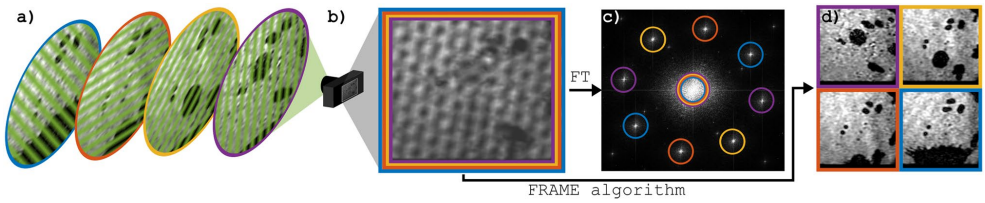


Figure 1. A conceptual illustration of FRAME. (**a**) Four laser pulses, with unique spatial modulations have passed a falling drop. (**b**) The pulses reach the camera and all add up to the detector image. (**c**) The Fourier transform of the detector image reveals the image information being separated according to the spatial modulation of respective pulse. (**d**) By multiplying the image in (**b**) with respective modulation and applying a low-pass filter the individual frames are reconstructed.

sequence gets longer. As for most, if not all, single-exposure techniques, the trade-off between sequence length and image quality is thus inherent to FRAME. For single-exposure techniques that separate pulses into distinct parts of the detector—space division techniques—a k times increase in the number of frames, n, decreases the number of pixels per frame, k times. For FRAME there is no such simple relation between n and pixels per frame. As mentioned, the decrease in the average image quality of the sequence in FRAME is a consequence of having to use fewer Fourier components per reconstructed frame. In practice this means that more high frequency components are excluded as the number of frames increase. This approach is analogous to standard image transform compression, such as JPEG, where removing spatial frequencies with low amplitude—in practice, often the high frequency components—is the means used to achieve compression[18]. Thus far, the low sequence length has thus far been regarded as a limitation of the FRAME concept[7]. In a 2017 FRAME experiment a femtosecond laser system was used to create a video sequence of a light pulse in flight with a temporal resolution of 200 fs[15]. The pulse train was constructed by splitting the output of a single femtosecond laser using beam splitters into four pulses, thus setting the sequence length. However, to be able to investigate whether longer sequences are possible with FRAME, an optical arrangement that does not rely on beam splitters is needed. With beam splitters, the setup would grow in size and complexity with each additional pulse and thus be too bulky and impractical to handle for sequence lengths above ~ 10.

Here we demonstrate a new multiplexing optical arrangement for illumination-based FRAME, that allows for videos consisting of up to 1024 image frames: a 250-fold increase compared to what has been previously demonstrated. This is made possible by replacing the traditional beam splitter arrangement with only two optical components; a DOE and a digital micromirror device (DMD). Compared to an optical setup based on beam splitters the current system is both significantly more compact and up to 30000 times more light efficient for longer sequence lengths, which has allowed for the investigation and validation of FRAME's compatibility with longer image sequences.

Setup and experimental work

FRAME is not one fixed technique, but rather a multitude of experimental methods in combination with a specific post-processing algorithm, used to reconstruct individual frames F_i from an original experimental image I[19–22]. Figure 1 shows a flow scheme that explains the principles of FRAME. First of all, the goal of the experimental part is to have a set of light pulses, each with a unique spatial modulation,

$$m_i = cos(xv'_{i,x} + yv'_{i,y} + \phi_i),$$

reaching a detector after interacting with a sample. In the expression, (x, y) are the spatial coordinates, ϕ_i an unknown phase, $v'_{i,x}$ and $v'_{i,y}$ the components of the spatial frequency $\vec{v}'_i$, and i an index running from 1 to n. Each pulse will be encoded with image information in accordance with the time it interacted with the sample. In Fig. 1a four such intensity-modulated pulses are shown, each carrying information about a falling water droplet at different times. The pulses reach the detector—within a single exposure—and add up to form I in Fig. 1b. By Fourier transforming the acquired image (and taking the absolute value of the result), the image in Fig. 1c is obtained. In this domain the majority of the information about the individual frames is confined to small areas, or clusters, each containing the image information of one video frame. In Fig. 1c the first order clusters are separated and can be seen as distinct spots. The central coordinates of these are the frequency components $\pm\vec{v}_i$ of the modulations observed by the detector. The *observed* modulations $\pm\vec{v}_i$ (in px^{-1}) corresponds to the *illumination* modulations $\pm\vec{v}'_i$ (in m^{-1}). The zeroth order clusters will never be separated, but mixed around the origin in the center. To reconstruct the frame $F_{i=j}$ from I, i.e. to go from Fig. 1b–d, involves multiplying I by a modulation matrix $M_{v_j,\theta}$, with frequency v_j and phase θ, followed by the application of a low-pass filter (LPF). Multiplying I with a modulation matrix shifts the entire Fourier domain such that the corresponding frequency component ends up in the origin in the center. A subsequent application of a LPF removes all but the primary frequency components for the frame, resulting in it's reconstruction. The complete algorithm step to reconstruct the j'th frame is:

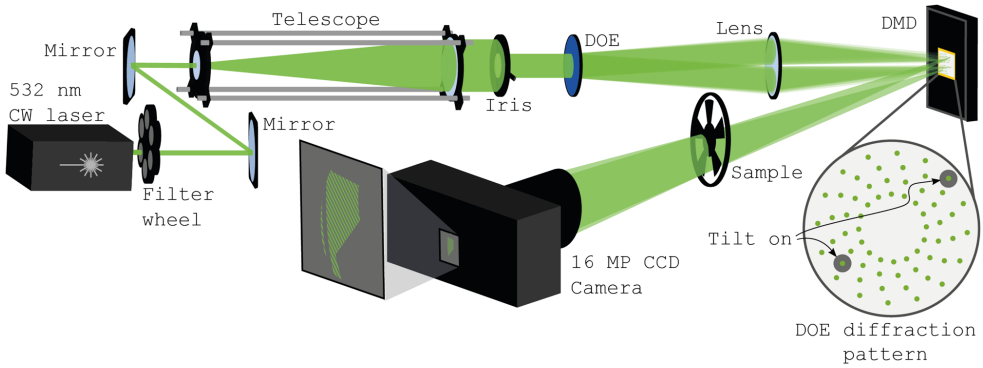


Figure 2. A schematic drawing of the optical arrangement. The DOE and DMD are used to create a train of uniquely modulated pulses, which illuminate the sample. The circular inset shows the DOE diffraction pattern together with a highlighted beam pair, experimentally selected by setting a proper tilt-on pattern on the DMD.

$$F_j = \sqrt{LPF_\sigma(I \odot M_{v_j,\theta})^2 + LPF_\sigma(I \odot M_{v_j,\theta+\pi/2})^2}, \tag{1}$$

where σ is the full width at half maximum (FWHM) of the LPF. The maximum σ that can be used, while still avoiding crosstalk between frames, is dictated by the distance between the spots in Fig. 1c.

For FRAME to work in the context of videography, each light pulse must be temporally controlled and spatially modulated. The modulation can either be achieved by imaging Ronchi gratings with different orientation and/or frequency onto the sample or by the interference of two coherent beams. A benefit with the latter approach is that the intensity modulation naturally forms a pure sine-wave, which will only generate zeroth and first order clusters in Fourier space, whereas Ronchi gratings produce a square wave modulation, also yielding additional higher order cluster, which may interfere in the image post-processing.

The setup we designed and built to investigate whether FRAME is compatible with long sequences relies on the method of overlapping coherent beams to create an interference pattern. Figure 2 shows a schematic drawing of the setup, which utilizes a continuous 532 nm diode laser (O-Like, 200 mW laser module), a diffractive optical element (DOE) (HOLO/OR), an imaging lens (Thorlabs, LA1979-A) and a digital micro-mirror device (DMD) (Vialux, V-9501 VIS) to create the spatially modulated pulse train. The camera used is a B4822 from Imperx with a 3256 by 4880 pixels ($\approx$ 16 MP) KAI-16070 charge coupled device (CCD) detector with 12 bits of grayscale. The setup also includes a filter wheel to adjust the laser intensity, adjustable mirrors for aligning purposes and a telescope in conjuction with an iris for expansion and cropping of the beam. This way the beam incident on the DOE has the desired diameter and a near top-hat intensity profile. A DOE is a commercially available optical element where a piece of glass has been etched on the micrometer scale in order to control the behaviour of an incident wavefront[23]. This allows for the possibility of tailoring a predetermined diffraction pattern without losing light intensity, allowing for the use of DOEs in many different areas of research such as beamshaping[24], optical tweezer applications[25] and micropatterning[26]. The DOE for the current experiments was chosen such that the original 532 nm beam is split into 64 copies with varying horizontal and vertical angles such that the resulting diffraction pattern has the form of the inset of Fig. 2. This diffraction pattern is then focused onto the DMD by a lens ($f = 200$ mm).

The DMD consists of 1920 by 1080 micromirrors that can be tilted to direct the beams towards either the sample (tilt on) or a beam dump (tilt off). If the tilt is in its on state, the beams are directed towards the sample onto which the image of the DOE is formed. By setting a proper pattern of tilt on and tilt off on the DMD, two beams can be made to propagate towards the sample and interfere, thus creating a spatial modulation. Changing the pattern on the DMD so that another pair of beams are allowed to propagate towards the sample changes the spatial modulation of the sample illumination. By repeating this procedure in a rapid succession the desired modulated pulse train is created and the light intensity of each modulated pulse stays at a constant 1/32 of the initial light input (the DOE divides the light into 64 beams of equal intensity). Due to the large number of possible combinations of beam pairs, these need to be chosen strategically in order to maximize the use of Fourier space.

The number of spatial modulations that can be achieved depends on the characteristics of the DOE. Generally, the more beams the DOE splits the original beam into, the more unique beam combinations exist. However, the geometry of the DOE beam pattern also matters. The inset in Fig. 2 shows the DOE pattern that we opted to use. This pattern was chosen with the intent of creating sequences of 32 frames and it consists of 64 beams, arranged in four rings of 16 beams each. However, using 64 beams there are not only 32 possible beam pairs but $\binom{64}{2} = 2016$ possible combinations. For the purpose of counting the number of pairs that yield unique spatial modulations these beam pairs can be divided into two categories. The first category contains the 32 pairs that

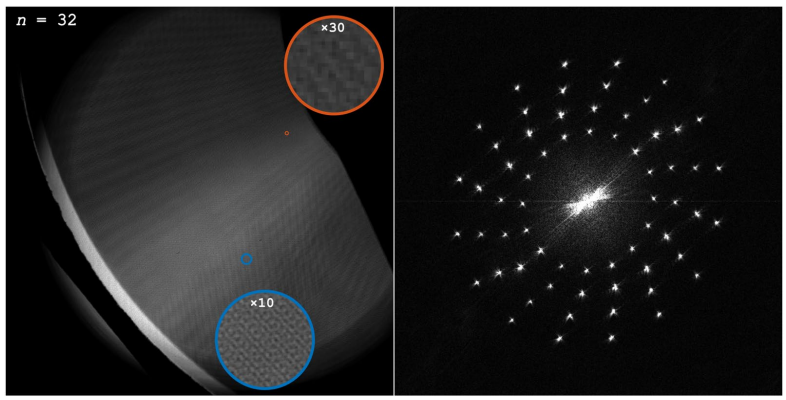


Figure 3. The original detector image of the fan (left) and the absolute value of its Fourier transform (right), when using a sequence length of 32 frames.

consist of a beam and its "twin", i.e. all pairs that are identical under a 180° rotation about the centre. All 32 beam pairs in this category yield unique spatial modulations. The remaining $2016 - 32 = 1984$ pairs make up the second category. Here, each combination of two beams has a duplicate (mirrored) pair that will give rise to the exact same interference pattern. Therefore, the beams in the two categories can in total be combined into $1984/2 + 32 = 1024$ pairs that yield unique interference patterns, setting the maximum sequence length of the setup.

Due to the multitude of available unique modulations, there are many ways to choose n beam combinations for sequence lengths of $n < 1024$. In general, beam pairs were selected to maximize the distance between frequency components in Fourier space, but some combinations that would end up in particularly ill suited regions of Fourier space were removed, in favour of other combinations. Ill suited regions are, e.g., near the origin in the center where all the zeroth order clusters add up, as well as where information about the target's stationary parts accumulate. The frames in each sequence were reconstructed using an implementation of Eq. (1), with the filter's FWHM (σ) being set by the minimum distance between any two spots in Fourier space. A selection of the reconstructed frames are presented in the following section.

Analysis and results

Perceived image quality. In order to investigate how the image quality of the reconstructed frames varies as a function of sequences length, a computer fan, spinning at 3000 RPM, was recorded. Being repeatable, the fan is a suitable sample for this purpose, since essentially the same dynamic event can be filmed multiple times. Also, the low structural complexity of the sample makes it easy to detect flaws and artefacts in the captured frames. The sample was filmed at 10 kfps in sequences consisting of 32, 64, 128, 256, 512, and 1024 frames. The original detector image and its Fourier transform are displayed in Fig. 3 for a sequence of 32 frames. For each sequence eight evenly spaced frames are presented in Fig. 4.

In order to make out a trend each frame was attributed, by visual inspection, an image quality on a five-graded scale from "good" to "useless". Several examples of each quality can be seen in Fig. 4. As expected, due to a decreasing filter size, the trend from this subset is that the image quality gets worse as the number of multiplexed frames, n, increases. The more complete and quantitative, but less qualitative, picture of this trend is given in Fig. 5. Here the Fourier transforms of the original images of each sequence are shown, together with spot markings, colored according to the image quality of the corresponding reconstructed frame. The overall trend is a trade-off between sequence length n and image quality, which can noted as the quality drops steadily from mostly "good" when $n = 32$ to mostly "awful" when $n = 1024$. Figure 5 further shows how the distance between spots decreases with n and, consequently, the low-pass filter radii must be reduced with n in order to avoid crosstalk between neighbouring frames. At the extreme case where $n = 1024$, most of the frequency components are indistinguishable from each other and the spatial resolution of the extracted data is thus very low. However, the results are promising as they show the possibility to encode a significantly higher amount of information into a single photograph using structured illumination than previously achieved and that a more strategic placement of the coded image information could enable better image qualities at long sequences.

Spatial resolution and data storage capabilities. The FRAME technique exploits the fact that natural images are mostly represented by low spatial frequencies in order to store image data at the vacant high spatial frequencies. FRAME uses intensity modulation to achieve this aim, although similar results can be achieved using e.g. space division methods[7]. As mentioned above, with space division techniques the signals from different time stamps are directed onto different (spatial) regions of the sensor. Each such image frame thus has

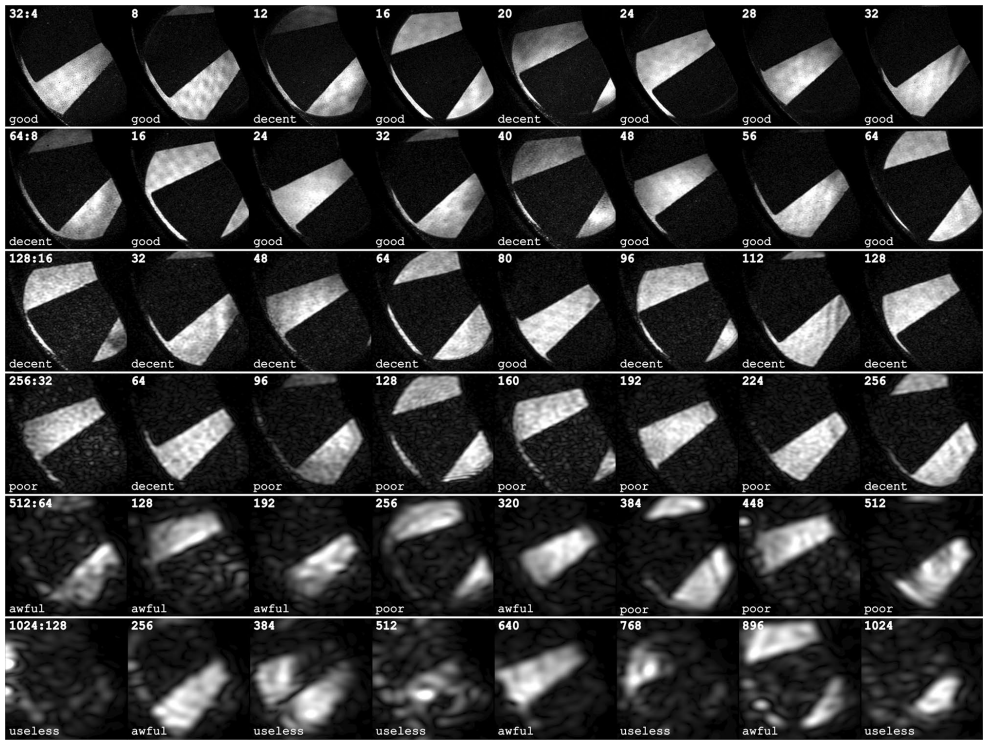


Figure 4. Examples of reconstructed frames for sequences of different length. Each row contains frame number $i\frac{n}{8}$ $(i = 1 \ldots 8)$ from a video sequence consisting of n frames (indicated for each row). The estimated image quality of each frame is indicated as "good", "decent", "poor", "awful" or "useless".

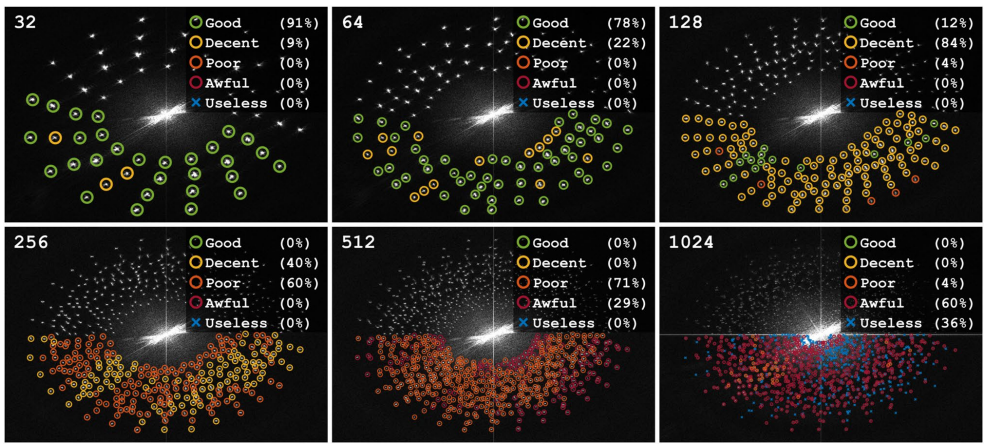


Figure 5. The Fourier transforms of the six original (unprocessed) images, with frequency spots marked according to the image quality of the corresponding reconstructed frame. The images have been cropped to 64% of their original size. The legends show the percentages of each image quality.

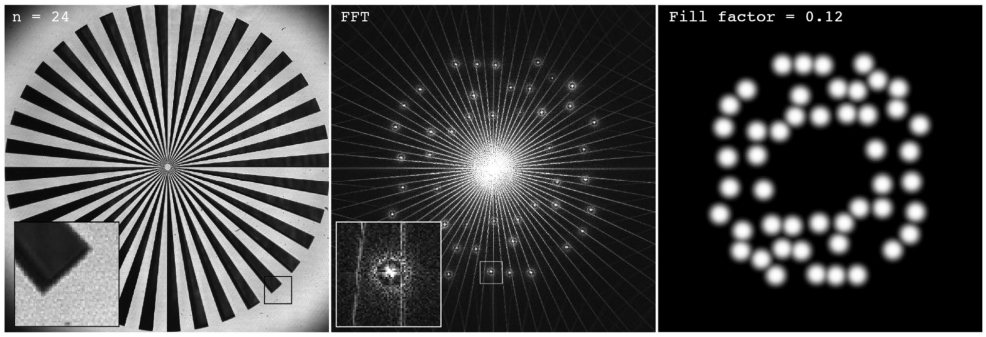


Figure 6. The original detector image of the sector star target (left) and the absolute value of its Fourier transform (middle), when using a sequence length of 24 frames. (Right) Visual representation of all the low-pass filters applied, the sum of which equals the fill factor.

a reduced number of pixels N_{frame} that, at best, is equal to the total number of pixels of the sensor N divided by the sequence length (number of divisions) n, i.e. $N_{\text{frame}} = N/n$. Note that this upper limit of the amount of pixels per image frame is only achieved when the fill factor is 100%, meaning a perfect distribution of the frames across the sensor. Under these ideal conditions, it can be shown that the linear spatial resolution achievable with this approach is reduced by a factor of $1/\sqrt{n}$. To date, the corresponding relationship between image resolution and sequence length for FRAME is not known and requires a more in-depth analysis to be understood. To elucidate, the outcome of the FRAME post-processing of a multiplexed image is a set of image frames, which, in contrast to space division techniques, each have the same number of pixels N_{frame} as the original (sensor) image, i.e. $N_{\text{frame}} = N$. The entire set of images in the extracted series therefore contains $n \cdot N$ pixels - a factor of n greater than the corresponding value for the space division approach. This value, which suggests that FRAME greatly enhances the sensors data storage capabilities, is, however, somewhat misleading since the spatial lock-in algorithm used to extract the data, or more specifically the low-pass filter, leads to a reduced spatial resolution. Instead, each of these oversampled images could be represented by a reduced number of pixels and consequently, a more accurate estimate of the "effective" number of pixels, $N_{\text{eff}}(n)$, in the entire data set is $\alpha(n) \cdot n \cdot N$, where $\alpha(n)$ is the coefficient of oversampling ($0 < \alpha \leq 1$). We will now estimate N_{eff} as a function of n in order to investigate FRAME's data storage capabilities.

To estimate $\alpha(n)$, a stationary sector star target was filmed at 10 kfps in 15 sequences of different lengths, ranging from $n = 2$ to $n = 256$ frames. The sector star target, which can be seen in Fig. 6, is 10 mm in diameter and consists of 36 black blades with a spatial frequency ranging from 1.15 line pairs per millimeter (lp/mm) at the outer edge to 57.4 lp/mm close to the center. From these measurements, the modulation transfer function (MTF) was extracted and used to find the spatial frequencies at which the reconstructed image frames show 10% contrast between the alternating white and black regions. Note that although the sector star target is a two-dimensional object, the resulting analysis yields a one-dimensional number that represents the overall spatial resolution—in both the x- and y-direction - of the image. This 1D value was then measured for all reconstructed frames in all 15 sequences. The top row in Fig. 7 shows four reconstructed frames from sequences of different lengths together with circles marking the 10% contrast. The general trend can be seen in these images; the longer the sequence, the lower the spatial resolution for the image frames. The sequence with 256 frames was left out of the analysis, since the constituting frames had a 10% cut-off frequency below 1.15 lp/mm (lowest spatial frequency of star target).

The MTF analysis of the acquired data shows that the spatial resolution ($R(n)$) for the image frames extracted using FRAME reduce more rapidly with sequence length n compared to the idealized space division case (Fig. 8a). When combined with the Nyquist theorem, which states that at least two pixels are needed to resolve a single line-pair[27], these values of 1D spatial resolution can be used to determine the minimum number of pixels, N_{min}, needed to display the 10×10 mm^2 camera field-of-view (essentially a conversion into 2D) according to:

$$N_{min}(n) = (2 \cdot R(n) \cdot 10)^2. \tag{2}$$

The ratio between $N_{min}(n)$ and the sensor's total number of pixels, N, thus equals the coefficient of oversampling α for a given sequence consisting of n images:

$$\alpha(n) = \frac{N_{min}(n)}{N}. \tag{3}$$

Examples of the extracted data before and after being downsampled are shown in Fig. 7 together with their corresponding α values. The effective number of pixels, N_{eff}, required for the full FRAME video sequence can thereafter be calculated according to:

$$N_{\text{eff}}(n) = n \cdot N_{min}(n) = n \cdot \alpha \cdot N. \tag{4}$$

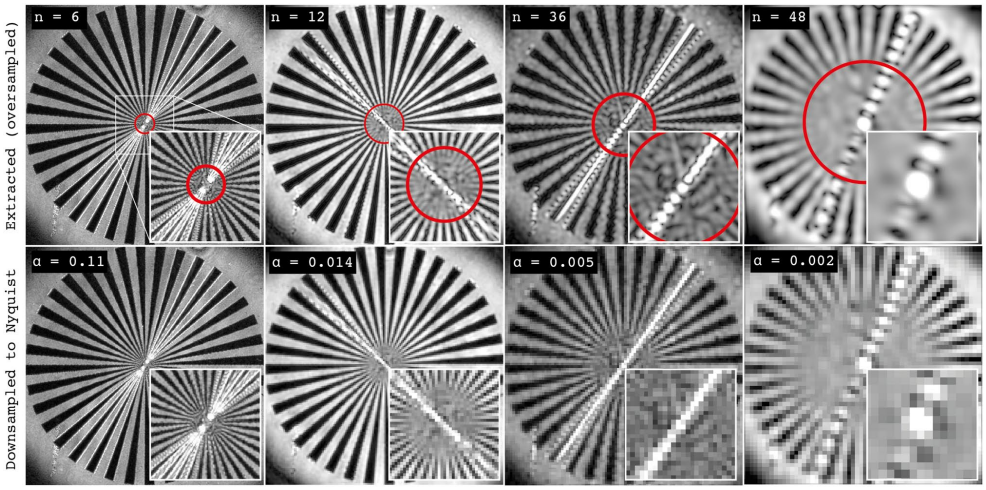


Figure 7. Top row: examples of reconstructed (oversampled) images of the sector star target, showing the location for the 10% contrast cut-off frequency for sequences with 6, 12, 36 and 48 frames. Bottom row: Pixel resolution digitally reduced to match the Nyquist sampling. Note how the image resolution has not changed compared to the top row due to the oversampling nature of the extraction algorithm.

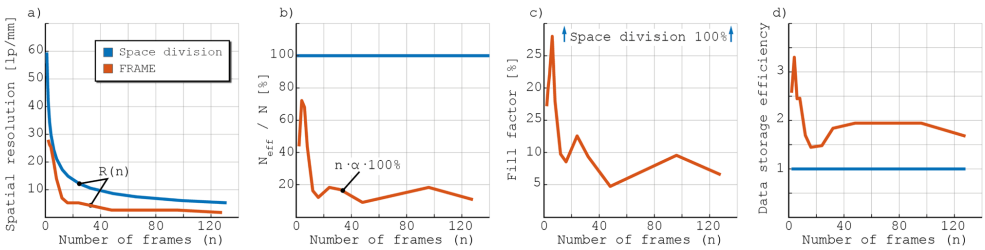


Figure 8. (**a**) 1D spatial resolution, $R(n)$, as a function of number of frames, n, for both FRAME (experimental data) and space division (theoretical data). (**b**) Effective number of pixels, N_{eff}, i.e. the pixels needed to represent the entire data set (video sequence), as a function of n. FRAME reaches, at best, 72% of the sensor's full capacity. (**c**) Fill factor (percentage of the sensor's number of pixels exploited), as a function of n. Here it is assumed that a constant fill factor of 100% can be achieved using space division. By comparing with graph (**b**), one can note that for $n = 3$, 72% of sensor's full pixel capacity is reached using only a fill factor of 28%. (**d**) The relationship between the trend in (**b**) and that in (**c**), i.e. the data storage efficiency for either technique. While FRAME does not surpass the sensor's available number of pixels ($N_{\text{eff}} < N$), the curve shows that image data is more efficiently stored in the frequency domain than in the spatial domain.

Figure 8b displays $N_{\text{eff}}(n)/N$ in percentage for both space division and FRAME. Unlike the idealized space division case that has a constant $N_{\text{eff}}(n) = N$, FRAME does not exploit all of the sensor's pixels but reaches a maximum of 72% for $n = 3$ and plateaus at 10–20% for higher n. This trend implies that FRAME fails to use the sensor's full data storage capacity, which is expected as FRAME cannot have a fill factor of 100% (see e.g. Fig. 6). Instead, the fill factor, which for FRAME is the total pixel area covered by the LPFs divided by the total area of the sensor, reach a maximum of 28% at $n = 3$, after which it drops to 5–10% at higher n (Fig. 8c). The ratio between the effective number of pixels, N_{eff}, and the fill factor indicates the data storage efficiency; for the space division approach, this value cannot exceed unity whereas the corresponding values for the FRAME measurements reach a maximum of about 3.3 (Fig. 8d). This implies that image data is more efficiently stored in the frequency domain than in the spatial domain, or, in other words, that although FRAME makes use of fewer pixels ($N_{\text{eff}} < N$), it exploits them more efficiently. We attribute this trend to the fact that the power of image information is not homogeneously distributed in reciprocal space but primarily concentrated at low spatial frequencies.

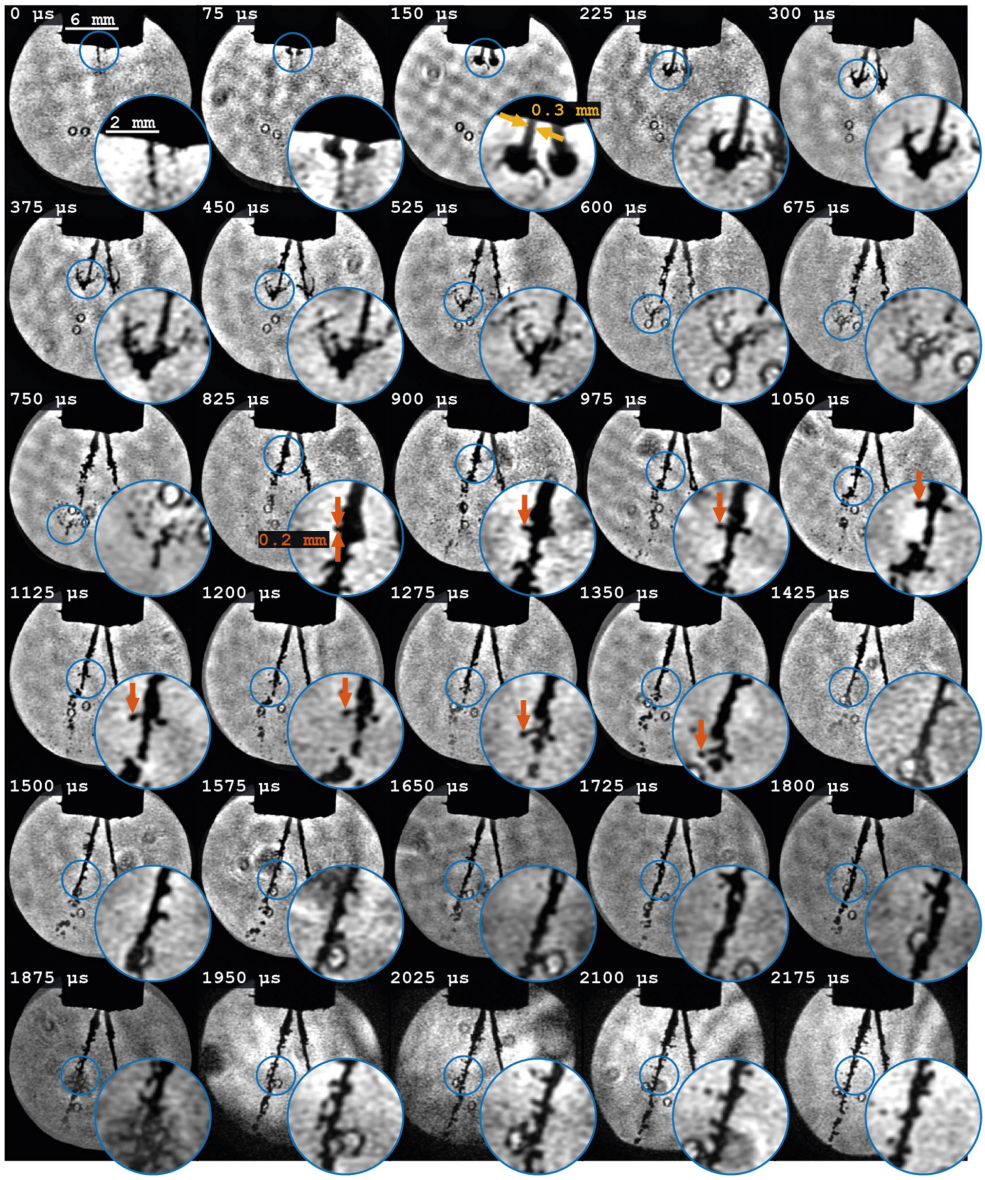


Figure 9. Reconstructed frames from the spray experiment. The insets have been magnified 3 times. The orange arrows indicate a traceable structure of widths of 0.2 mm, i.e. the estimated resolution limit of the system for the a sequence length of $n = 32$, while the yellow arrows indicate the 0.3 mm width of the jet.

Application on a stochastic target. Unlike pump-probe techniques, FRAME offers the possibility to film transient, unrepeatable events. An example of such a one-time event is the injection of fuel into the cylinders of an engine. To find out whether FRAME can be used to acquire relatively long video sequences of such

stochastic and one-time events, a nozzle producing a two spray plumes of water was filmed. The liquid pressure was set to 1.5 bar, making the jets travel at around 17 m/s. To approximately match the duration of the transient event the sequence length was set to 32 frames and the recording speed was set to the highest possible with the setup; 13.3 kfps.

The first 30 frames, out of the 32, can be found in Fig. 9, where magnified insets have been added to highlight the ability to capture detailed structural information even at relatively long sequences. For example, the spray plumes, which are approximately 0.3 mm wide, are clearly resolvable, yet finer details can as well be resolved. Theoretically, an object requires a minimum of two pixels (one line pair) to be resolved and therefore structures as small as $1lp/R(n = 32) \approx 200\mu\, m$ should be detectable, according to the results in section "Spatial resolution and data storage capabilities". The formation and breakup of such a small liquid structure can be observed and traced over the 8 frames from 825 to 1350 µs.

Discussion

Though being able to reach extraordinary video recording speeds, FRAME has thus far only been demonstrated with a relatively short sequence length. Here we have demonstrated the compatibility of illumination-based FRAME with long sequence length consisting of up to 1024 individual images: a 250-fold increase from previous demonstrations. This was made possible by greatly reducing the number of essential optical components in the setup to only 3; a DOE, an imaging lens and a DMD. While past solutions to achieve such a long sequence of image frames would have required nearly 2000 beam splitters to form the pulse train, this division is instead accomplished by only *one* DOE. This improvement greatly reduces the physical footprint of the FRAME setup and has a $3 \cdot 10^4$ higher light-division efficiency.

We have demonstrated the applicability of the setup on two dynamic targets; a computer fan and a spray. The known geometry and time evolution of the fan allowed for the attribution of a perceived image quality to each reconstructed frame. Even though the spatial resolution of the individual images decreases as the sequence length increases, the fan blades are clearly traceable for the majority of images even in the 1024 long image sequence (Supplementary Videos 1–6 and Fig. 4). In contrast to the simple geometric structure of the fan, the measurements on the atomizing spray system demonstrated the system's ability to visualize more complex transient structures and fine image details at sequence lengths up to $n = 32$ (Supplementary Video 7).

Due to the fixed etched pattern on the DOE, which was designed for the $n = 32$ case, the spread of image information in the Fourier domain could not be optimized for higher n, ultimately leading to either cross-talk between the individual image frames or extremely narrow low-pass filters. By redesigning the DOE pattern, the data can be distributed more strategically in reciprocal space, which would boost the fill-factor and thus the overall image quality for $n > 32$. The presented results should therefore not be considered as an upper limit for the image quality and sequence length of FRAME but rather as a demonstration of the previously unknown image storage capabilities made possible by the technique.

The development of faster imaging systems is indicative of the growing need to follow physical, chemical and biological processes on time-scales that have not previously been possible. However, the stochastic nature of these processes makes it difficult to synchronize a short acquisition time with the event of interest. Circumvention of this problem necessitates either adjustable acquisition timing or, preferably, longer video sequences. The presented work demonstrates, for the first time, the opportunity of acquiring long video sequences with FRAME for the investigation of such events.

Received: 5 June 2020; Accepted: 16 October 2020
Published online: 03 November 2020

References

1. Schultze, M. *et al.* Delay in photoemission. *Science* **328**, 1658–1662 (2010).
2. Zewail, A. H. Femtochemistry: atomic-scale dynamics of the chemical bond. *J. Phys. Chem. A* **104**, 5660–5694 (2000).
3. Berera, R., van Grondelle, R. & Kennis, J. T. Ultrafast transient absorption spectroscopy: principles and application to photosynthetic systems. *Photosynth. Res.* **101**, 105–118 (2009).
4. Etoh, T. G. *et al.* The theoretical highest frame rate of silicon image sensors. *Sensors* **17**, 483 (2017).
5. Haight, R. *et al.* Picosecond time-resolved photoemission study of the inp (110) surface. *Phys. Rev. Lett.* **54**, 1302 (1985).
6. Downer, M. & Shank, C. Ultrafast heating of silicon on sapphire by femtosecond optical pulses. *Phys. Rev. Lett.* **56**, 761 (1986).
7. Liang, J. & Wang, L. V. Single-shot ultrafast optical imaging. *Optica* **5**, 1113–1127 (2018).
8. Li, Z. *et al.* Single-shot tomographic movies of evolving light-velocity objects. *Nat. Commun.* **5**, 3085 (2014).
9. Wang, X. *et al.* High-frame-rate observation of single femtosecond laser pulse propagation in fused silica using an echelon and optical polarigraphy technique. *Appl. Opt.* **53**, 8395–8399 (2014).
10. Kakue, T. *et al.* Digital light-in-flight recording by holography by use of a femtosecond pulsed laser. *IEEE J. Sel. Top. Quant. Electron.* **18**, 479–485 (2011).
11. Suzuki, T. *et al.* Single-shot 25-frame burst imaging of ultrafast phase transition of ge2sb2te5 with a sub-picosecond resolution. *Appl. Phys. Express* **10**, 092502 (2017).
12. Nakagawa, K. *et al.* Sequentially timed all-optical mapping photography (stamp). *Nat. Photon.* **8**, 695 (2014).
13. Gao, L., Liang, J., Li, C. & Wang, L. V. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature* **516**, 74–77 (2014).
14. Wang, P., Liang, J. & Wang, L. V. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nat. Commun.* **11**, 1–9 (2020).
15. Ehn, A. *et al.* Frame: femtosecond videography for atomic and molecular dynamics. *Light Sci. Appl.* **6**, e17045 (2017).
16. Kristensson, E. *et al.* Instantaneous 3d imaging of flame species using coded laser illumination. *Proc. Combust. Inst.* **36**, 4585–4591 (2017).
17. Dorozynska, K. & Kristensson, E. Implementation of a multiplexed structured illumination method to achieve snapshot multispectral imaging. *Opt. Express* **25**, 17211–17226 (2017).

18. Wallace, G. K. The jpeg still picture compression standard. *IEEE Trans. Consumer Electron.* **38**, xviii–xxxiv (1992).
19. Gragston, M., Smith, C., Kartashov, D., Shneider, M. N. & Zhang, Z. Single-shot nanosecond-resolution multiframe passive imaging by multiplexed structured image capture. *Opt. Express* **26**, 28441–28452. https://doi.org/10.1364/OE.26.028441 (2018).
20. Dorozynska, K., Kornienko, V., Aldén, M. & Kristensson, E. A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Opt. Express* **28**, 9572–9586. https://doi.org/10.1364/OE.384535 (2020).
21. Li, Z. *et al.* Simultaneous multispectral imaging of flame species using frequency recognition algorithm for multiple exposures (frame). *Combust. Flame* **192**, 160–169. https://doi.org/10.1016/j.combustflame.2018.02.009 (2018).
22. Deng, C., Hu, X., Zhang, J. S. Y. & Dai, Z. Z. Q. Snapshot hyperspectral imaging via spectral basis multiplexing in fourier domain. *Opt. Express* **26**, 32509–32521 (2018).
23. Wyrowski, F. Diffractive optical elements: iterative calculation of quantized, blazed phase structures. *J. Opt. Soc. Am. A* **7**, 961–969. https://doi.org/10.1364/JOSAA.7.000961 (1990).
24. Gu, H., Chen, M., Wang, Q. & Tan, Q. Design of two-dimensional diffractive optical elements for beam shaping of multicolor light-emitting diodes. *Appl. Opt.* **57**, 2653–2658. https://doi.org/10.1364/AO.57.002653 (2018).
25. Zheng, M. J., Ogura, Y. & Tanida, J. Application of optical tweezers using DOE and SLM to control of beads with information-DNA for photonic DNA computing. In Mu, G., Song, F., Yu, F. T. S. & Jutamulia, S. (eds.) *Inf. Opt. Photon. Technol. II*, vol. 6837, 74 – 82, https://doi.org/10.1117/12.767629. International Society for Optics and Photonics (SPIE, 2008).
26. Kuroiwa, Y., Takeshima, N., Narita, Y., Tanaka, S. & Hirao, K. Arbitrary micropatterning method in femtosecond laser microprocessing using diffractive optical elements. *Opt. Express* **12**, 1908–1915. https://doi.org/10.1364/OPEX.12.001908 (2004).
27. Nyquist, H. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Electr. Eng.* **47**, 617–644 (1928).

Acknowledgements
European Research Council (803634).

Author contributions
S.E. and E.K. designed the experiment. S.E. performed most of the experimental work and analysis and wrote the majority of the paper with input from E.K. and V.K. V.K. also helped with the control of the hardware.

Funding
Open Access funding provided by Lund University.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-75603-7.

Correspondence and requests for materials should be addressed to E.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper III

1712 Vol. 10, No. 7 / July 2022 / Photonics Research Research Article

PHOTONICS Research

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

Vassily Kornienko,* David Andersson, Mehdi Stiti, Jonas Ravelid, Simon Ek, Andreas Ehn, Edouard Berrocal, and Elias Kristensson

Division of Combustion Physics, Department of Physics, Lund University, Lund, Sweden
**Corresponding author: vassily.kornienko@forbrf.lth.se*

Received 13 December 2021; revised 28 March 2022; accepted 29 March 2022; posted 2 May 2022 (Doc. ID 451108); published 30 June 2022

Fast transient events, such as the disintegration of liquid bodies or chemical reactions between radical species, involve various processes that may occur at different time scales. Currently, there are two alternatives for monitoring such events: burst- or high-speed imaging. Burst imaging at ultrahigh speeds (∼100 MHz to THz) allows for the capture of nature's fastest processes but only for a narrowly confined period of time and at a repetition rate of ∼10 Hz. Monitoring long lasting, rapidly evolving transient events requires a significantly higher repetition rate, which is met by existing ∼kHz to 1 MHz high-speed imaging technology. However, the use of such systems eliminates the possibility to observe dynamics occurring on the sub-microsecond time scale. In this paper, we present a solution to this technological gap by combining multiplexed imaging with high-speed sensor technology, resulting in temporally resolved, high-spatial-resolution image series at two simultaneous time scales. We further demonstrate how the collection of such data opens up the tracking of rapidly evolving structures up to MHz burst rates over long durations, allowing, for the first time, to our knowledge, the extraction of acceleration fields acting upon the liquid bodies of an atomizing spray in two dimensions at kHz frame rates.

https://doi.org/10.1364/PRJ.451108

1. INTRODUCTION

Digital high-speed cameras have been used to monitor a plethora of transient events ranging from laser ablation and plasma formation [1,2] to cell dynamics [3,4] and the investigation of geomaterials [5]. The demand for affordable methods to investigate even faster events in high quality has been, and is, ever increasing, serving as the driving force behind the extensive development of single sensor digital cameras. This has resulted in a 10^4 increase in frame rates over the last 30 years [6,7].

Various electronic sensor architectures such as charged coupled devices (CCDs), complementary metal oxide sensors (CMOSs), and *in situ* storage image sensors (ISISs) have been invented with the purpose of boosting sensor frame rate. However, this usually comes at a cost in the form of a reduced sequence depth and/or pixel number [5]. For example, the state-of-the-art Photron SA-Z records videos with a sensor size of 1024 × 1024 pixels at a speed of 20 kfps (fps, frames per second) [8], while the corresponding value when operating at a frame rate of 2.1 Mfps is a sensor size of 128 × 8 pixels. The speed of this camera still allows it to work in near continuous mode, resulting in sequence depths of almost 9×10^7 images (at its maximum speed of 2.1 Mfps). Dao *et al.*, on the other hand, demonstrated a fast sensor technology with a similar pixel resolution (32 × 32 pixels) capable of an acquisition speed of up to 25 Mfps. However, the sequence depth at these speeds is significantly reduced, reaching 1220 frames, and the data acquisition can instead be classified as high-speed burst videography [9]. In 2020, Suzuki *et al.* presented an even faster sensor, capable of capturing a burst of 368 images (50 × 108 pixels) at 100 Mfps, thus yielding a 3.68 μs long video [7].

This apparent trade-off among speed, sensor size, and sequence length will unavoidably dictate the level of resolvable spatial and temporal characteristics of the phenomena under study. Currently available fast single sensors are able to produce short, low-resolution burst videos at low burst repetition rates and are thus useful for globally monitoring rapidly evolving events whose dynamics are localized in time and where the onset can be predicted or triggered [10–14]. In contrast, sensors with a reduced frame rate can provide long video sequences at high pixel resolution for monitoring structural changes occurring on a ∼10 μs time scale [15–17], yet those sensors are limited to more slowly evolving events.

To overcome some of the rapid single-sensor-based limitations, alternative videography solutions have emerged. For example, multi-channel intensified cameras can acquire a burst of up to 16 images (1024 × 1024 pixels) at 200 Mfps [18]. This

2327-9125/22/071712-11 Journal © 2022 Chinese Laser Press

type of technology, where the event is imaged onto multiple sensors in series, allows for high pixel resolution and imaging speed, but at the cost of low burst repetition rates (10 Hz for the multi-channel intensified camera Imacon 468, and 0.5 Hz for the Icarus 2, an X-ray nanosecond gated camera) [14,19,20]. Another way of overcoming electronic sensor-based trade-offs is with the use of image multiplexing concepts. These techniques can include the application of spatial masks on images, followed by algorithmic extraction [21,22] or via spectral filtering of different colors [23], enabling alternative means to produce fast burst videos in the Mfps regime. Such solutions have, however, been able to provide only low burst repetition rates due to the technological restrictions set by either the camera (e.g., streak camera) or the employed illumination (e.g., LEDs).

To date, there is no existing video technology that has demonstrated the ability to circumvent the aforementioned technological trade-offs to monitor small-scale structural details, whose evolution occurs at a nanosecond time scale while the entire event may unfold over multiple orders of magnitude longer in time. Such events, however, do occur within many fields of the natural sciences such as within transient spray diagnostics [24,25], combustion and plasma diagnostics [26–29], nanostructure dynamics [30], or velocimetry in general [31,32].

The ability to both track extremely fast small-scale dynamics and to continuously monitor the slower large-scale motion would require technology that provides (1) a high frame rate (within a single burst), (2) high burst repetition rate, and (3) imaging with high pixel and spatial resolution. In this paper, we combine multiplexed imaging with high-speed single sensor technology to meet all these requirements, allowing for the tracking of transient events occurring on the MHz time scale with standard, non-intensified kHz–MHz camera technology. The capabilities of the system are demonstrated, to the best of the authors' knowledge, for the first time by extracting 2D accelerometry maps at kHz rates over a period of time spanning the entire event of interest (>1 s).

2. EXPERIMENTAL SETUP

To obtain high-speed accelerometry data of dynamic transient events at kHz time scales, we hybridize high-speed illumination-based multiplexed imaging with high-speed CMOS camera technology. The optical setup presented herein allows for three burst images with a temporal resolution of up to 52 ns (19 MHz) at a burst repetition rate of up to 1 Mfps (depending on the desired field of view).

The system that attains these capabilities is divided into three units: (i) a high-speed passive imaging unit whose main component is a high-speed CMOS camera [Photron SA5, top of Fig. 1(a)], (ii) a high-speed illumination system, consisting of three individual high-repetition-rate pulsed nanosecond laser sources [Thorlabs NPL45- and 49-B, bottom of Fig. 1(a)], and (iii) an image multiplexing unit built of optical components that image unique spatial modulations onto the sample.

The two electronic units, (i) and (ii), are synchronized in time via a triggering system such that within a single camera exposure, each laser is triggered once at separated time intervals [inset of Fig. 1(a)]. This pulse train is optically relayed onto the

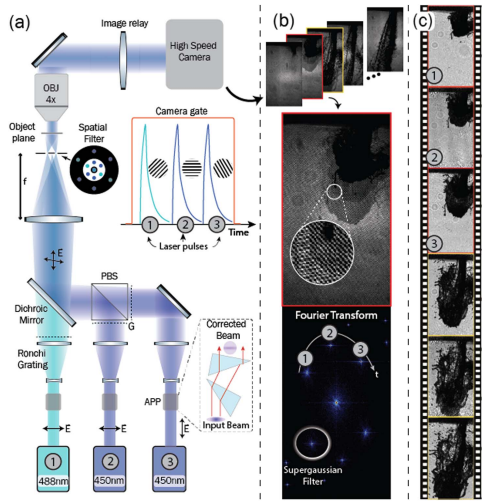


Fig. 1. System capable of monitoring microscopic MHz dynamics with kHz technology. (a) The light from each nanosecond pulsed laser source, triggered w.r.t the high-speed camera, is spatially modulated with Ronchi gratings (20 lp/mm) and combined into a pulse train incident on the sample. The imaged event is then relayed to the high-speed camera via a microscopy configuration. Due to the laser pulses' temporal separation [inset of (a)] the event of interest will be illuminated at three distinct times within a single camera exposure. (b) Each high-speed camera image consists of three superimposed spatially modulated images of which an example is framed in red. Its Fourier transform depicts distinct peaks, within which the individual pulse information is contained. Performing a spatial lock-in followed by low-pass filtering on the peaks extracts said information. (c) Carrying out this process (including background correction) on each high-speed camera image results in a video sequence where bursts of three images (FRAME triplets) are taken at a repetition rate set by the camera, allowing, e.g., the extraction of 2D accelerometry data from each camera image. The complete video of (c) is included as Visualization 4.

sample via a polarizing beam splitter (PBS), dichroic mirror, and lens ($f = 150$ mm). The now illuminated sample is then imaged onto the high-speed camera with a microscope objective (4× Olympus plan achromat microscope objective) and a relay lens ($f = 180$ mm).

All three illumination events within the rapid succession of pulses reach the camera within a single exposure and need to be separated digitally. To enable this separation and thereby identify the information retained within each pulse, a multiplexed imaging technique named frequency recognition algorithm for multiple exposures (FRAME) [22,33,34] is used. With FRAME, the intensity profile of each laser pulse is spatially modulated at a unique angle [inset of Fig. 1(a)], optically achieved by imaging Ronchi gratings (20 lp/mm) onto the sample plane (zeroth order and undesired harmonics are rejected using spatial filtering).

Upon video capture, the high-speed camera produces a series of multiplexed images where each includes three

superimposed, spatially modulated, and temporally separated illumination events of the dynamic sample [Fig. 1(b)]. This can be visualized in the example image, framed in red, where the event's progression is seen by the stepwise varying gray scales, overlayed with unique modulation combinations. As a result of these modulated structures, the Fourier transform of the image depicts three peaks separated from the DC component, each corresponding to the image produced from a single laser pulse. Algorithmic separation is then achieved by spatially locking-in to each peak, low-pass filtering (with a 2D Gaussian to the power of eight), then inverse Fourier transforming, resulting in three images of the event's progression [red framed images of Fig. 1(c)], dubbed herein as a FRAME triplet. Iterating this procedure on each image of the raw high-speed video results in an image sequence that contains MHz dynamic information (up to 19 MHz) at the repetition rate offered by the CMOS camera [video reel of Fig. 1(c)], which in the current case corresponds to a maximum of 1 Mfps.

3. VERSATILE MONITORING OF HIGH-SPEED EVENTS

To demonstrate the capabilities of this high-speed hybrid imaging system, accelerometry is performed upon the acquired microscopic video recordings of 60 ms long injection events set about by a commercial fuel injector (Bosch EV1 four-hole nozzle running at 5 bar water injection pressure). These transient liquid dynamics are imaged at a camera speed, i.e., burst repetition rate, of 21 kHz (variable C in Fig. 2) and varying burst rates (variable L in Fig. 2). The complete videos, comprising about 3 × 1300 images, of these and additional injection events are included in Visualization 1, Visualization 2, Visualization 3, Visualization 4, Visualization 5, Visualization 6, Visualization 7.

A prerequisite for accelerometry measurements is high-contrast, blur-free images such that, e.g., the accurate segmentation of structures at varying scales can be performed. The short illumination time in conjunction with the high photon flux offered by the light sources (∼50 ns and ∼2.5 nJ in this case) is able to provide such images by (1) eliminating motion blur despite the relatively long camera exposure and (2) maintaining a uniform intensity profile over the entire image for high-contrast shadows of the event. In addition, the lock-in filtering procedure associated with FRAME suppresses image blur caused by multiple light scattering [35].

A further requirement for measuring velocity vector fields of a transient event (the precursor for acceleration vector fields) is versatility in the time domain. Indeed, for different versions of non-seeded image flow analyses to function efficiently, the distance traveled by the event between two images must be optimized such that the correlation between the spatial structures remains high for the given distance traveled [36,37]. Furthermore, to extract velocity vector fields at more than two time scales (one slow [25] and the other fast [24]), one needs to vary the inter-pulse delay such that the time between any pair of pulses is optimized for vector extraction at the given time scale. Figure 2(b) illustrates these capabilities offered by the high-speed FRAME imaging system. The independence of the light sources within the illumination system allows for a varied inter-pulse time (20 μs between the first two and 2 μs

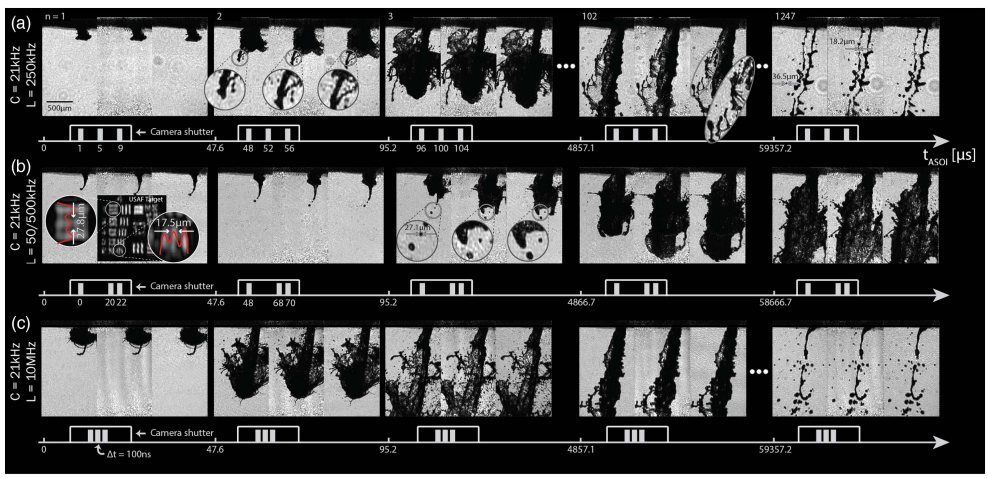


Fig. 2. Microscopic imaging of transient fluid dynamics at high speed. (a) Imaging the spray injection event at a camera speed of 21 kHz and a shutter time of 12 μs within which three pulses, with constant inter-pulse time (250 kHz FRAME triplet), are recorded. (b) Non-linear temporal triggering where the first two pulses arrive at the sample with an inter-pulse time of 20 μs (50 kHz), while the last two are incident with an inter-pulse time of 2 μs (500 kHz). At the beginning of the event, 50 kHz suffices to track the fluid's evolution within the camera exposure, while already at 95 μs ASOI, higher speeds are necessary. The zoom-ins of the USAF 1951 resolution target correspond to the resolution limits in x and y of the current optical configuration. (c) Imaging the injection event with 21 kHz camera speed and 10 MHz FRAME triplet (100 ns inter-pulse time) while retaining the image spatial resolution of (b). See Visualization 1, Visualization 2, Visualization 5, and Visualization 6 for videos of the entire events.

between the last two), all at a burst repetition rate set by the camera speed.

The temporal pulse shape of the illumination sources, on the other hand, allows for a maximum burst rate of 19 MHz (Appendix A). Due to the nature of the light sources and the multiplexing methodology (i.e., each light source and its corresponding event information are independent of each other), the image quality does not deteriorate as the burst rate increases. Figure 2(c) shows this trait for a burst rate of 10 MHz, where the image spatial resolution of 29 and 18 lp/mm [inset of Fig. 2(b)] and high image contrast are conserved. On the other hand, imaging at increased burst repetition rates usually results in a reduced pixel number of the high-speed camera, affecting only the field of view. Hence, imaging at higher rates overall (burst and burst repetition rates) does not inherently imply a deteriorated spatial resolution or image quality.

A. High-Speed Tracking of Liquid Dynamics over the Entire Event at Various Spatial Scales

An inherent problem with tracking the acceleration of fast transient events is the requirement of three measurements close enough in time such that a finite difference approximation of the velocity's time derivative can be accurately extracted. Previous attempts at capturing the data necessary for such an analysis have primarily focused on ballistic imaging techniques using a multiple number of time-gated cameras (ICCDs) [38] or by combining two CCD cameras in PIV mode with polarization-based filtering, resulting in two velocity maps close in time [39]. Although these techniques provide excellent temporal and spatial resolution per acquisition, both ICCDs and interline transfer (double-framed) CCDs often have a restricted frame rate and would be able to extract accelerometry data only at a repetition rate of 10^0 – 10^2 Hz. The approach presented in this paper overcomes this problem by combining the image multiplexing ability offered by FRAME, i.e., high burst rate, with the high frame rate offered by modern high-speed CMOS cameras. The rate at which instantaneous acceleration vector fields can be extracted is thus dictated by the repetition rate of the CMOS camera that greatly exceeds that of previously demonstrated approaches.

Estimating the acceleration vector fields from the FRAME triplets is performed by tracking spray features in time using an in-house developed segmentation and template matching post-processing step. The left column of Fig. 3(a) depicts a FRAME triplet with the post-processing segmentation edge of each measurement outlined in green. The segmentation accuracy can be varied with different parameters such as the radius of the low-pass filter used in the lock-in algorithm or the parameters used in the segmentation algorithm such as thresholds, and filling and/or morphological operation parameters.

Templates of a preset size ($M \times N$ pixels) are extracted for, depending on the desired vector density, every 2nd to 10th

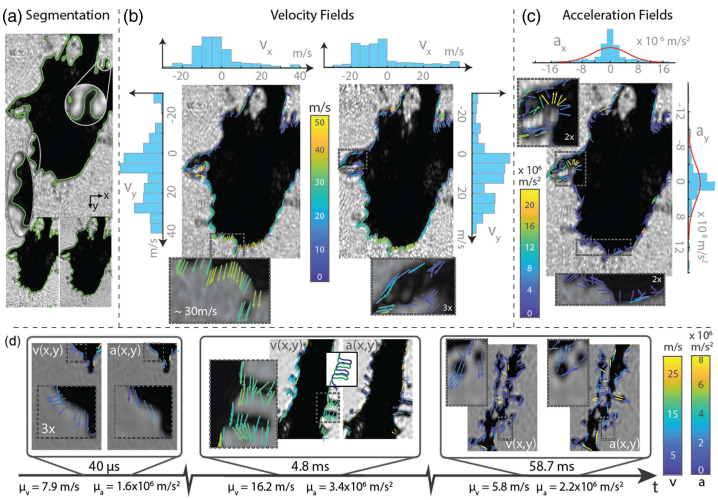


Fig. 3. Extracting 2D accelerometry data from a single camera exposure. (a) Thresholding and morphologically closing results in the green segmentation of the shadowgram. (b) Template matching between two frames results in 2D velocity vector fields (color coded arrows), calculated at the pixel level along the segmentation edge. Here we choose to calculate the vectors every fourth pixel, yielding a vector density of 0.25 px^{-1}. Histogramming their sizes indicates downward movement towards the left (negative x direction but positive y direction) at an average speed of $\sim\sqrt{10^2 + 15^2} = 18$ m/s. (c) The matching of two velocity vector fields results in the acceleration (or, equivalently, force) fields that act upon the liquid body. The centering of the histograms around zero shows that the the vast majority of the spray does not accelerate. The Gaussian distribution of vectors, however, demonstrates the stochastic nature of the turbulence that is present. Furthermore, the zoom-in reveals a possible primary breakup event that is underway. (d) This process of force extraction is performed over the complete series of images from the start to the end of injection, showing the evolution of the spray's dynamics for the entire event. The sensitivity of the system/algorithm is withheld for the varying velocity and acceleration regimes over the course of the event, from low to high velocity/acceleration at varying times. For a complete video of the injection event overlayed with accelerometry vectors, see Visualization 8.

pixel along the segmentation edge. These templates are then laterally shifted over the subsequent images where the shift that maximizes the cross-correlation is set as the distance traveled by the spatial feature. Based on the lateral shift and the time separation between images, velocity vectors for each template are calculated. In contrast to seeded flow velocimetry where the displacement of seeded particles is monitored with grid-based cross-correlation algorithms, this technique yields a vector field that reveals the structural displacement of features along the entire edge of the liquid structures.

Through said template matching of the triplet, two velocity vector maps can be extracted from a single exposure [Fig. 3(b)]. Here the velocity vectors are drawn and color-coded w.r.t their speed, and histograms of their x and y components are also displayed alongside. In the leftmost zoom-in, the highly uniform velocity vectors of the spray front are highlighted, indicating a region of high stability that is moving with the inertial speed of the entire liquid body. In contrast, the rightmost zoom-in of Fig. 3(b) depicts a liquid ligament that is traveling away from the main body, indicating a possible breakup event. In general, however, the statistics of the v_x histograms show that the spray travels slightly towards the left, while the v_y histograms indicate a movement downward at an average speed of about 15 m/s.

By pairing vectors from the two velocity vector plots, average acceleration vectors can be extracted over the duration of the pulse train triplet [Fig. 3(c)]. In the case of fluid dynamics, this map can then be directly related to physical quantities such as pressure p and viscosity μ via, e.g., the Navier Stokes equation [39]

$$\frac{\partial v_i}{\partial t} = -v_j \frac{\partial v_i}{\partial x_j} - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \mu \frac{\partial^2 v_i}{\partial x_j \partial x_i}, \tag{1}$$

allowing for the validation of models related to the dynamics of the spray front. For more details on the algorithm for the extraction of accelerometry data, see Appendix B.

As a result of the constant velocity field of the spray front, and its large inertial mass, an expected low acceleration (between zero and 1×10^6 m/s^2) is seen in the corresponding region [lower zoom-in of Fig. 3(c)]. On the other hand, the large inward pointing acceleration vectors in the upper zoom-in display a quickly changing spray morphology in that region, further confirming the possibility that a primary breakup event is about to occur. In general, however, the acceleration histograms of Fig. 3(c) indicate a mean of zero acceleration in both x and y directions. Despite this, the stochastic nature of the turbulences experienced by the edges of the liquid body results in a Gaussian distribution of forces around this mean.

B. Maintained Sensitivity despite Largely Varying Event Regimes

Due to the aforementioned marriage among high burst rate, burst repetition rate, and long sequence depth offered by the presented videography technique, this type of accelerometry analysis can be performed for the entirety of the event of interest. In the case of Fig. 3(d), the last image corresponds to camera image number 4400 of a sequence running at a burst rate of 500 kHz and burst repetition rate of 75 kHz. As expected, it is visible that the liquid body undergoes many different regimes over the course of the entire event [40]—low mean acceleration and velocity 40 μs after the start of injection (ASOI), high, in general, velocity and increased turbulence in

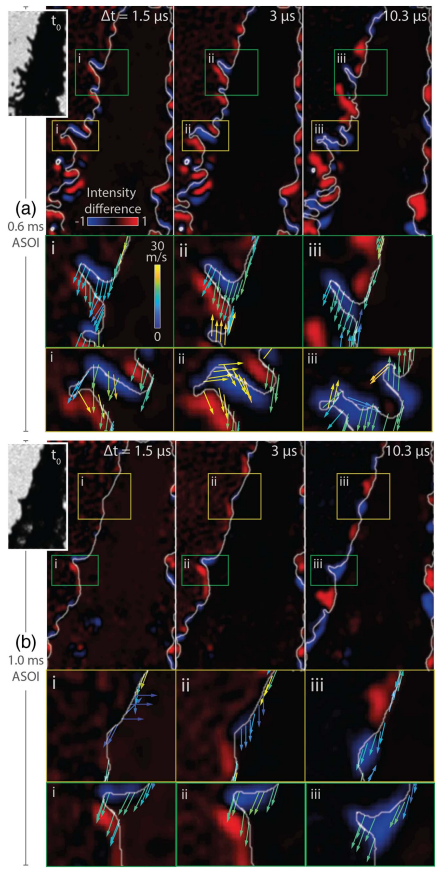


Fig. 4. Sensitivity of velocimetry to different inter-pulse durations. (a), (b) Two instances from one injection event are analyzed for three different, simultaneously acquired values of inter-pulse delay time (Δt). Intensity differences between the two individual images are depicted in blue and red, whereas black areas reflect static structures. The necessity of acquiring with multiple inter-pulse times is particularly clear for the yellow zoom-ins of (b). With a small Δt (i) the lack of structural change forces the velocity vectors to match with noise. However, as Δt increases (ii, iii), structural differences emerge, resulting in velocity vectors that are instead related to the spray dynamics. Similarly, the yellow zoom-ins of (a) show that the matching algorithm will give varying results for differing inter-frame times. In contrast, the two green zoom-ins depict structures with velocities that are more robust to such changes in Δt. The two instances together illustrate how the instabilities of the injected liquid may have different temporal characteristics throughout the duration of the event. These variations imply that a single, static value of Δt is not ideal to accurately track all the dynamics.

the middle parts (4.8 ms ASOI), and a high degree of breakup activity where air resistance dominates after the nozzle has closed (58.7 ms ASOI). This demonstrates the importance of attaining tracking ability over the entire lifetime of the event to follow the physical cause and effect relationships acting upon the dynamics. For the entire video overlayed with the accelerometry data, refer to Visualization 8.

The accuracy of the velocity vectors, extracted using our template matching algorithm, increases as the tracked features undergo large lateral shifts in between consecutive frames. On the other hand, the instabilities and turbulence associated with liquid atomization cause the large, intact, liquid core to disintegrate and break up into smaller, liquid bodies. Such structural deformations, which the algorithm interprets as rigid body motion, become more pronounced as the inter-frame time Δt increases, thus counteracting any gain in accuracy obtained through large lateral shifts.

Another aspect regarding the accuracy of the velocity vectors concerns the large spread in time scales involved in the liquid atomization process. Obtaining an accurate finite difference approximation for both the rapid changes associated with the primary breakup as well as the motion of the main liquid body requires various inter-image times, which is achievable using FRAME. This aspect is exemplified in Fig. 4, where the varying structural speeds observed within the same exposure require different values of Δt to be accurately identified and measured.

C. Single Droplet Tracking

The high contrast of the acquired images in conjunction with the marginal loss in image spatial resolution following the low-pass filtering allows for simultaneous segmentation, force tracking, and extraction of quantitative data from features at different spatial scales. Figure 5 reveals the force dynamics acting on a single droplet that has broken off from the main body.

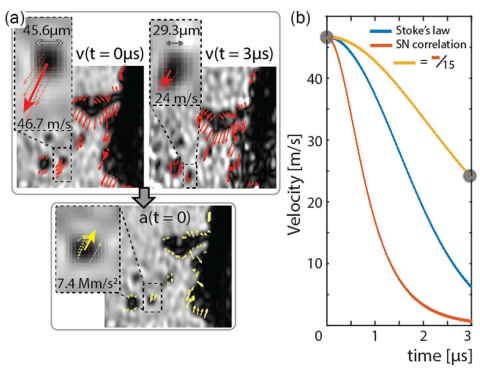


Fig. 5. High-speed tracking of air resistance on a single droplet. (a) The direction of the forces acting on a single droplet, depicted by the solid yellow arrow (itself equal to the mean of the dashed arrows) indicates strong air resistance. (b) Despite this, the measured decrease in velocity from 46.7 to 24 m/s (depicted as a black circle) is less than expected from air resistance simulations performed using Stoke's law and Schiller–Naumann (SN) correlation. Due to the deformation of the liquid droplet to a more aerodynamic shape (seen in the second velocity image), the correct drag coefficient is 15 times less than that proposed by SN correlation. See Appendix C for details.

Modeling of the velocity progression of this single droplet (simplified to a spherical solid drop of radius of 45.6 μm and initial speed of 46.7 m/s) with a simple Stokes or Schiller–Naumann (SN) model, results in a decrease in velocity to 6 versus 1 m/s, respectively, after 3 μs, significantly lower than the measured value of 24 m/s [blue and red lines of Fig. 5(b) versus the black circle]. However, we see that the droplet does not remain spherical throughout its journey but instead attains a more aerodynamic form after 3 μs [zoom-in of Fig. 5, $v(t = 3\ \mu s)$], resulting in a reduced drag coefficient. In conjunction with the acceleration vectors pointing in the direction of air resistance, it is possible to infer this inconsistency between model and experiment by taking into account the deformation of the liquid droplet to a more aerodynamic form (SN/15) caused by air resistance [41,42]. See Appendix C for more details.

4. DISCUSSION

The underlying principle of sensor-based high-speed cameras is the ability to shift photoelectrons between electronic components in a structured manner at extremely high speeds. This physical movement of electrons, however, fundamentally limits sequence depth, sensor size, and imaging speed [43], confining these types of technologies to the imaging of either kHz events over a duration of several seconds or MHz events over the duration of microseconds. On the other hand, illumination-based techniques can reach THz imaging speeds, but the fundamental necessity to record the entire sequence on a single sensor within a single exposure usually severely limits their sequence depths and/or spatial resolution. This results in a strict trade-off between imaging speed and temporal length of the image sequence. In this work, we combined illumination-based multiplexed imaging with sensor-based high-speed cameras to perform accelerometry on liquid bodies at video rates significantly faster than previously demonstrated.

The optical hybrid configuration presented herein has several technical advantages over conventional high-speed imaging techniques. First, hybridization of FRAME is not limited to single sensor technology but is compatible with, e.g., rotating mirror cameras (such as Cordin 560 or Brandaris 128), ultrafast framing cameras (Imacon) or ISIS. Second, the independence of the pulse train light sources allows for non-linear temporal sampling of the event. As a result, the time separation between pulses can be tailored to track events ranging from low to high speeds [2,17,44] or events that include both time scales simultaneously [3,4,27]. Moreover, this flexibility allows for the acquisition of suitable inputs to acceleration extraction algorithms such as the one presented herein or in Ref. [36], a task that is, as has been shown, quite sensitive to the Δt between images. Third, although higher frame rates ($\Delta t \sim 10^{-9}$ s) can be achieved with, e.g., ultrafast framing cameras or double-pulsed PIV systems, the burst repetition rate of such systems—the rate at which velocity or acceleration data can be measured—is usually in the range of 10^0 – 10^2 Hz. Hence, the extraction of accelerometry data of rapid transient events has, thus far, been

restricted to a single-shot basis [38,39]. In contrast, the system presented herein relies on neither double-frame image technology nor intensified (time-gated) cameras, thereby granting a repetition rate up to six orders of magnitude greater. We take advantage of this boost in frame rate enabling the near continuous extraction of acceleration data throughout the entire transient event. Finally, since all images are acquired on a single sensor, calibration procedures to achieve pixel-to-pixel overlapping and compensation for different magnifications are not necessary, which greatly facilitates and decreases the uncertainties related to region matching analyses [45].

Due to the electronic speed limit of CCD and CMOS sensors, a notable amount of development has instead gone into increasing the number of pixels, their sensitivity, and noise characteristics while still maintaining high speeds (e.g., a 100 kfps 2 Mpx sensor was realized in 2017 [46]). This general trend greatly benefits hybridization with multiplexed imaging techniques such as FRAME, as this would allow more dynamic range per multiplexed image, more images per exposure (significantly more than three has been experimentally demonstrated [34]), and higher spatial resolution, enhancing the temporal characteristics of both cameras and FRAME itself. On the other hand, fast illumination technology in the form of high-repetition-rate femtosecond laser sources or burst lasers is becoming more commercially available than ever before, opening up high-repetition-rate THz burst accelerometry of, e.g., chemical reactions [47].

5. CONCLUSION

While the development of double-framed cameras and burst lasers has enabled velocimetry, the ability to measure acceleration has, thus far, been restricted to solutions using multiple camera optical configurations running at low repetition rates. Furthermore, the technological gap restricting simultaneous access to slow kHz and fast MHz–THz processes, limits the possibility to efficiently track a large variety of naturally occurring events. Hybridizing illumination-based imaging with high-speed sensor technology overcomes these restrictions, with the added benefit of being able to reap the benefits of technological advancement from the respective fields. All in all, the ability to measure properties such as acceleration of fast but long duration events is of fundamental importance to understand and model the forces acting upon physical processes, and we envision the measurement of such quantities to become a major part of future research.

APPENDIX A: TEMPORAL RESOLUTION

The temporal resolution of an imaging system should be defined in such a way as to guarantee unique information in each image. To guarantee this, we have previously defined pulse limits that set strict boundaries to where the next pulse is allowed to be placed in time [22]. These limits are derived from Gaussian pulses and their FWHM, within which 76% of the pulse's power lies [inset of Fig. 6(b)]. By applying this 76% limit on any other pulse shape, this definition guarantees a 76% degree of uniqueness to the image stored within this pulse.

Another way to see it is that videography entails a non-delta spike sampling of the time axis of an event with 2D images.

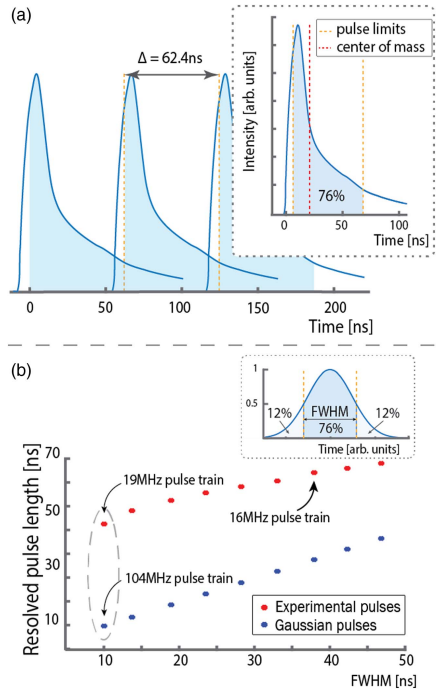


Fig. 6. 19 MHz temporal resolution. (a) Due to their shape, the pulses from these nanosecond laser sources have a center of mass [red dashed line of inset] separated from the peak. Hence, the pulse limits (yellow dashed lines), within which 76% of the light intensity lies, are much further apart than the FWHM of the peak. Placing three such pulses in a pulse train, such that they adhere to the defined temporal resolution condition, allows for a 16 MHz pulse train. (b) Pulse limits versus FWHM: the asymmetric pulse shape of the lasers used here causes the pulse limits to be much larger than their Gaussian counterparts (red dots versus blue dots of the plot).

Even though active techniques have the ability to overlap pulses in time, according to the Nyquist theorem, one cannot gain much from this approach. To avoid oversampling, we must define pulse limits, within which the resulting image can be seen as a unique sampling point with the integration time being the distance between said pulse limits.

The pulse characteristics of the lasers used herein were measured with a photodiode and a 6 GHz oscilloscope resulting in pulses of the form of the inset in Fig. 6(a). The pulse limits were then extracted giving a sampling integration time Δ. The maximum imaging speed of a pulse train consisting of such pulses is then given by $1/\Delta$, in this case, 16 MHz. The maximum attainable temporal resolution with these Thorlabs nanosecond laser sources was measured to be 19 MHz.

In contrast, had these pulses been Gaussian with similar FWHMs, the maximum temporal resolution would have been 104 MHz, highlighting how the sampling integration window

of the illuminating pulse can cause significant variation in the temporal resolution, as is appropriate.

APPENDIX B: ACCELERATION EXTRACTION

To extract velocity and in turn acceleration vectors from the acquired images, the image analysis technique of template matching is used. In general, a subarea of an initial image is chosen, dubbed herein as a template. This template is then used to calculate cross-correlations between itself and a second image (within a preset search area). A vector is subsequently drawn from the position of the original template to the point of highest cross-correlation in the second image. For this to function, there are a few necessary steps to perform.

1. Segmentation

Background artifacts that arise due to non-uniform illumination and the low-pass filtering of Gaussian noise in the FRAME algorithm can interfere with correlation calculations. Therefore, we segment the images to logical maps via thresholding (with values around 50%–70% of the maximum), morphological closing (disc radius of one to five pixels), and hole filling. The edge pixels of the resulting logical maps [indicated in green in Fig. 3(a)] are used as the starting points for the template matching algorithm.

2. Template Matching for Velocity Vectors

Template matching is an established three-step approach to extract velocity vectors from image pairs [24,38].

1. **Extraction of a suitable template**. An initial template is chosen by extracting a matrix of image values around a given edge pixel (green boxes in Fig. 7). The size of the matrix is a preset value of 10×10 – 30×30 pixels in these cases.
2. **Searching for the corresponding point**. The aforementioned template is then used in the next image to search for its "corresponding" point. This is done by moving the template around its central point within a given search area (red dashed boxes in Fig. 7) and extracting the cross-correlation between the template and searched position in the second image. The position of maximum cross-correlation (orange boxes in Fig. 7) is chosen as the corresponding point. Here it is important that the time difference between the two images is such that there is a non-zero shift in the spray while minimal structural change has occurred within the search area (see Fig. 4 for details).
3. **Extracting a velocity**. The velocity vector (red arrows in Fig. 7) is calculated as the distance and direction from the central point of the template's initial to its final position. This distance is precalibrated with resolution targets, and the speed is subsequently calculated by dividing with the inter-image time.

This three-step approach is performed for, if desired, each pixel along the edge of segmentation. Hence, the spatial resolution of the vector field, i.e., the density of the extracted vectors, is limited only by the post-segmentation spatial resolution of the image.

3. Calculating Acceleration

Using two subsequent images with associated velocity vectors, acceleration vectors valid over the time span of the three images

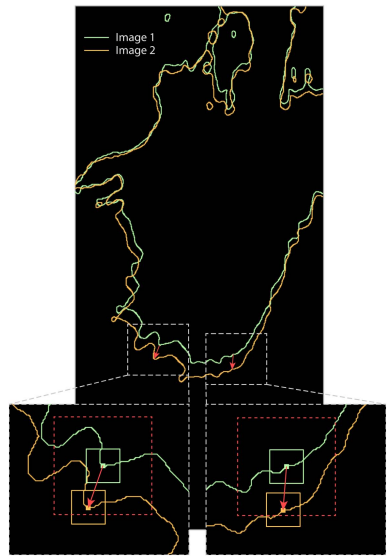


Fig. 7. Extraction of velocity vectors for each pixel along the segmentation edge. Two examples of where the initial template (green boxes) is scanned within a search area (red dashed boxes) across the subsequent image (outlined in orange). The template is placed at the position of maximum cross-correlation (orange boxes). The velocity vector is then calculated as the distance between the two center points divided by the inter-image time. This process is performed for every pixel along the segmentation edge.

used can be produced. By matching the end point of velocity vectors of the first image to the closest starting point of a vector in the second image, each point in the spray front can be tracked over the course of the three images. Once two velocity vectors are matched, and thus deemed to arise from the same segment, the average acceleration of the spray segment during the course of the three images can be calculated by taking the time derivative of the velocities:

$$\tilde{a}(t) = \frac{\mathrm{d}\tilde{v}}{\mathrm{d}t} = \frac{\mathrm{d}v_x}{\mathrm{d}t}\hat{x} + \frac{\mathrm{d}v_y}{\mathrm{d}t}\hat{y} = \frac{v_{x2} - v_{x1}}{t}\hat{x} + \frac{v_{y2} - v_{y1}}{t}\hat{y}, \quad \textbf{(B1)}$$

where the subscript indicates direction (x,y) and image (1,2), and t is the time between the two velocity vector maps. The entire code is included in a git repository and is updated continuously: https://gitter.forbrf.lth.se/david.andersson/vector-plotting.

4. Validity of the Algorithm

The presented algorithm was tested under two conditions: (1) experimental data of a rigid sphere falling under gravitational pull with negligible air resistance and (2) a synthetically created sphere also falling under gravitational pull (9.82 $\mathrm{m/s^2}$). In both cases, the inter-image time is set to 2 ms. The results are presented in Fig. 8.

The raw data input in Fig. 8(b) are a logical map depicting a falling circle. Hence, the segmentation is always perfect, resulting in an extraction of velocities where the distribution of the vectors is zero (blue and green arrows in the extracted velocities). Matching the velocities with acceleration vectors yields the desired value of 9.82 m/s^2 with, again, a zero distribution, hence validating a lack of systematic bias in the algorithm.

The imaging of the falling rigid sphere in the experiment, on the other hand, does not result in a perfect circle, and hence, the segmentation and template matching algorithm will extract a distribution of velocities [extracted velocities of Fig. 8(a)], which in turn results in a distribution of accelerations ($\sigma_a \sim \sqrt{\sigma_{v1}^2 + \sigma_{v2}^2}/\Delta t$). This distribution, however, is centered around a mean gravitational acceleration of 9.74 m/s^2, and a −0.2 m/s^2 acceleration in the x direction, plausible values within our laboratory.

A more in-depth error analysis attained by propagating a pixel error in segmentation, σ_x, into an acceleration error, σ_a, is also performed. The acceleration, a, is given by

$$a = \frac{x_3 - 2x_2 + x_1}{\Delta t^2}, \tag{B2}$$

where $x_{1,2,3}$ are edge points from the three segmented images used to extract acceleration, and Δt is the time between each pair of images. The resulting propagated error will hence be

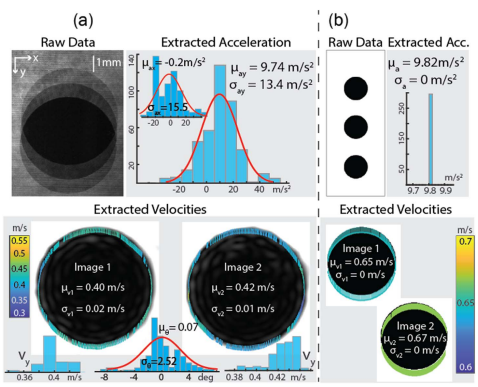


Fig. 8. Validation of the velocity and acceleration extraction algorithm. (a) Experimental validation using the known gravitational acceleration acting upon a rigid sphere. The velocities are first extracted, yielding the displayed vector fields and distributions, where the angular distribution (θ) histogram is given by $\arctan(v_x/v_y)$ (a zero angle points directly in the direction of the fall). Matching the velocity vectors results in the extracted acceleration distribution with a mean acceleration of 9.74 m/s^2 in the y direction and −0.2 m/s^2 in the x direction. σ_{ax} and σ_{ay} correspond to a rounded two pixel uncertainty in the image segmentation process, while σ_θ corresponds to a single pixel uncertainty. (b) Synthetic data of a falling sphere (logical maps) accelerating at 9.82 m/s^2 are inputted into the algorithm, resulting in distributions with zero width in velocities (magnitude and direction) and acceleration, validating the vector field extraction algorithm.

$$\sigma_a = a\sqrt{4\left(\frac{\sigma_x}{x_3 - 2x_2 + x_1}\right)^2 + 2\left(\frac{\sigma_{\Delta t}}{\Delta t}\right)^2}, \tag{B3}$$

where $x_3 - 2x_2 + x_1$ is given by

$$x_3 - 2x_2 + x_1 = a\Delta t^2. \tag{B4}$$

Given $\sigma_{\Delta t} = 0$, σ_a can be expressed as

$$\sigma_a = a\sqrt{4\left(\frac{\sigma_x}{aC\Delta t^2}\right)^2} = \sqrt{4\left(\frac{\sigma_x}{C\Delta t^2}\right)^2}, \tag{B5}$$

where C is a calibration factor converting pixels to meters for a given imaging system, in our case, 77,200 px/m. By solving for σ_x in Eq. (B5), the measured σ_{ax} of 13.4 [Fig. 8(a)] corresponds to a pixel error of 2.07 in the segmentation process. This low value highlights how the spatial resolution of the image extraction process is maintained as well as the efficiency of the segmentation/template matching algorithm.

As for the vectors' direction, the distribution of $\theta = \arctan(v_x/v_y)$ is centered around a value of 0.07° (where a value of zero points directly downwards). Performing the same type of analysis as above, the spread of 2.52° corresponds to a rounded pixel error of one, again highlighting the accuracy of the presented process.

APPENDIX C: DRAG ON A SINGLE DROPLET

A droplet's trajectory can be modeled by solving the equation of motion for a droplet subject to gravity and the drag force:

$$m_d \frac{d\vec{v}_d}{dt} = -\frac{\pi}{8}\rho_g C_d d^2 ||\vec{v}_d - \vec{v}_g||(\vec{v}_d - \vec{v}_g) + m_d\vec{g}, \tag{C1}$$

with m_d the droplet mass, v_d the droplet velocity, ρ_g the air density, d the droplet diameter, and C_d the drag coefficient. Figure 5 presents the evolution of the droplet's velocity as a function of time for different drag coefficients. We have considered Stoke's law and the SN correlation:

$$C_d = \frac{24}{R_e},$$

$$C_d = \begin{cases} \frac{24(1+0.15R_e^{0.687})}{R_e} & R_e \le 1000 \\ 0.44 & R_e > 1000 \end{cases},$$

where R_e is the Reynolds number. As observed, none of the models manages to explain the decrease in speed observed experimentally. Indeed, the observed decrease can be explained only by a drag coefficient 15 times lower than those of the SN model. This decrease in the drag coefficient is linked to spray effects and to the deformation of the droplet, which tends to become more aerodynamic [41,42].

Funding. Vetenskapsrådet (2021-04652, 2019-05183); European Research Council (803634, 852394).

Acknowledgment. The authors thank Prof. Mattias Richter for use of the high-speed camera, and Saeed Derafshzan for fruitful input on velocimetry history and techniques.

Disclosures. The authors declare no competing interests.

Data Availability. Data and code underlying the results presented in this paper are available from the corresponding author upon request.

REFERENCES

1. T. Hirata and Z. Miyazaki, "High-speed camera imaging for laser ablation process: for further reliable elemental analysis using inductively coupled plasma-mass spectrometry," Anal. Chem. **79**, 52–147 (2007).
2. J. E. Field, E. Amer, P. Gren, M. A. Zafar, and S. M. Walley, "High-speed photographic study of laser damage and ablation," Imaging Sci. J. **63**, 119–136 (2015).
3. H. Mikami, C. Lei, N. Nitta, T. Sugimura, T. Ito, Y. Ozeki, and K. Goda, "High-speed imaging meets single-cell analysis," Chem **4**, 2278–2300 (2018).
4. C. K. Haluska, K. A. Riske, V. Marchi-Artzner, J.-M. Lehn, R. Lipowsky, and R. Dimova, "Time scales of membrane fusion revealed by direct imaging of vesicle fusion with high temporal resolution," Proc. Natl. Acad. Sci. USA **103**, 15841–15846 (2006).
5. H. Xing, Q. Zhang, C. Braithwaite, B. Pan, and J. Zhao, "High-speed photography and digital optical measurement techniques for geomaterials: fundamentals and applications," Rock Mech. Rock Eng. **50**, 1611–1659 (2017).
6. H. Etoh, "A high-speed video camera operating at 4500 fps," J. Inst. Telev. Eng. Jpn. **46**, 543–545 (1992).
7. M. Suzuki, Y. Sugama, R. Kuroda, and S. Sugawa, "Over 100 million frames per second 368 frames global shutter burst CMOS image sensor with pixel-wise trench capacitor memory array," Sensors **20**, 1086 (2020).
8. J. Manin, S. A. Skeen, and L. M. Pickett, "Performance comparison of state-of-the-art high-speed video cameras for scientific applications," Opt. Eng. **57**, 121706 (2018).
9. V. T. S. Dao, N. Ngo, A. Q. Nguyen, K. Morimoto, K. Shimonomura, P. Goetschalckx, L. Haspeslagh, P. De Moor, K. Takehara, and T. G. Etoh, "An image signal accumulation multi-collection-gate image sensor operating at 25 Mfps with 32 × 32 pixels and 1220 in-pixel frame memory," Sensors **18**, 3112 (2018).
10. O. Shpak, M. Verweij, H. J. Vos, N. de Jong, D. Lohse, and M. Versluis, "Acoustic droplet vaporization is initiated by superharmonic focusing," Proc. Natl. Acad. Sci. USA **111**, 1697–1702 (2014).
11. K. Kooiman, H. J. Vos, M. Versluis, and N. de Jong, "Acoustic behavior of microbubbles and implications for drug delivery," Adv. Drug Delivery Rev. **72**, 28–48 (2014).
12. I. Beekers, M. Vegter, K. R. Lattwein, F. Mastik, R. Beurskens, A. F. van der Steen, N. de Jong, M. D. Verweij, and K. Kooiman, "Opening of endothelial cell–cell contacts due to sonoporation," J. Controlled Release **322**, 426–438 (2020).
13. A. Gleason, C. Bolme, H. Lee, B. Nagler, E. Galtier, D. Milathianaki, R. Kraus, J. Eggert, D. Fratanduono, G. Collins, R. Sandberg, W. Yang, and W. Mao, "Ultrafast visualization of crystallization and grain growth in shock-compressed SiO_2," Nat. Commun. **6**, 8191 (2015).
14. P. A. Hart, A. Carpenter, L. Claus, D. Damiani, M. Dayton, F.-J. Decker, A. Gleason, P. Heimann, E. Hurd, E. McBride, S. Nelson, M. Sanchez, S. Song, and D. Zhu, "First X-ray test of the Icarus nanosecond-gated camera," Proc. SPIE **11038**, 110380Q (2019).
15. F. H. Zhang and S. T. Thoroddsen, "Satellite generation during bubble coalescence," Phys. Fluids **20**, 022104 (2008).
16. A. Roth, D. Frantz, W. Chaze, A. Corber, and E. Berrocal, "High-speed imaging database of water jet disintegration part I: quantitative imaging using liquid laser-induced fluorescence," Int. J. Multiph. Flow **145**, 103641 (2021).
17. M. Alsved, A. Matamis, R. Bohlin, M. Richter, P.-E. Bengtsson, C.-J. Fraenkel, P. Medstrand, and J. Löndahl, "Exhaled respiratory particles during singing and talking," Aerosol Sci. Technol. **54**, 1245–1248 (2020).
18. M. Versluis, "High-speed imaging in fluids," Exp. Fluids **54**, 1458 (2013).
19. C. T. Chin, C. Lancée, J. Borsboom, F. Mastik, M. E. Frijlink, N. de Jong, M. Versluis, and D. Lohse, "Brandaris 128: a digital 25 million frames per second camera with 128 highly sensitive frames," Rev. Sci. Instrum. **74**, 5026–5034 (2003).
20. X. Chen, J. Wang, M. Versluis, N. de Jong, and F. S. Villanueva, "Ultra-fast bright field and fluorescence imaging of the dynamics of micrometer-sized objects," Rev. Sci. Instrum. **84**, 063701 (2013).
21. J. Liang and L. V. Wang, "Single-shot ultrafast optical imaging," Optica **5**, 1113–1127 (2018).
22. V. Kornienko, E. Kristensson, A. Ehn, A. Fourriere, and E. Berrocal, "Beyond MHz image recordings using LEDs and the frame concept," Sci. Rep. **10**, 16650 (2020).
23. K. Maassen, F. Poursadegh, and C. Genzale, "Spectral microscopy imaging system for high-resolution and high-speed imaging of fuel sprays," J. Eng. Gas Turbine Power **142**, 091004 (2020).
24. D. Sedarsky, S. Idlahcen, J.-B. Blaisot, and C. Rozé, "Velocity measurements in the near field of a diesel fuel injector by ultrafast imagery," Exp. Fluids **54**, 1–12 (2013).
25. L. Yin, M. Lundgren, Z. Wang, P. Stamatoglou, M. Richter, Ö. Andersson, and P. Tunestål, "High efficient internal combustion engine using partially premixed combustion with multiple injections," Appl. Energy **233–234**, 516–523 (2019).
26. S. Li, D. Sanned, J. Huang, E. Berrocal, W. Cai, M. Aldén, M. Richter, and Z. Li, "Stereoscopic high-speed imaging of iron microexplosions and nanoparticle-release," Opt. Express **29**, 34465–34476 (2021).
27. Y. Tang, C. Kong, Y. Zong, S. Li, J. Zhuo, and Q. Yao, "Combustion of aluminum nanoparticle agglomerates: from mild oxidation to microexplosion," Proc. Combust. Inst. **36**, 2325–2332 (2017).
28. M. Uddi, N. Jiang, E. Mintusov, I. V. Adamovich, and W. R. Lempert, "Atomic oxygen measurements in air and air/fuel nanosecond pulse discharges by two photon laser induced fluorescence," Proc. Combust. Inst. **32**, 929–936 (2009).
29. G. D. Stancu, F. Kaddouri, D. A. Lacoste, and C. O. Laux, "Atmospheric pressure plasma diagnostics by OES, CRDS and TALIF," J. Phys. D **43**, 124002 (2010).
30. S. Wang, J. Kang, Z. Guo, T. Lee, X. Zhang, Q. Wang, C. Deng, and J. Mi, "In situ high speed imaging study and modelling of the fatigue fragmentation of dendritic structures in ultrasonic fields," Acta Mater. **165**, 388–397 (2019).
31. A. van der Bos, M.-J. van der Meulen, T. Driessen, M. van den Berg, H. Reinten, H. Wijshoff, M. Versluis, and D. Lohse, "Velocity profile inside piezoacoustic inkjet droplets in flight: comparison between experiment and numerical simulation," Phys. Rev. Appl. **1**, 014004 (2014).
32. J. J. Philo, M. D. Frederick, and C. D. Slabaugh, "100 kHz PIV in a liquid-fueled gas turbine swirl combustor at 1 MPa," Proc. Combust. Inst. **38**, 1571–1578 (2021).
33. A. Ehn, J. Bood, Z. Li, E. Berrocal, M. Aldén, and E. Kristensson, "Frame: femtosecond videography for atomic and molecular dynamics," Light Sci. Appl. **6**, e17045 (2017).
34. S. Ek, V. Kornienko, and E. Kristensson, "Long sequence single-exposure videography using spatially modulated illumination," Sci. Rep. **10**, 18920 (2020).
35. Y. N. Mishra, E. Kristensson, M. Koegl, J. Jonsson, L. Zigan, and E. Berrocal, "Comparison between two-phase and one-phase SLIPI for instantaneous imaging of transient sprays," Exp. Fluids **58**, 110 (2017).
36. D. Sedarsky, J. Gord, C. Carter, T. Meyer, and M. Linne, "Fast-framing ballistic imaging of velocity in an aerated spray," Opt. Lett. **34**, 2748–2750 (2009).
37. J. Fielding, M. B. Long, G. Fielding, and M. Komiyama, "Systematic errors in optical-flow velocimetry for turbulent flows and flames," Appl. Opt. **40**, 757–764 (2001).
38. D. Sedarsky, M. Rahm, and M. Linne, "Visualization of acceleration in multiphase fluid interactions," Opt. Lett. **41**, 1404–1407 (2016).

39. K. Christensen and R. Adrian, "Measurement of instantaneous Eulerian acceleration fields by particle image accelerometry: method and accuracy," Exp. Fluids **33**, 759–769 (2002).
40. J. Eggers and E. Villermaux, "Physics of liquid jets," Rep. Prog. Phys. **71**, 036601 (2008).
41. C. Shao, K. Luo, and J. Fan, "Detailed numerical simulation of unsteady drag coefficient of deformable droplet," Chem. Eng. J. **308**, 619–631 (2017).
42. I. Silverman and W. Sirignano, "Multi-droplet interaction effects in dense sprays," Int. J. Multiph. Flow **20**, 99–116 (1994).
43. T. Etoh, A. Nguyen, Y. Kamakura, K. Shimonomura, T. Le, and N. Mori, "The theoretical highest frame rate of silicon image sensors," Sensors **17**, 483 (2017).
44. J. E. Chomas, P. A. Dayton, D. May, J. Allen, A. Klibanov, and K. Ferrara, "Optical observation of contrast agent destruction," Appl. Phys. Lett. **77**, 1056–1058 (2000).
45. R. Adrian, "Twenty years of particle image velocimetry," Exp. Fluids **39**, 159–169 (2005).
46. T. Kondo, Y. Takemoto, N. Takazawa, M. Tsukimura, H. Saito, H. Kato, J. Aoki, S. Suzuki, Y. Gomi, S. Matsuda, and Y. Tadaki, "A 3D stacked global-shutter image sensor with pixel-level interconnection technology for high-speed image capturing," Proc. SPIE **10328**, 1032804 (2017).
47. A. H. Zewail, "Femtochemistry: atomic-scale dynamics of the chemical bond," J. Phys. Chem. A **104**, 5660–5694 (2000).

Paper IV

scientific reports

Check for updates

OPEN

High-speed videography of transparent media using illumination-based multiplexed schlieren

Simon Ek✉, Vassily Kornienko, Adrian Roth, Edouard Berrocal & Elias Kristensson

Schlieren photography is widely used for visualizing phenomena within transparent media. The technique, which comes in a variety of configurations, is based on detecting or extracting the degree to which light is deflected whilst propagating through a sample. To date, high-speed schlieren videography can only be achieved using high-speed cameras, thus limiting the frame rate of such configurations to the capabilities of the camera. Here we demonstrate, for the first time, optically multiplexed schlieren videography, a concept that allows such hardware limitations to be bypassed, opening up for, in principle, an unlimited frame rate. By illuminating the sample with a rapid burst of uniquely spatially modulated light pulses, a temporally resolved sequence can be captured in a single photograph. The refractive index variations are thereafter measured by quantifying the local phase shift of the superimposed intensity modulations. The presented results demonstrate the ability to acquire a series of images of flame structures at frame rates up to 1 Mfps using a standard 50 fps sCMOS camera.

Schlieren is an imaging technique, invented more than one century ago, that allows faint refractive index variations in transparent media to be visualized[1–4]. With time, alternative optical configurations with similar imaging capabilities have emerged, all of which rely on detecting the refraction that occurs at the interface between regions of different refractive index. Such variations can be caused by changes in pressure, temperature or constituents, allowing phenomena such as shock-waves[5–7], heat transfer[8,9], combustion[10,11] and gas flow[12,13] to be studied.

To visualize the variations in refractive index, the fundamental schlieren configuration is based on transmission imaging with a collimated light source followed by spatial filtering to optically reject part of the light, according to degree of deflection[2,14]. Such a configuration thus permits direct visualization of regions within the sample where the light was refracted. In 1949, Burton et al.[15] demonstrated the Moiré Fringe Method (MFM) as an alternative schlieren configuration wherein two optically miss-matched transmission gratings are placed on either side of the object, thus blocking non-deflected light whilst permitting photons that deviate from the original trajectory. The benefit with MFM over conventional schlieren is that it can provide a larger field-of-view (FOV). More recently, Background Oriented schlieren (BOS)[4,16–18] has emerged as an attractive alternative to schlieren imaging, having a reduced experimental complexity compared to illumination-based schlieren. Instead of optically rejecting light that does not deviate from the original trajectory, BOS uses a textured background. Light refraction yields local distortions in this texture and by quantifying the distortion the refractive index gradients are deduced. Any textured—random[17], periodic[19] or structured[20]—background can be used, allowing e.g. supersonic air-craft shock waves to be observed by using the textured ground as background[16]. In an early version of BOS, labeled synthetic schlieren, a grating was used as background, but unlike the closely related MFM this technique utilizes an undistorted image of the background as a reference, instead of a second grating[21]. The technique shares a common limitation with schlieren and MFM in that it can only detect refractive index gradients perpendicular to the mask/grating edges. Fast Checkerboard Demodulation (FCD), demonstrated by Wildeman in 2018, is a way of overcoming this limitation, by using a checkered background and subsequent Fourier demodulation of the captured image[20]. The technique produces two unique images, corresponding to refractive index gradients in x- and y-direction, respectively, thus allowing the full 2D vector displacement field to be calculated.

Schlieren imaging does not require any particular detector. Thus, by utilizing high-speed sensor technology and intense illumination sources, rapid transient events can be studied with schlieren[4]. High-speed schlieren

Division of Combustion Physics, Lund University, Professorsgatan 1, Lund, Sweden. ✉email: simon.ek@forbrf.lth.se

videography in the 10^3–10^6 Hz video rate range can be achieved using fast sensor technology[22–25] while, with current technology, faster video rates require multiple intensified detectors[26]. However, despite a continuous development and strive towards faster sensor technology, all detection-based video configurations have a fundamentally limited frame rate due to the need to physically shift the generated photoelectrons[27]. This technological limitation has spurred the creation of alternative methods for videography that, instead of relying on rapid readout, rely entirely on high-speed *illumination*[28–30]. The acquisition rate of such illumination-based video systems are thus fundamentally limited by the pulse duration of the illumination and, with the advent of commercially available ultrafast pico- and femtosecond laser systems, has thereby enabled videography at THz video rates.

In this paper we present, to the best of our knowledge, the first demonstration of illumination-based schlieren videography, where the frame rate is governed by the pulse duration of the light source alone. The presented approach is based on an intensity modulation-scheme, which is a versatile optical method that is used for several different imaging applications, such as topography[31], super-resolution microscopy[32], schlieren[19,20,33,34], interferometry[35,36] and thermometry[37]. Here we combine its ability to provide schlieren images through phase-sensitive analysis[20] with its ability to optically multiplex temporally separated image data[28,38–40] into a powerful imaging concept. Past approaches for image multiplexing using this scheme have thus far been based on quantifying the amplitude of the intensity modulation; to gain schlieren data, we here demonstrate the ability to extract phase information from such multiplexed data. In the experiments we further verify the applicability of both coherent and incoherent light and demonstrate schlieren videography at acquisition speeds up to 1 Mfps, using a standard 50 fps sCMOS camera. We expect this imaging concept to open up for studies of transparent ultrafast, sub-nanosecond phenomena in the near future.

Method

Frequency recognition algorithm for multiple exposures. Frequency Recognition Algorithm for Multiple Exposures (FRAME) is an optical multiplexing technique that allows multiple image signals to be acquired within a single camera exposure (with a single detector), without having to read out each image individually[28]. This ability is enabled by marking each signal carrying light pulse with a unique spatial code, which is eventually used to identify and reconstruct the complete sequence of signals in the post-processing of the multiplexed camera acquisition. The character of the obtained sequence depends on how the light pulses are separated from each other in the optical part of the FRAME system. In the particular case of temporal separation the result is a video sequence[28], but the separation could also be e.g. spatial[41] or spectral[42,43]. For video sequences the minimum inter-frame time is only limited by the pulse duration, and the frame rate can therefore be significantly higher than that of the camera. To this date, frame rates of up to 5 THz have been achieved with the technique[28].

Figure 1 shows the signal extraction process of FRAME and how schlieren data can be measured using the method. Figure 1a shows a trans-illumination image of a premixed propane/air Bunsen burner flame that has been captured using the LED-based FRAME setup in Fig. 2. This image constitutes the *unprocessed* original data from which schlieren data will be extracted. The exposure time (10 µs) of this recording is too short for the flame luminescence to leave any noticeable imprints in the image and the faint outline of the flame is caused primarily by refraction of the light that transmits the flame. The main reason refraction becomes visible by eye (albeit only faintly) is because the incident light is guided through a transmission grating, which is subsequently imaged onto the detector. As seen in the zoomed in sections of Fig. 1a this spatially modulated illumination makes the image take on a striped appearance. It is this type of modulation that, not only makes refraction observable, but also makes optical multiplexing possible with FRAME. The reason multiplexing is possible is illustrated in Fig. 1b, where the image signal shows up not just in the centre of the Fourier domain, but also at positions corresponding to the angle and period of the modulation. In this context these signal copies are referred to as Fourier clusters. Even if the detector is subject to a series of signal-carrying light pulses, the respective Fourier clusters will have an insignificant overlap if the modulations of the pulses are sufficiently different from each other. This separation of Fourier clusters allow for their corresponding signals to be individually reconstructed, by applying a lock-in algorithm and a low-pass filter. The process is described in more detail in[28,38–43]. As more images are stored on the detector the distance between the Fourier clusters will unavoidably decrease, and eventually the size of the applied low-pass filter will have to be decreased to avoid cross-talk between signals. Hence, there is an inherent trade-off between the number of multiplexed images and the spatial resolution of the individual images associated with the technique. However, since for natural images most information is located close to the low frequency components in Fourier space, the losses can in practice be small, even for a large number of images[39].

Phase extraction methodology. When a spatially intensity modulated light field (such as that used in FRAME) propagates inside a transparent object, refraction leads to an observable displacement of the sinusoidal pattern. The amount of displacement at the location $\boldsymbol{r}$ in the recorded image $I(\boldsymbol{r})$ (see Fig. 1a), is determined in relation to a blank reference image $I_0(\boldsymbol{r})$. The goal of synthetic schlieren imaging (which includes BOS) is to find the full 2D displacement field $\boldsymbol{u}(\boldsymbol{r})$ that transforms $I_0(\boldsymbol{r})$ into $I(\boldsymbol{r})$[20], i.e. the $\boldsymbol{u}(\boldsymbol{r})$ that fulfills

$$I(\boldsymbol{r}) = I_0(\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r})).$$

The entity $\boldsymbol{u}(\boldsymbol{r})$ is of interest, since it to first order is proportional to the refractive index gradients of the object under study. By applying Fourier domain filtering, centered on the modulation frequency in the recordings, the following signals are obtained:

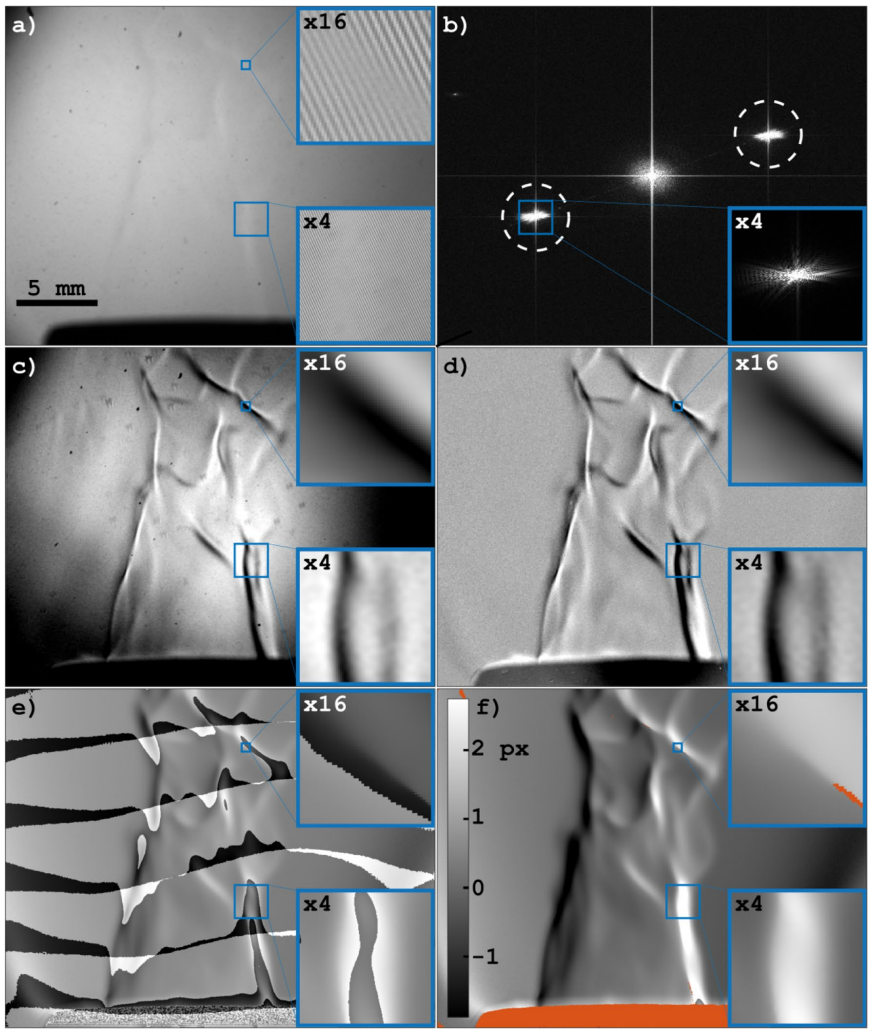


Figure 1. Illustration of the phase analysis process. (**a**) The image of the spatially modulated light pulse guided through a flame. The outlines of the flame are barely discernible, but leave a faint imprint in the modulated structure in the shape of a phase displacement. (**b**) The absolute value Fourier transform of the original image. The encircled Fourier clusters contain the information used to reconstruct the complex flame signal, carried by the light pulse. The size of the low-pass filter used for reconstruction is indicated by the circles. (**c**) The amplitude of the reconstructed signal. (**d**) The amplitude of the signal minus the amplitude of a blank recording. (**e**) The phase shift of the signal, i.e. the difference in phase between the signal and a blank recording. (**f**) The unwrapped phase shift of the signal. The grayscale bar gives the phase shift as displacement in pixels. The orange parts are areas where the amplitude of the signal was too low for proper determination of the phase. The zoomed in sections in (**a**) and (**c**)–(**f**) are of identical position. See also Supplementary video 1 for a comparison between pure trans-illumination imaging and schlieren based on modulated light.

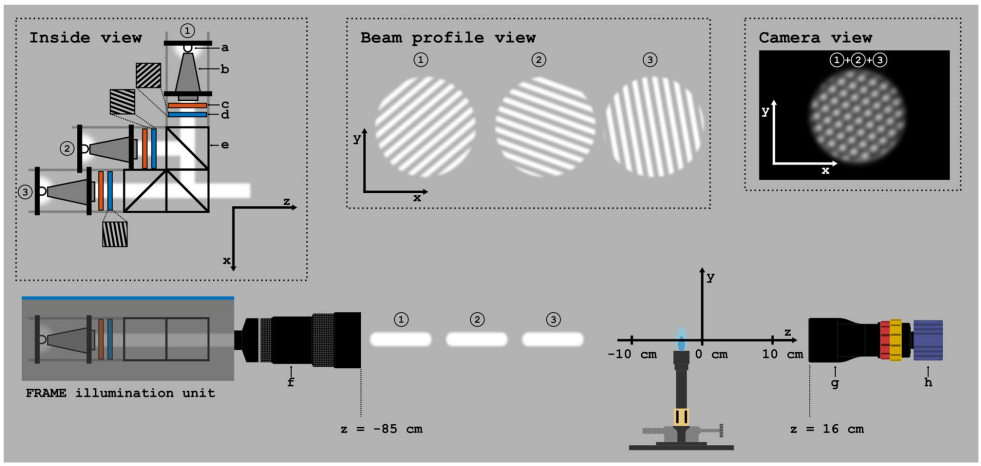


Figure 2. The experimental setup. A train of (up to) three modulated white light pulses illuminate a Bunsen burner flame and are recorded by a LaVision Imager M-lite 10-bit CMOS 5 MP camera (h). The pulse train is constructed by a FRAME illumination unit, where each optical arm consists of a fast LED (**a**), a tapered light pipe with 3x magnification (**b**), a diffusor (**c**), and a Ronchi grating (**d**). Recombination of pulses is performed by cube beam-splitters (**e**). The illumination unit is equipped with a Nikon 50–300 mm camera objectiv (**f**), while the camera (**h**) is equipped with an Edmund Optics Gold TL 0.5x telecentric objective, which has a fixed field-of-view of 26×26 mm^2. (**g**). The Ronchi gratings are projected at the plane $z = 0$ cm. The modulation line widths and the shortness of the pulses are greatly exaggerated, while the distance from illumination unit to $z = 0$ cm is reduced for illustration purposes.

$$\begin{aligned} s(\boldsymbol{r}) &= a(\boldsymbol{r})e^{i\boldsymbol{k}(\boldsymbol{r}-\boldsymbol{u}(\boldsymbol{r}))} \\ s_0(\boldsymbol{r}) &= a(\boldsymbol{r})e^{i\boldsymbol{k}\boldsymbol{r}}, \end{aligned} \tag{1}$$

where $\mathbf{k}$ is the wave vector associated with the superimposed intensity modulation[44]. Taking the phase difference

$$\phi(\boldsymbol{r}) = \arg(s(\boldsymbol{r})) - \arg(s_0(\boldsymbol{r})) = -\boldsymbol{k}\boldsymbol{u}(\boldsymbol{r}) \tag{2}$$

of the signals in Eq. 1 thus results in an equation that makes it possible to solve for the component of $\boldsymbol{u}(\boldsymbol{r})$ parallel to $\boldsymbol{k}$. To obtain the full 2D displacement field two such equations, with $\boldsymbol{k_1} \times \boldsymbol{k_2} \neq 0$, are required.

Schlieren imaging, where the path-integrated variations of the refractive index is visualized in 2D, can thus be achieved by means of intensity-modulated illumination followed by a phase-sensitive analysis of the acquired data[21,45]. The phase-determination process is demonstrated in Fig. 1, where a single intensity-modulated light field has been used to probe a Bunsen burner flame. The Fourier transform of the acquired data shows the frequency component of the superimposed sine wave (Fig. 1b). By applying a 2D spatial lock-in analysis on the acquired image the complex signal $s(\boldsymbol{r})$ (Eq. 1), containing both amplitude and phase information, carried by the modulated light pulse is retrieved. Past FRAME experiments have all been based on the extraction of the amplitude of the signal and performing such analysis on the data in Fig. 1a reveals otherwise barely discernible structures of the flame (Fig. 1c). Interfering background variations, rendered by an uneven illumination, can be mitigated through flat-field correction, i.e. subtraction of a blank recording (Fig. 1d). However, even though the amplitude image reveals structural information about the probed sample, it appears to primarily bring out finer details. For example, the unburnt- and the burnt regions appear to have similar contrast values in general, which is unexpected. We attribute these discrepancies to the fact that local variations in the amplitude of the signal are not directly related to the light displacement, i.e. schlieren data. However, as described above, the phase of the signal, rather than the amplitude, is the entity of interest for schlieren multiplexing as it carries information that can be transferred into the degree of light displacement. This transfer is achieved by calculating difference between the phase of the signal (i.e. the accumulated phase) and that of a blank recording (Eq. 2). Unfortunately the solution for the accumulated phase has many possible values (multiples of 2π) and the outcome of the phase analysis (Fig. 1e) often needs to undergo an unwrapping procedure, where all appropriate multiples of 2π are considered and discontinuities (so-called 2π-jumps) are minimized[46,47]. By properly extracting and unwrapping the accumulated phase across the FOV schlieren data is obtained, in which also smooth spatial structures in the reaction zone are visible (Fig. 1f).

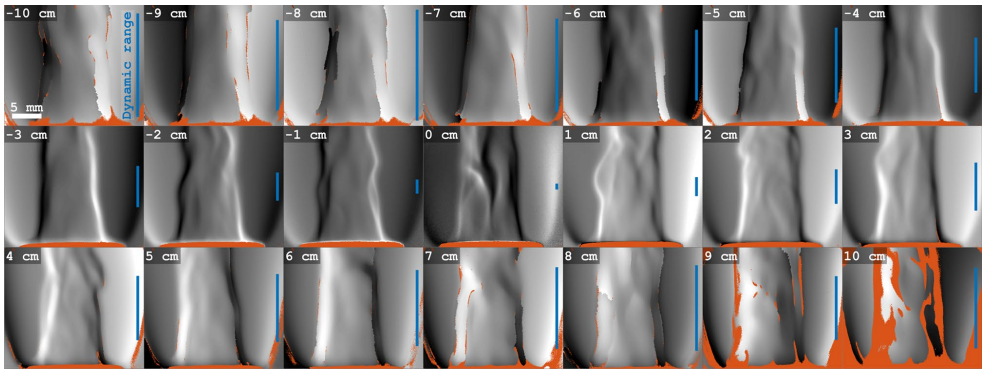


Figure 3. Phase shift images for varying target positions. The images have been acquired and reconstructed according to the FRAME method, followed by phase unwrapping. Negative distances correspond to positions closer to the illumination. Orange areas indicate where the amplitudes were too low for accurate phase determination. The blue lines show the relative dynamic range of respective image, i.e. the difference between the maximum and the minimum phase value.

Experimental setup. To investigate the feasibility of combining FRAME image multiplexing and schlieren imaging, a LaVision FRAME illumination unit was used[38]. Apart from this unit, the setup consists of a 50 fps 10-bit 5 MP sCMOS camera (LaVision, Imager M-lite) equipped with a telecentric objective lens (Edmund Optics, Gold TL 0.5x) and a Bunsen burner that can be translated along the optical axis. The general layout of the setup is sketched in the lower part of Fig. 2. The modulated light pulses are supplied by the FRAME illumination unit, whose optical layout is shown in the upper left corner ("Inside view"). It consists of three optical arms, which are combined using beam splitters. At the start of each arm sits a Ce:YAG white light LED (Cree XHP35) with a 10 to 90% rise time of 205 ns, a 100 to 10% fall time of 500 ns and temporal jitter FWHM of 35 ns. The LEDs can be controlled individually, making it possible to set the duration of each pulse as low as 0.2 µs without restrictions in the inter-frame time. Ronchi gratings (50 lp/mm) are used in each optical arm to imprint a spatial modulation onto the intensity profile of the light pulses ("Beam profile view"). When observed through the camera, the grating period corresponds to a modulation period of 3.4 pixels per line pair. Once combined, the light pulses pass through a Nikon 50-300 mm camera objective, which images the gratings at the image plane (z = 0 cm). The temperature gradients in the Bunsen flame refract the light, thus inducing phase-shifts in the superimposed spatial modulations, which are imaged on the detector within the same acquisition ("Camera view"). The FOV of the proof-of-concept system is currently set to 26×26 mm^2, restricted by the telecentric lens. Note, however, that schlieren based on intensity-modulation can be used to visualize larger objects[34], as demonstrated in Fig. S4, where the FOV is 110×110 mm^2 and Supplementary video 2 and 3.

Results

The effect of object position on sensitivity. One of the main drawbacks with a BOS system is that the image plane of the textured background (or modulated illumination) must not coincide with the object plane of the sample under study[4,48]. As the analysis of the displacement relies on visualizing fine distortions in the background pattern, the object inevitably needs to be placed out-of-focus from the cameras perspective. This requirement leads to a trade-off between sensitivity and spatial resolution; a large distance between object- and image plane yields greater displacements but allows only for visualization of the main features of the sample. The opposite case, where the object- and the image plane coincide, reveals finer structural details but the sensitivity in quantifying the displacement may be significantly reduced (especially for flatter objects). Our setup is a variation of a traditional BOS arrangement, meaning that it also is subject to this trade-off. To investigate the relation between spatial resolution and measurement sensitivity inherent in our system, 21 schlieren images of a Bunsen burner flame were acquired at different object positions, see Fig. 3.

The results in Fig. 3 show the expected trend; a greater distance between the image plane and the position of the object yields a more pronounced phase shift of the incident intensity modulation. When the burner is placed at $z = 0$ cm, i.e. at the image plane, the dynamic range of the phase shift is reduced by a factor of ~17, compared to $z = \pm 10$ cm. At the same time, positioning the object too far away yields a reduction in the amplitude of the superimposed intensity modulation that, in turn, makes accurate phase determination challenging. Regions where the signal amplitude is less than the average noise level—and thus deemed unfit for accurate phase reconstruction—are indicated by the orange areas, which can be seen to increase in size as the burner is placed farther from the image plane. These results are in consonance with established BOS theory, which states that the sensitivity of the system is approximately proportional to the distance between the object of study and the reference background[49]. As for all BOS systems, the trade-off between sensitivity and the distance between the object and the focal plane must thus be considered. For the following results, the Bunsen burner was placed at $z = -3$ cm.

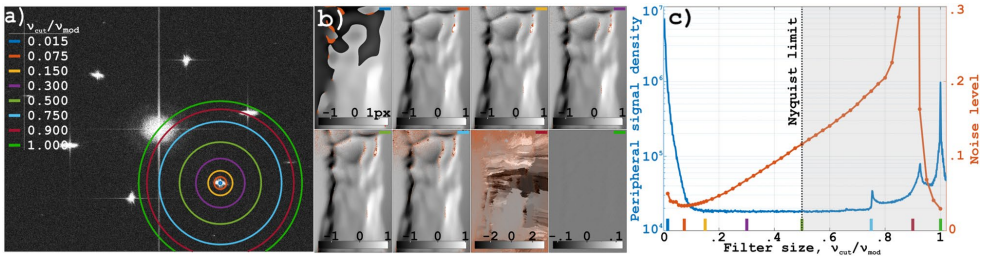


Figure 4. The effect of filter size on image quality. (**a**) The absolute value of the Fourier transform of the original image. The rings indicate the sizes of the filters used to reconstruct the signal in the Fourier cluster they are centered on. In the legend the filter specific ratio between its cut-off frequency and the modulation frequency is given. (**b**) Light displacement in pixels for different filter sizes. The colors in the upper right corners show the filter size, according to the legend in (**a**). (**c**) The measurement's peripheral signal density, defined as the total signal at the (one pixel wide) filter edge divided by the area of the filter edge (blue line, left y-axis), and the noise level, defined as the average standard deviation of 10 000 random small areas (10 × 10 px), in images as those in (**b**), for 49 different filter sizes (orange, right y-axis).

The effect of low-pass filter size. Another factor that influences the spatial resolution in the schlieren images is the characteristics of the low-pass filter used to isolate and extract the multiplexed images in the data post-processing. In general, a wider filter permits higher spatial frequencies but increases the risk of cross-talk between neighboring images (in the Fourier domain). However, a too wide filter would not necessarily amount in more structural details, as explained by the Nyqvist-Shannon sampling theorem[50,51]. This theorem states that a single intensity modulated field with a carrier frequency of ν_{mod} can, at best, be used to extract spatial variations corresponding to $\nu_{mod}/2$.

To investigate how the filter size influences the quality of the extracted multiplexed schlieren images,a set of 49 low-pass filters with different bandwidths were used to extract the bottom right sub-image of Fig. 6. The bandwidth of each filter is characterized by a cut-off frequency ν_{cut}, which can be expressed as a fraction of the modulation frequency ν_{mod} (see Fig. 4a).The schlieren images in Fig. 4b show that using a filter with $\nu_{cut} = 0.015\nu_{mod}$ yields an unsatisfactory result, having both a poor spatial resolution and an unsuccessful unwrapping. However, already for $\nu_{cut} \geq 0.075\nu_{mod}$ the unwrapping becomes stable and the signal appears to be fully captured; increasing ν_{cut} further does not reveal any new structures in this particular object, but instead the noise level increases. This effect is quantitatively shown in the graph in Fig. 4c, where the orange line represents the noise level in the extracted images (measured by taking the average standard deviation of 10 000 randomly selected areas of 10x10 pixels) as a function of ν_{cut}. The graph shows that the noise level reaches a minimum when applying a filter with a $\nu_{cut} \approx 0.075\nu_{mod}$. For lower values of ν_{cut}, the noise increases due to unsuccessful phase unwrapping and hence the introduction of artificial sharp edges in the image. For higher values of ν_{cut}, cross-talk with neighboring images are partly responsible for the increase in the noise level, yet we attribute this increase primarily to the fact that image signals, in general, drop as $1/\nu$ and a larger filter bandwidth would therefore mostly add noise. Plotted in Fig. 4c is also the signal density at different distances from the Fourier cluster, shown as a blue line. Close to the center of the Fourier cluster the signal density is high, but decreases rapidly as the distance to the centre increases. From $\nu_{cut} \approx 0.15\nu_{mod}$ to $\nu_{cut} \approx 0.75\nu_{mod}$ there is a plateau of low signal density, which indicates that the contribution in this area is primarily the constant background noise. Beyond $\nu_{cut} \approx 0.75\nu_{mod}$ the signal density increases as a result of other Fourier clusters being included, i.e. cross-talk. These results suggest that a bandwidth of $0.075\nu_{mod}$ is sufficient for the current measurements.

The effect of modulation period on sensitivity. A model that generates synthetic data, closely resembling the experimental data, was developed to study how the period ($1/\nu_{mod}$) of the intensity modulations affects the signal-to-noise ratio (SNR) of the detected phase shift. The synthetic data contains a central region where a variable phase is added, yielding a displacement of the modulations in this region (see Fig.S1–S3). After applying the phase analysis process, described above, the SNR is determined as the measured difference in phase between the two regions divided by the average standard deviation over the entire image. This analysis is repeated as the displacement is scanned from 0.01 pixels to the modulation period, which, in turn, is varied from 2 to 10 pixels. The sensitivity, as a function of modulation frequency, is then determined by calculating the median SNR over each scan (see Fig. S2), yielding the trend shown in Fig. 5.From this trend it can be concluded that the sensitivity of schlieren signals based on intensity-modulation depends highly on the modulation period, with a maximum at 3 pixels. The experimental modulation period (3.4 pixels) is thus close to the optimum period. According to these results, sub-pixel displacements—down to 0.02 pixels (SNR ≈ 2)—can be achieved under the current experimental conditions. See Supplementary information for additional information and example images.

Single-shot optically multiplexed schlieren videography. The purpose of FRAME is to simultaneously store more than one signal (image) on a single detector—an ability enabled by means of optical multiplex-

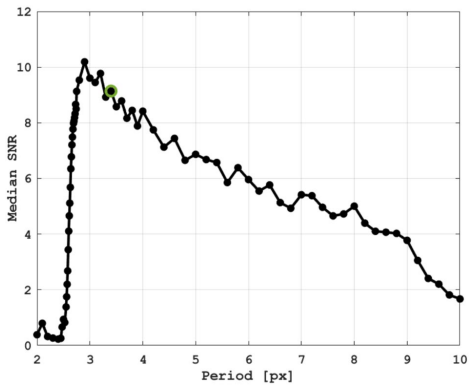


Figure 5. Modelled signal-to-noise ratio as function of modulation period. Each data point is a modelled median signal-to-noise ratio for a specified modulation period. The model is described in the Supplementary information. The highlighted data point (green circle) corresponds to the modulation period of the experimental data. The analysis shows that the measurement sensitivity is poor at low- and high modulation frequencies (near the Nyquist limit) and that a periodicity of 3 pixels yields highest sensitivity.

ing. To date, all experiments using FRAME have been based on quantifying the local amplitude of the superimposed intensity modulated pattern associated with each multiplexed image. This amplitude is either reduced by the sample through e.g. light extinction[38,39] or is increased due to e.g. fluorescence[43]. In this study we investigate the possibility to quantify the local phase of a set of temporally distinct, yet multiplexed intensity modulated image signals, which ultimately would open up for ultrafast schlieren videography. To investigate this and determine whether refractive index gradients in transparent objects can be visualized by means of FRAME, 4 video sequences—each composed of 3 frames - with a respective inter-pulse time of 1, 10, 100 or 1000 µs of a Bunsen burner flame was captured (Fig. 6). To avoid temporal overlap between images in the two fastest sequences, the pulse duration was set to be equal to the inter-pulse time (1 and 10 µs). For the other two sequences the pulse durations were set to 30 µs, to avoid motion blur and overexposing the detector. These differences in exposure time (pulse duration) will influence the SNR, which inevitably becomes lower as the exposure time is reduced. This reduced SNR can be noticed both in the unprocessed camera recordings (see magnified regions) as well as in the extracted sequences.

The steepest temperature gradient in the Bunsen flame is at the reaction zone (flame front), which is the sub-millimeter boarder between the burnt and the unburnt air/propane mixture[52]. The majority of light refraction occurs in this interface and the schlieren images are therefore expected to primarily carry information about this region. The unprocessed image (as recorded by the camera) for each sequence is presented in Fig. 6, wherein a faint imprint of the wrinkled structure of the reaction zone can be discerned (likely as Moiré patterns). This imprint diminishes as the Δt between the multiplexed frames increases because of the motion of the flame, which will smooth out structural details. By applying the spatial lock-in analysis on the acquired image and determining the local phase shifts, the progression of the complexly wrinkled flame front is revealed with high contrast. At short Δt the wrinkling of the flame is frozen in time, as illustrated by the spatially overlapping highlighted structures (colored lines), while the corresponding structures do not overlap in space for sequences with a Δt of 100 and 1000 µs. The extracted schlieren images show no signs of cross-talk, which would be manifested as intensity-modulated structures. These results and findings demonstrate that optical multiplexing of temporally separated images can indeed be united with schlieren sensitive imaging.

One drawback with the image-coding methodology associated with FRAME multiplexing is that it is one-dimensional, i.e. each light pulse is intensity modulated by a single spatial frequency v_{mod}, which does not permit the estimation of the displacement vector component parallel to the stripes[20]. Since the image-coding of FRAME relies on the multiplexed images having different modulated patterns, this inability poses a problem; the individual frames will have uneven capabilities in resolving structural details depending on the orientation of the superimposed intensity modulation. This difference in resolving capabilities can be seen in the image sequence with a Δt of 1 µs, where, despite the inter-frame time being sufficiently short to temporally freeze the rapidly evolving flame structures, certain structures in which the flame fronts appear different can be identified. Such structural differences can be mitigated by reducing the angle between the intensity modulations in the illumination, yet with an increased risk of image cross-talk.

Schlieren videography with full vector displacement. A more complete solution to the problem of measuring spatial variations in refractive index orthogonal to the modulation was presented in 2018 by Wildeman[20]; by using a checkered background pattern, i.e. composed of multiplicatively combined vertical and horizontal binary lines, the phase shift (and hence the light displacement) can be estimated in both the x- and

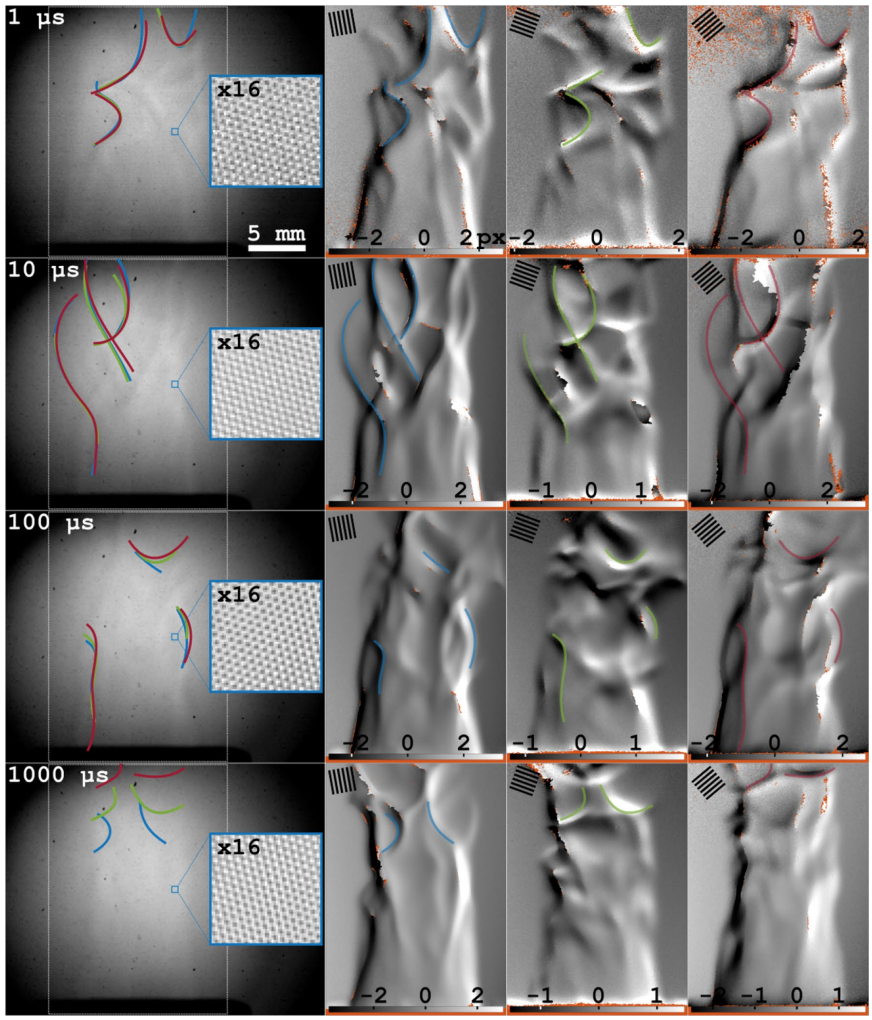


Figure 6. Video sequences of flame induced light displacement. The left column show the original image for each sequence. The blue, green and red lines highlight structures that are considered identical in image 1, 2 and 3, respectively. These colored lines are present both in the original image and in the image they originate from. The inter-frame time of each sequence is shown in the upper left corner of the original image. The orientations of the spatial modulations, used for acquisition, are shown for respective image.

y-direction. To investigate the feasibility of combining this approach for full vector displacement with temporal image multiplexing, a laser-based FRAME setup was used, capable of producing four intensity-modulated laser pulses with an intra-pair time separation of 6 µs. Using this setup, the Bunsen flame could thus be illuminated with a set of two temporally separated intensity modulated pairs of laser pulses with orthogonal modulations.

Figure 7a displays an image of the Bunsen flame captured using the laser-based FRAME setup. The images in Fig. 7b show the magnitude of the light displacement induced by the flame at unique times and in unique directions. The direction and magnitude of the displacements are shown as arrow plots in Fig. 7c. In Fig. 7d the arrow plots associated with (approximately) the same time have been combined by adding the components of the vectors, yielding two arrow plots showing both direction and magnitude of the full 2D vector displacement induced by the flame at two different occasions. These results illustrate the possibility of using image multiplexing in order to, in 2D, track complex structural details of transparent objects that evolve at high speeds.

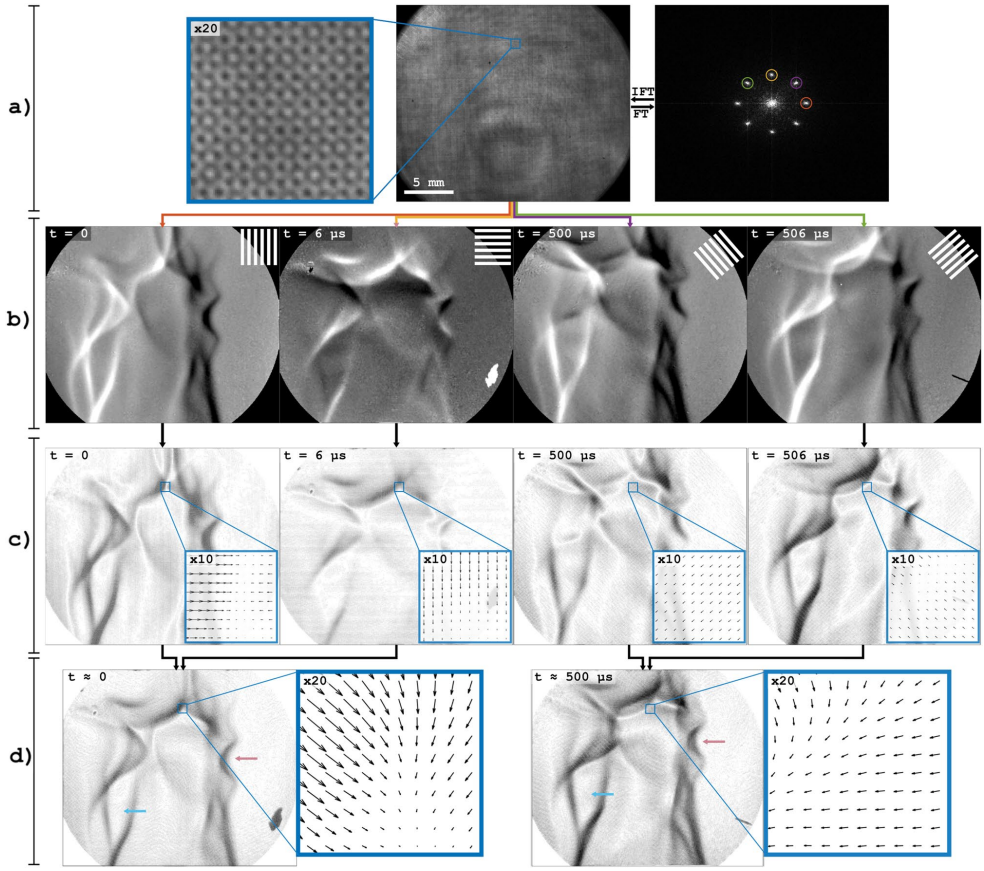


Figure 7. Videography using orthogonal modulations. (**a**) The original image and the absolute value of its Fourier transform. (**b**) Images showing the magnitude of the flame induced displacement. (**c**) Magnitude and direction of the displacement, shown through arrow plots. (**d**) The displacement vectors of both signal pairs, i.e. signals recorded with the minimum possible temporal separation of 6 µs, are added to get the full 2D vector displacement at both recording instances. The red and blue arrows point at (what is considered) identical structures in the flame. The zoomed in sections are of identical position.

Discussion

The family of schlieren-like techniques is a powerful set of tools that can be used to visualize and even measure refractive index gradients in transparent media. However, no schlieren technique has previously been demonstrated in combination with videography based on image multiplexing. Schlieren videography has thus been confined to acquisition speeds dictated by camera frame rates, limiting what phenomena can be studied accordingly. In this study we exploit intensity-modulated illumination in order to (1) qualitatively visualize the refractive index gradients of the object and (2) multiplex a set of temporally separated images from which schlieren data can be extracted. This FRAME-based schlieren technique resembles synthetic schlieren and FCD, but differs in one crucial aspect; instead of a static backdrop it utilizes a dynamic background, in the form of a burst of modulated light pulses. The implication of this is that the frame rate is not set by the camera, but by how short light pulses the illumination source can deliver. By instead relying on rapid illumination, slow yet high-resolution sensors can be used for fast videography. Further, by using both LED and laser sources to produce the modulated light burst, we show that the functionality of the technique is independent on whether or not the light source is coherent.

A drawback with the FRAME technique is that the number of frames in the sequence is limited. However, it has previously been demonstrated that hundreds of signals can be multiplexed in this way[39], and our results (Fig. 4b) suggest that it could be sufficient with a Fourier space area as small as 100×100 pixels ($\nu_{cut} = 0.075\nu_{mod}$ corresponds to a filter radius of ~50 pixels) to capture all relevant signal, while still avoiding cross-talk between

signals. Given these conditions, approximately 30 image signals could be multiplexed on a 4 Mp sensor using 4 px/lp modulations (corresponding to Fourier clusters 1000 pixels from the centre, so that $n \approx 1000\pi/100 \approx 30$). However, perhaps the main challenge associated with increasing the number of multiplexed signals by a certain factor x is that the level of the individual signals will have to decrease by the same factor to avoid overexposing the detector. The noise that the signals are competing against is however global, and will therefore not decrease, meaning that the SNRs will decrease by the factor x. Choosing the smallest possible filter size, to reject as much noise as possible, while still capturing the full signal, will thus be crucial for long video sequences. For more details on the challenges—both concerning the illumination as well as the analysis—with multiplexing long video sequences, see Ref.[39].

In previous FRAME studies, the amplitude has been the sought part of the reconstructed signals. For most applications, e.g. those based on absorption, extinction and fluorescence, the amplitude is the important aspect of the signal, but when studying transparent objects it is not ideal, even though the flame refractive index gradients can be seen to induce amplitude variation (Fig. 1c,d). This is because the relation between object's refractive index and the observed amplitude is dependent on how the optical signal is digitized, and hence not easily interpreted. The phase shift, however, is directly related to the displacement of light, which is a result of the encountered refractive index gradients. It can thus reveal information about the structure of the object under study. Here we show that the phase can be retrieved after multiplexing 4 images on the same detector.

Based on the results in this study it is evident that FRAME can be used to generate schlieren images of dynamic transparent events, revealing how the refractive index within an object changes. To the best of the authors knowledge this is the first time multiplexed illumination-based schlieren videography has been demonstrated. Since, for these techniques, the achievable frame rate is solely limited by how short light pulses one can, produce, this paves the way for future ultra-high speed schlieren videography.

Data availability

Data underlying the results presented in this paper are available from the corresponding author upon request.

Received: 26 August 2022; Accepted: 26 October 2022
Published online: 08 November 2022

References

1. Rienitz, J. Schlieren experiment 300 years ago. *Nature* **254**, 293 (1975).
2. Settles, G. S. *Schlieren and Shadowgraph Techniques: Visualizing Phenomena in Transparent Media* (Springer Science & Business Media, 2001).
3. Toepler, A. J. I. *Beobachtungen nach einer neuen optischen methode: Ein beitrag experimentalphysik* (Max Cohen & Son, 1864).
4. Settles, G. S. & Hargather, M. J. A review of recent developments in schlieren and shadowgraph techniques. *Meas. Sci. Technol.* **28**, 042001 (2017).
5. Estruch, D., Lawson, N., MacManus, D., Garry, K. & Stollery, J. Measurement of shock wave unsteadiness using a high-speed schlieren system and digital image processing. *Rev. Sci. Instrum.* **79**, 126108 (2008).
6. Kleine, H. & Grönig, H. Color schlieren methods in shock wave research. *Shock Waves* **1**, 51 (1991).
7. Hayasaka, K., Tagawa, Y., Liu, T. & Kameda, M. Optical-flow-based background-oriented schlieren technique for measuring a laser-induced underwater shock wave. *Exp. Fluids* **57**, 1 (2016).
8. Tanda, G., Fossa, M. & Misale, M. Heat transfer measurements in water using a schlieren technique. *Int. J. Heat Mass Transf.* **71**, 451 (2014).
9. Narayan, S., Srivastava, A. & Singh, S. Rainbow schlieren-based investigation of heat transfer mechanisms during isolated nucleate pool boilingphenomenon: Effect of superheat levels. *Int. J. Heat Mass Transf.* **120**, 127 (2018).
10. Lyn, W. & Valdmanis, E. The application of high speed Schlieren photography to diesel combustion research. *J. Photogr. Sci.* **10**, 74 (1962).
11. Ben-Yakar, A., Kamel, M., Morris, C., & Hanson, R. Hypersonic combustion and mixing studies using simultaneous oh-plif and schlieren imaging. In *36th AIAA Aerospace Sciences Meeting and Exhibit,* (1998), p. 940.
12. Agrawal, A. K., Butuk, N. K., Gollahalli, S. R. & Griffin, D. Three-dimensional rainbow schlieren tomography of a temperature field in gas flows. *Appl. Opt.* **37**, 479 (1998).
13. Siewert, E. *et al.* Visualization of gas flows in welding arcs by the Schlieren measuring technique. *Weld. J.* **93**, 1S (2014).
14. Degen, N. An overview on Schlieren optics and its applications: Studies on mechatronics. ETH-Zurich (2012).
15. Burton, R. A. A modified schlieren apparatus for large areas of field. *JOSA* **39**, 907 (1949).
16. Heineck, J. T., Banks, D., Schairer, E. T., Haering, E. A., & Bean, P. Background oriented schlieren (bos) of a supersonic aircraft in flight. In *AIAA flight testing conference,* (2016), p. 3356.
17. Raffel, M. Background-oriented schlieren (BOS) techniques. *Exp. Fluids* **56**, 1 (2015).
18. Richard, H. & Raffel, M. Principle and applications of the background oriented schlieren (BOS) method. *Meas. Sci. Technol.* **12**, 1576 (2001).
19. Hatanaka, K., Saito, T. Background oriented schlieren method using multi-scale periodic pattern. In *29th International Symposium on Shock Waves 1,* pp. 453–458.
20. Wildeman, S. Real-time quantitative Schlieren imaging by fast Fourier demodulation of a checkered backdrop. *Exp. Fluids* **59**, 1 (2018).
21. Dalziel, S. B., Hughes, G. O., & Sutherland, B. R. Synthetic schlieren. In *Proceedings of the 8th International Symposium on Flow Visualization,* , vol. 62 (1998).
22. Wilson, S., Gustafson, G., Lincoln, D., Murari, K. & Johansen, C. Performance evaluation of an overdriven LED for high-speed schlieren imaging. *J. Vis.* **18**, 35 (2015).
23. Ciccarelli, G., Johansen, C. & Kellenberger, M. High-speed flames and DDT in very rough-walled channels. *Combust. Flame* **160**, 204 (2013).
24. Mizukaki, T., Wakabayashi, K., Matsumura, T. & Nakayama, K. Background-oriented schlieren with natural background for quantitative visualization of open-air explosions. *Shock Waves* **24**, 69 (2014).
25. Winter, K. O. & Hargather, M. J. Three-dimensional shock wave reconstruction using multiple high-speed digital cameras and background-oriented schlieren imaging. *Exp. Fluids* **60**, 1–13 (2019).
26. Ben-Yakar, A. & Hanson, R. Ultra-fast-framing schlieren system for studies of the time evolution of jets in supersonic crossflows. *Exp. Fluids* **32**, 652 (2002).
27. Etoh, T. G. *et al.* The theoretical highest frame rate of silicon image sensors. *Sensors* **17**, 483 (2017).

28. Ehn, A. *et al.* FRAME: Femtosecond videography for atomic and molecular dynamics. *Light Sci. Appl.* **6**, e17045 (2017).
29. Liang, J. & Wang, L. V. Single-shot ultrafast optical imaging. *Optica* **5**, 1113 (2018).
30. Nakagawa, K. *et al.* Sequentially timed all-optical mapping photography (STAMP). *Nat. Photon* **8**, 695 (2014).
31. Cortizo, E., Moreno Yeras, A., Lepore, J. R. & Garavaglia, M. Application of the structured illumination method to study the topography of the sole of the foot during a walk. *Opt. Lasers Eng.* **40**, 117 (2003).
32. Gustafsson, M. G. L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci.* **102**, 13081 (2005).
33. Joannes, L., Dubois, F. & Legros, J.-C. Phase-shifting schlieren: High-resolution quantitative schlieren that uses the phase-shifting technique principle. *Appl. Opt.* **42**, 5046 (2003).
34. Psota, P. *et al.* Quantitative Schlieren imaging based on fringe projection. *EPJ Web Conf.* **264**, 01034 (2022).
35. Takeda, M., Ina, H. & Kobayashi, S. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *J. Opt. Soc. Am.* **72**, 156–160 (1982).
36. Qian, K., Seah, H. S. & Asundi, A. Fault detection by interferometric fringe pattern analysis using windowed Fourier transform. *Meas. Sci. Technol.* **16**, 1582 (2005).
37. Kristensson, E., Ehn, A., Bood, J. & Aldén, M. Advancements in Rayleigh scattering thermometry by means of structured illumination. *Proc. Combust. Inst.* **35**, 3689 (2015).
38. Kornienko, V., Kristensson, E., Ehn, A., Fourriere, A. & Berrocal, E. Beyond MHz image recordings using LEDs and the FRAME concept. *Sci. Rep.* **10**, 1 (2020).
39. Ek, S., Kornienko, V. & Kristensson, E. Long sequence single-exposure videography using spatially modulated illumination. *Sci. Rep.* **10**, 1 (2020).
40. Kornienko, V. *et al.* Simultaneous multiple time scale imaging for kHz and MHz high-speed accelerometry. *Photon. Res.* **10**, 1712 (2022).
41. Kristensson, E., Li, Z., Berrocal, E., Richter, M. & Aldén, M. Instantaneous 3D imaging of flame species using coded laser illumination. *Proc. Combust. Inst.* **36**, 4585 (2017).
42. Dorozynska, K., Kornienko, V., Aldén, M. & Kristensson, E. A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Opt. Express* **28**, 9572 (2020).
43. Dorozynska, K. & Kristensson, E. Implementation of a multiplexed structured illumination method to achieve snapshot multispectral imaging. *Opt. Express* **25**, 17211 (2017).
44. Takeda, M. & Mutoh, K. Fourier transform profilometry for the automatic measurement of 3-D object shapes. *Appl. Opt.* **22**, 3977 (1983).
45. Sutherland, B. R., Dalziel, S. B., Hughes, G. O. & Linden, P. Visualization and measurement of internal waves by 'synthetic schlieren' Part 1. Vertically oscillating cylinder. *J. Fluid Mech.* **390**, 93 (1999).
46. Kasim, M. Fast 2D phase unwrapping implementation in MATLAB. GitHub (2017).
47. Van der Walt, S. *et al.* scikit-image: Image processing in Python. *Peer J.* **2**, e453 (2014).
48. Bichal, A. & Thurow, B. On the application of background oriented schlieren for wavefront sensing. *Meas. Sci. Technol.* **25**, 015001 (2013).
49. Meier, G. Computerized background-oriented schlieren. *Exp. Fluids* **33**, 181 (2002).
50. Shannon, C. E. Communication in the presence of noise. *Proc. IRE* **37**, 10 (1949).
51. Nyquist, H. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Electr. Eng.* **47**, 617 (1928).
52. Kohse-Höinghaus, K., Barlow, R. S., Aldèn, M. & Wolfrum, J. Combustion at the focus: Laser diagnostics and control. *Proc. Combust. Inst.* **30**, 89 (2005).

Author contributions

S.E. and E.K. wrote the main manuscript text. S.E. prepared all figures in the manuscript. S.E. carried out the LED-based experiments with aid from V.K, while V.K. and E.K. carried out the laser-based experiments. S.E. and A.R. performed the data analysis. E.B., E.K. and V.K. designed the FRAME equipment. All authors reviewed the manuscript.

Funding

Open access funding provided by Lund University. European Research Council (ERC) (803634). Swedish Research Council (2019-05183 and 2016-03894).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-23198-6.

Correspondence and requests for materials should be addressed to S.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper V

Letter Vol. 46, No. 22 / 15 November 2021 / Optics Letters 5723

Optics Letters

Improved temporal contrast of streak camera measurements with periodic shadowing

Yupan Bao,[1] Vassily Kornienko,[1] David Lange,[1] Wolfgang Kiefer,[2,3] Tina Eschrich,[4] Matthias Jäger,[4] Joakim Bood,[1] Elias Kristensson,[1] and Andreas Ehn[1,*]

[1]Department of Combustion Physics, Lund University, P.O. Box 118, S-221 00 Lund, Sweden
[2]Institute for Physical and Theoretical Chemistry, University of Würzburg, 97074 Würzburg, Germany
[3]Eisingen Laboratory for Applied Raman Spectroscopy, 97249 Eisingen, Germany
[4]Leibniz Institute of Photonic Technology, Albert Einstein St. 9, 07745 Jena, Germany
*Corresponding author: andreas.ehn@forbrf.lth.se

Received 13 August 2021; revised 23 September 2021; accepted 25 September 2021; posted 28 September 2021; published 15 November 2021

Periodic shadowing, a concept used in spectroscopy for stray light reduction, has been implemented to improve the temporal contrast of streak camera imaging. The capabilities of this technique are first proven by imaging elastically scattered picosecond laser pulses and are further applied to fluorescence lifetime imaging, where more accurate descriptions of fluorescence decay curves were observed. This all-optical approach can be adapted to various streak camera imaging systems, resulting in a robust technique to minimize space-charge induced temporal dispersion in streak cameras while maintaining temporal coverage and spatial information.

https://doi.org/10.1364/OL.438034

Streak cameras [1–3] were invented by Courtney-Pratt in 1949 [4,5] in order to capture fast transient phenomena, which are beyond the temporal resolution of traditional mechano-optical cameras. Their ability to simultaneously provide spatial and temporal information at sub-picosecond (even up to attosecond [6,7]) resolution was first used to characterize ultrashort laser pulses generated by mode-locked laser systems [8]. Over the years, the streak camera has become a robust workhorse for capturing rapid dynamics of luminous phenomena within the fields of biology, chemistry, physics, and medicine, with applications ranging from fluorescence lifetime imaging [9,10] to ultrafast electron microscopy [1,2,11].

The working principle of a streak camera is illustrated in Fig. 1(a) [12]. The signal of interest is first imaged onto a slit where the one-dimensional (1D) spatial information is retained. Signal photons are converted into photoelectrons via a photocathode. These will be accelerated through a mesh before entering the sweeping tube, where a voltage ramp is applied, giving the electrons a tangential velocity perpendicular to the direction of propagation. The signal electrons, which reach the sweeping tube at different times, experience different electric fields. Consequently, they will arrive at the detector at different positions, resulting in a time-dependent spatial distribution of the signal. The resulting image, hence, retains the slit's spatial information in one dimension and the signal's temporal information in the second dimension.

A side effect of the voltage ramp inside the sweeping tube is that any initial velocity distribution of electrons entering the tube translates directly into a temporal dispersion in the final image output. Hence, a wider velocity distribution, caused by, e.g., the space-charge effect at the photocathode [14], directly translates into a broader temporal profile (pulse broadening), degrading the temporal resolution of the streak camera. Various solutions have been proposed and implemented to minimize this effect. Some examples include: inserting a reflectron to compensate for the energy spread of the initial photoelectrons by giving them a higher energy and longer flight path [15,16]; shortening the distance between the photocathode and mesh to reduce space-charge related electron pulse broadening in said region [2,3,17]; implementing specially designed photocathodes, lens systems, and streaking tubes with optimized working voltage [18,19]; and adding a two-dimensional (2D) to 1D fiber array with certain time delays between adjacent fibers to reduce the cross talk between the input channels [20,21].

In this Letter, we propose an all-optical approach to further reduce the temporal dispersion of streak camera imaging in the form of periodic shadowing (PS) [12]. Due to the low experimental complexity of this approach, it can be readily adapted to any streak camera imaging system, regardless of the sweeping method or implemented opto-electronic improvement while maintaining both temporal coverage and spatial information.

PS of the signal is here implemented in two configurations: either with a Ronchi grating [12] or an optical fiber bundle [13]. An example of a single-shot streak camera signal of two consecutive scattered laser pulses coded with PS is shown in Fig. 1(c). Its Fourier transform [Fig. 1(d)] displays separated Fourier components of the tagged signal, where the signal of interest is marked by an ellipse. The tagged signal can be extracted by first shifting its Fourier components to the origin and subsequently applying

0146-9592/21/225723-04 Journal © 2021 Optica Publishing Group

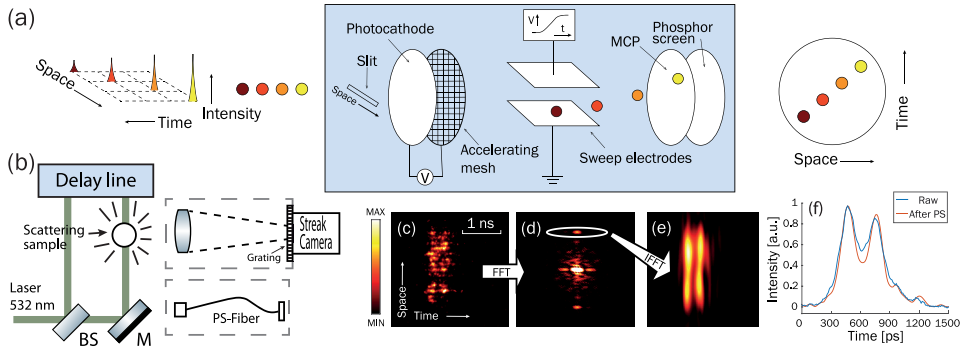


Fig. 1. (a) Simplified sketch of a streak camera. (b) A sketch of the experimental setup: BS, beam splitter; M, mirror; streak camera, OptoScope S20 with FTSI1-ST, operated in single-shot mode; laser, Ekspla SL334; grating, Ronchi Grating (5, 7, 4.8 lp/mm); PS-fiber, equivalent to the one found in [13]. The grating and PS fiber-based setups are interchangeable, without affecting the rest of the setup. (c) An example of a single-shot streak camera signal of two successive scattered laser pulses, using periodic shadowing (PS). (d) The Fourier transform of (c) where the modulated signal is encircled by the ellipse. (e) Inverse Fourier transform of the first order component filtered out by the Gaussian ellipse filter from (d). (f) Comparison of the normalized intensity of the laser pulses obtained before and after postprocessing in the Fourier domain.

a Gaussian filter. The inverse Fourier transform of the recentered and filtered image results are shown in Fig. 1(e). More technique details of PS can be found in Supplement 1. A drawback of PS is the sacrifice of spatial resolution due to the application of a low-pass filter. Hence, for streak camera applications, where the temporal resolution is of importance, an elliptical low-pass filter is chosen. Its large semi-major axis retains the temporal resolution of the streak camera while the small semi-minor axis effectively isolates the PS signal from the DC background. A comparison of the normalized intensity of the signal processed with/without PS shows clear signal reduction in the region between the two pulses [Fig. 1(f)]. Furthermore, an inherent advantage of PS is the automatic subtraction of the DC component. This is advantageous especially when the background is nonlinear to the signal, where conventional methods to record the background, e.g., a black recording without any signal, become impractical.

The improvement of temporal contrast when using PS on streak camera measurements is demonstrated by (1) imaging a single and two consecutive laser pulses incident on a scattering medium and (2) measuring the temperature-dependent fluorescence lifetime of a Rhodamine 610 solution.

The results of applying PS (7 lp/mm) to the imaging of a scattered laser pulse at a streak rate of 7 ps/px, correcting for streak camera jitter, and averaging over 100 shots are shown in Fig. 2(a). Here, the effect of PS is to eliminate artefacts that arise in the beginning and end of the peak (the total integrated area is 85.4% that of the non-PS peak), effectively reducing the FWHM from 320 to 280 ps. The inset in Fig. 2(a) illustrates the automatic baseline correction obtained from PS due to the elimination of DC background after applying the low-pass filter. The trend of lower FWHM for the PS signal continues for all streak rates until it converges with the non-PS signal at 1.4 ps/px [Fig. 2(b)]. At this streak rate, the camera has sufficient temporal resolution for single pulse measurements such that the pulse broadening caused by space-charge effects is negligible compared to the pulse length. However, in combination with

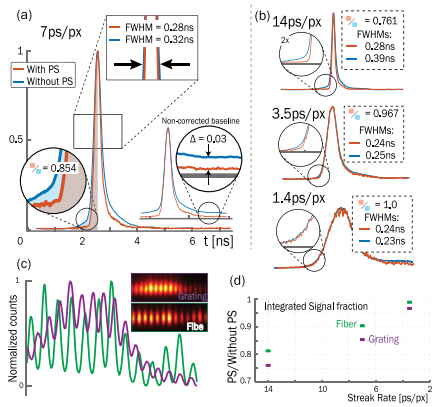


Fig. 2. Results of applying PS to the imaging of a single scattered laser pulse at (a) a streak rate of 7 ps/px and (b) higher streak rates. (c) Comparision of the modulation depth of the signal between the use of grating and fiber bundle. (d) Integration of the signal with divided by without PS at different streak rates. All data were averaged over 100 shots with jitter correction.

PS, the same temporal resolution could be achieved at lower streak rates, where higher temporal coverage and higher SNR are attained.

The measurements were also repeated with an optical fiber bundle, where the lower acceptance angle of the fiber, in conjunction with more efficient light shielding as compared to an imaging system based on a Ronchi grating, further eliminates background scattered light from entering the streak camera. This has the effect of increasing the modulation depth of the PS pattern [Fig. 2(c)], allowing for a larger share of dynamic range allocated to the modulated signal. Figure 2(d) compares the integration fraction (PS/without PS) for the grating (in this

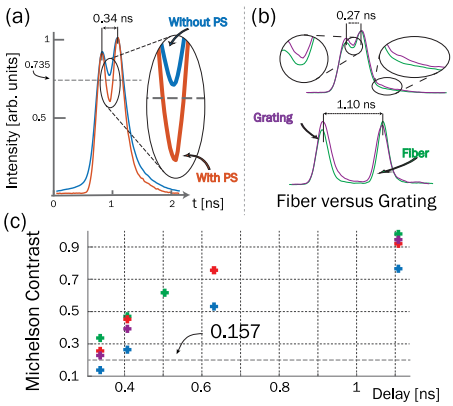


Fig. 3. (a) Two scattered Gaussian laser pulses are brought from a temporally unresolved to resolved regime using PS (with the streak of 7 ps/px). (b) Performing the same investigation with a fiber bundle further increases the temporal resolution due to the larger modulation depth. (c) Michelson contrast of the raw data (blue), grating of 5 lp/mm (red) and 7 lp/mm (purple), and fiber bundle or 4.8 lp/mm (green).

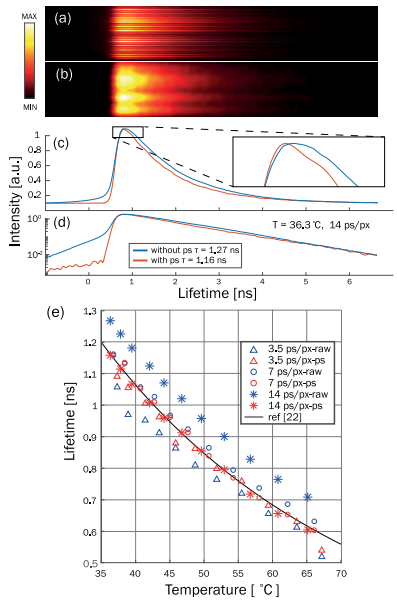


Fig. 4. Fluorescence lifetime measurements of Rhodamine 610. (a) The raw image. (b) The postprocessed image using PS. (c), (d) Intensity curves derived from the raw image (blue) and the processed image (magenta) in linear scale (c) and logarithmic scale (d). (e) Temperature-dependent lifetime calculated from the measured curve with different streak rates. The reference curve originates in the Arrhenius equation, $\frac{1}{\tau} = A \exp(\frac{-E_a}{RT})$ with E_a the apparent activation energy, equal to 19.2 kJ/mol.

case 5 lp/mm, indicated in purple) and fiber bundle (4.8 lp/mm indicated in green) as a function of streak rate. An integration fraction below unity, even for measurements that have reduced amounts of background light, (i.e., measurements with the fiber bundle), substantiates that PS not only eliminates background stray light that enters the camera but is also efficient at reducing the temporal dispersion caused by space-charge effects within the streak camera.

To further validate PS, the PS with streak camera combination was implemented in order to characterize a two-pulse picosecond pulse train. Inspired by the Rayleigh criteria for spatially resolved point sources, two Gaussian pulses are said to be temporally resolved herein if their summed intensity falls below a Michelson contrast of 0.735 (see Supplement 1). As a result of the increased temporal contrast attainable with PS, two consecutive scattered Gaussian laser pulses were brought from a temporally unresolved to resolved regime [Fig. 3(a)]. Although not as critical, this trend continues for increased inter-pulse delay times, for which a higher Michelson contrast is maintained [Fig. 3(c)]. Moreover, performing PS with a fiber bundle has the effect of even further increasing the contrast between two pulses [Figs. 3(b) and 3(c)].

Temperature-dependent fluorescence lifetime measurements of a Rhodamine 610 and water solution (3×10^{-9} M) were carried out with a 5 lp/mm Ronchi grating and streak rates of 14 ps/px, 7 ps/px, and 3.5 ps/px. A streak rate of 1.4 ps/px could not be applied to these measurements as the whole decay would not fit on the sensor. Furthermore, the temporal jitter is too large at this streak rate (in relation to the temporal coverage) to allow for efficient jitter correction. This fact exemplifies the need for PS, as it can increase the temporal resolution (in the same way that increasing streak rate does) while maintaining temporal coverage and the possibility to perform jitter correction. An example is given in Figs. 4(a)–4(d), where a significant amount of background has been eliminated with PS, especially in the rising edge. Although not as obvious, the interfering background on the falling edge has also been filtered out, increasing the accuracy of the decay curve for this Rhodamine sample. Accuracy in this part of the signal is critical for lifetime measurements, as a slight difference in its shape can cause a significant change in the calculated lifetime. For example, the difference in lifetime calculated from the curve processed with and without PS is around 100 ps (at a streak rate of 14 ps/px). This corresponds to more than 15% of the actual lifetime, about 600 ps, of Rhodamine 610 at 65°C. The calculated lifetimes from both the PS-corrected and original curves are plotted in Fig. 4(e) with reference data (from [22]). Good agreement between [22] and the PS measurements is observed. Detailed data of the decay curves of fluorescence lifetime measurements and parameter study of the size of the ellipse filter applied using PS can be found in Supplement 1.

In contrast to Fig. 2(b), where PS has a smaller effect for the calculation of FWHM at higher streak rates, in the context of lifetime measurements, the increased accuracy of lifetime evaluation is significant even for streak rates of 3.5 ps/px. This should be credited to the efficient removal of background interference in the form of stray light and space-charge effect, which in itself is a sensitive parameter for accurate lifetime determination. Moreover, unlike the measurements with streak rates of 14 and 7 ps/px, the lifetime calculated from the processed curve

is not only more accurate but also longer with a streak rate of 3.5 ps/px. This shows that PS not only deals with stray light, whose removal via the subtraction of a black field image always leads to a shorter lifetime, but also background interference introduced by the aforementioned space-charge effect.

The underlying principle of a streak camera is its ability to convert temporal into spatial information. This allows for its temporal resolution to be independent of exposure time, frame-rate, and pixel readout speed (as is the case with conventional cameras) setting its minimum temporal dispersion to the streak rate in seconds per pixel. However, the complex photon-electron conversion steps involved result in an actual temporal dispersion that is instead limited by, e.g., space-charge effects on the streak unit, blooming in the multi-channel plate, and imperfections in the imaging system.

The optical configuration of PS has several technical advantages over traditional, electronic-based streak camera improvements. First, applying a grating to the imaging configuration of the setup is a versatile approach that can be readily used and tailored to fit to a specific streak camera measurement configuration. Second, even though the camera can attain streak rates that are high enough to temporally resolve all laser pulse measurements presented herein, it does so at the cost of temporal coverage and decreased signal per pixel (resulting in decreased SNR). PS compensates for this by allowing for higher temporal resolution at a given streak rate, hence maintaining the higher signal per pixel available at lower streak rates. Third, the use of a fiber bundle, in itself resulting in an increased modulation depth and final temporal contrast as compared to a grating, grants this technique (in its fiber-based configuration) access to difficult *in situ* measurement configurations, such as a pressure/vacuum chamber with limited optical access. Finally, the significant decrease of pulse broadening, along with the elimination of the DC component, allows for accurate streak camera-based optical lifetime measurements, introducing an alternative to conventional lifetime measurement configurations.

The increased necessity of high-speed optical measurement techniques across the sciences has brought the streak camera into a diverse set of labs. However, at these breakneck speeds, its temporal resolution is usually limited by pulse broadening effects within the electronics. Even though there exist solutions, these are always targeted at the opto-electronic components, which have already been pushed to their extremes. The work presented herein offers a versatile, all-optical approach to simultaneously increasing the temporal resolution of a streak camera while maintaining temporal coverage and spatial information, adapting streak cameras to applications within a wider range of fast transient events.

Funding. European Research Council (669466, 803634, 852394); Vetenskapsrådet (2019-05183, 2015-05321); Knut och Alice Wallenbergs Stiftelse (2019.0084); Energimyndigheten (22538-4).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Some of the data underlying the results presented in this Letter are available in Supplement 1. The other data can be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. A. Campillo and S. Shapiro, IEEE J. Quantum Electron. **19**, 585 (1983).
2. C. Johnson, *High Speed Optical Techniques: Developments and Applications* (SPIE, 1977), Vol. **94**, pp. 13–18.
3. A. Huston, J. Phys. E **11**, 601 (1978).
4. J. Courtney-Pratt, Research **2**, 287 (1949).
5. J. Courtney-Pratt, Proc. R. Soc. London A **204**, 27 (1950).
6. J. Itatani, F. Quéré, G. L. Yudin, M. Y. Ivanov, F. Krausz, and P. B. Corkum, Phys. Rev. Lett. **88**, 173903 (2002).
7. M. Ivanov and O. Smirnova, Phys. Rev. Lett. **107**, 213605 (2011).
8. D. Bradley and G. H. New, Proc. IEEE **62**, 313 (1974).
9. K. Suhling, L. M. Hirvonen, J. A. Levitt, P.-H. Chung, C. Tregidgo, A. Le Marois, D. A. Rusakov, K. Zheng, S. Ameer-Beg, S. Poland, S. Coelho, R. Henderson, and N. Krstajic, Med. Photon. **27**, 3 (2015).
10. J. Qu, L. Liu, D. Chen, Z. Lin, G. Xu, B. Guo, and H. Niu, Opt. Lett. **31**, 368 (2006).
11. W. E. King, G. H. Campbell, A. Frank, B. Reed, J. F. Schmerge, B. J. Siwick, B. C. Stuart, and P. M. Weber, J. Appl. Phys. **97**, 111101 (2005).
12. E. Kristensson, J. Bood, M. Alden, E. Nordström, J. Zhu, S. Huldt, P.-E. Bengtsson, H. Nilsson, E. Berrocal, and A. Ehn, Opt. Express **22**, 7711 (2014).
13. M. Gong, H. Kim, J. Larsson, T. Methling, M. Aldén, E. Kristensson, C. Brackmann, T. Eschrich, M. Jäger, W. Kiefer, and A. Ehn, Opt. Express **29**, 7232 (2021).
14. Z. Tao, H. Zhang, P. M. Duxbury, M. Berz, and C.-Y. Ruan, J. Appl. Phys. **111**, 044316 (2012).
15. U. Boesl, R. Weinkauf, and E. W. Schlag, Int. J. Mass Spectrom. Ion Processes **112**, 121 (1992).
16. Y. Wang and N. Gedik, IEEE J. Sel. Top. Quantum Electron. **18**, 140 (2012).
17. B.-L. Qian and H. E. Elsayed-Ali, J. Appl. Phys. **91**, 462 (2001).
18. H. Niu, H. Zhang, Q. L. Yang, Y. P. Liu, Y. C. Wang, Y. A. Reng, and J. L. Zhou, Proc. SPIE **1032**, 472 (1989).
19. H. Niu, V. P. Degtyareva, V. N. Platonov, A. M. Prokhorov, and M. Y. Schelev, Proc. SPIE **1032**, 79 (1989).
20. N. Hirmiz, A. Tsikouras, E. J. Osterlund, M. Richards, D. W. Andrews, and Q. Fang, Opt. Express **27**, 22602 (2019).
21. A. Tsikouras, J. Ning, S. Ng, R. Berman, D. W. Andrews, and Q. Fang, Opt. Lett. **37**, 250 (2012).
22. R. Mercade-Prieto, L. Rodriguez-Rivera, and X. D. Chen, Photochem. Photobiol. Sci. **16**, 1727 (2017).

Paper VI

Ultrafast Science
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

The Space-Charge Problem in Ultrafast Diagnostics: An All-Optical Solution for Streak Cameras

Vassily Kornienko, Yupan Bao, Joakim Bood, Andreas Ehn, and Elias Kristensson*

Department of Physics, Lund University, Lund, Sweden.

*Address correspondence to: elias.kristensson@forbrf.lth.se

Citation: Kornienko V, Bao Y, Bood J, Ehn A, Kristensson E. The Space-Charge Problem in Ultrafast Diagnostics: An All-Optical Solution for Streak Cameras. *Ultrafast Sci.* 2024;4:Article 0055. https://doi.org/10.34133/ultrafastscience.0055

Submitted 4 October 2023
Accepted 19 December 2023
Published 30 January 2024

The field of ultrafast science is dependent on either ultrashort laser pulse technology or ultrafast passive detection. While there exists a plethora of sub-picosecond laser pulse solutions, streak cameras are singular in providing sub-picosecond passive imaging capabilities. Therefore, their use in fields ranging from medicine to physics is prevalent. Streak cameras attain such temporal resolutions by converting signal photons to electrons. However, the Coulomb repulsion force spreads these electrons spatiotemporally aggravating streak cameras' temporal resolution and dynamic range—an effect that increases in severity in ultrafast applications where electrons are generated nearly instantaneously. While many electro-optical solutions have been proposed and successfully implemented, this issue remains as a challenge for all sub-picosecond streak camera technology. Instead of resorting to electro-optical solutions, in this work, we present an all-optical approach based on the combination of photon tagging and spatial lock-in detection with a technique called periodic shadowing—that is directly applicable to all generations of streak cameras. We have demonstrated that this accessible all-optical solution, consisting of a single externally applied optical component, results in (a) a >3× improvement in dynamic range, (b) a 25% increase in temporal resolution, and (c) a reduction of background noise levels by a factor of 50, which, when combined, allows for a markedly improved accuracy in the measurement of ultrafast signals.

Introduction

The aim to understand ultrafast transient physics and chemistry on the sub-picosecond scale has resulted in research ranging from plasma physics [1–3] to the field of attochemistry [4]. In order to measure such events, fast detection is a necessity, be it achieved either passively or actively [5]. Such detection technology ranges from high-speed cameras with videography speeds of kilohertz up to 100 MHz [6–8], intensified cameras with nanosecond shutter speeds [9] to ultrafast laser-based methods that can attain femtosecond or even attosecond shutter times [10,11] and beyond terahertz videography speeds [12–14].

A unique class of fast detectors that is extensively used in experimental studies around the world is the streak camera. Streak cameras were invented in 1949 by Courtenay-Pratt [15,16] and were first used in order to characterize ultrashort laser pulses generated by mode-locked laser systems [17]. Today, they are used in physics applications ranging from visualizing the propagation of light [18,19], observing self-modulation of relativistic photon bunches [20] or for visualizing the coherent control of molecular rotations [21]. Outside of physics, biomedical applications also rely strongly on streak camera-based fluorescence lifetime imaging for understanding, e.g., protein–protein interactions in living cells [22,23].

In short, streak cameras function by firstly converting photons into electrons to then be accelerated and incident onto a so-called streak tube (Fig. 1). Here, they experience a fast voltage ramp such that they obtain a deflection angle dependent on their time of arrival at the tube's entrance. This has the effect of converting their temporal signature into a spatial distribution to then be recorded on a focal point array of choice, such as a charge-coupled device (CCD) chip.

The conversion of photons to electrons, followed by the application of ultrafast electric field ramps, is key for the extreme temporal resolutions attainable with streak cameras. However, due to this complicated structure, there are a multitude of sources that deteriorate their temporal resolution. These range from varying trajectory times depending on the photocathode location where electrons are generated [24] to the initial energy spread of the photoelectrons affecting their speed throughout the system (chromatic dispersion) [25]. The major culprit is however the so-called space-charge effect [26] (illustrated in the lower panel of Fig. 1). Briefly, the Coulomb interaction between electrically charged particles will cause electrons that are produced in a point (in space and time) to expand into a spherical cloud as they travel in time. This radial expansion is then recorded by the streak camera as a broadening of the temporal response profile (see the time axes of Fig. 1). It turns out that this broadening is extremely nonlinear, posing experimental difficulties when quantifying e.g., fluorescence/coherence lifetimes or plasma emission. Indeed, due to this nonlinearity of the response, one cannot assume a constant

temporal resolution to be applicable throughout an entire detected event.

Many electro-optical-based approaches have been proposed and successfully implemented in order to alleviate these issues. One example includes inserting a reflectron into the system, forcing a reverse acceleration upon the electrons counteracting the space-charge temporal distortion at the streak tube [27]. A severe trade-off between temporal resolution, linear response, and dynamic range is nonetheless always present.

In this work, we address the space-charge induced temporal distortion issue in streak cameras with an accessible all-optical approach based on the periodic shadowing (PS) technique. By eliminating the signal from electrons that have acquired a spatial deflection due to the space-charge effect and being read out as a temporal broadening, we show an increase in temporal resolution by ~25% along with a simultaneous 3-fold increase in dynamic range. The all-optical nature of this technique does not encroach on the electro-optical assembly of the camera, opening up for use on any streak camera system with similar results. Finally, the technique is applied to fluorescence lifetime measurements of Rhodamine B, where we show substantial improvements in the robustness of extracted lifetimes, opening up for direct application of this technique across the sciences.

The streak camera

The working concept of a streak camera is displayed as the schematic of Fig. 1, where the model under investigation for this work is an Optoscope SCMU-ST main streak camera unit with an FTSU-1 sweep unit. Four temporally and spatially separated light pulses (in this case color-coded photon bunches) are incident on a slit oriented along the axis where spatial information is desired. This slit is imaged onto a photocathode (in our case, an S20 photocathode with peak spectral response around 450 nm) where electrons are emitted whose energy and number depend on the photon intensity and wavelength. These photoelectrons are accelerated toward a fine structure accelerating mesh with a, in our case, ~3.5-kV potential. The accelerated electrons travel into a streak tube, where a time-dependent voltage is applied in the direction perpendicular to the slit axis. Consequently, the electrons will obtain a deflection angle depending on their time of arrival into the streak tube, effectively projecting temporal information onto a spatial coordinate. After the deflection, the photoelectrons are incident on a phosphor screen that converts them to photons. These photons are then imaged via an image intensifier unit consisting of a photocathode (in our case, an LNS20 cathode, with gain 900 V for all the presented measurements), microchannel plate and another phosphor screen (all fiber coupled). Finally, this signal is imaged onto a CCD camera (Anima PX/25). The result is a 2-dimensional image of the signal where the 1-dimensional spatial information is recorded along the slit axis and the temporal information is recorded purpendicular to the slit axis.

Since the temporal information has been projected onto a spatial axis, one can define a so-called technical temporal resolution, τ_0, which is related to the spatial resolution of the system (pixel size, magnification, and slit transfer function) and the temporal characteristics given by the voltage ramp (the streak speed given in picoseconds per millimeter). In essence, τ_0 corresponds to the streak camera's temporal response in its linear regime, that is, when the signal intensities lie within the linear response region of the system. In our case, at a streak speed of 10 ps/mm, $\tau_0 = 2$ ps, and since the spatial resolution is constant, the technical temporal resolution decreases proportionally with streak speed, i.e., at a streak speed of 25 ps/mm, $\tau_0 = 5$ ps, etc. Another consequence of the projection of temporal information onto a spatial axis is the so-called temporal coverage or time range (TR) of the streak camera. Due to the finite size of the CCD camera upon which the signal is imaged, there is only a certain temporal window that can be imaged for a given

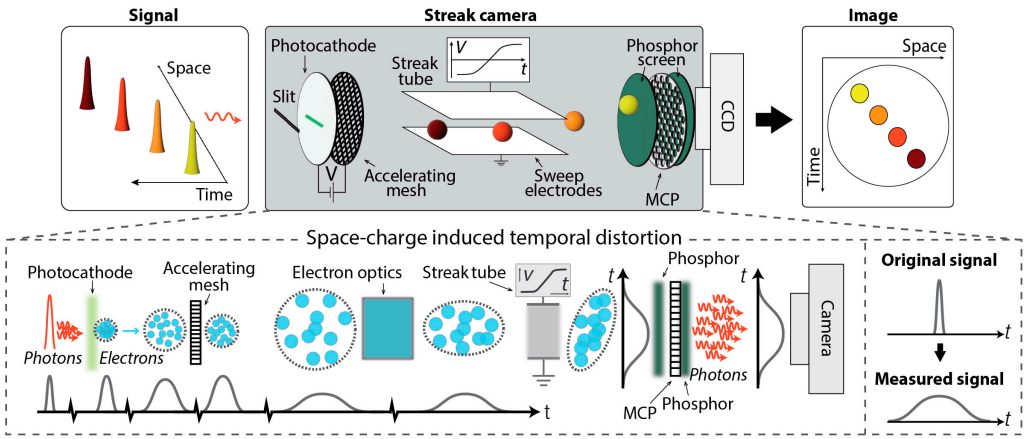


Fig. 1. Illustration of streak cameras and the space charge effect. Top panel: A signal consisting of 4 pulses separated in space and time are incident on the entrance slit of the streak camera. Photons are converted to electrons and then pass through a streak tube, where a fast voltage ramp is applied. This ramp will deflect electrons at an angle dependent on their time of arrival at the streak tube, effectively projecting the pulses' timestamp to a spatial axis. This schematic has been adapted from Hamamatsu's "Guide to Streak Cameras" [44]. Lower panel: The space charge effect, caused by the Coulomb repelling force between electrons, will cause the traveling electron bunches to expand radially, distorting the temporal footprint of the measured signal.

technical temporal resolution. With a, in our case, CCD size of 19.5 mm, a streak speed of 10 ps/mm will result in a TR of 195 ps. Any signal that lies before or after this window will not be recorded. For more details about the technical temporal resolution and TR, see Section S1, in particular Table S1 for conversion between streak rate, τ_0, and TR.

As a result of the time scales that streak cameras are designed to detect via the conversion of photons to electrons, many issues that affect their temporal response must be considered. Examples include chirp, optical/electron path lengths, fast electronics, varying photon kinetic energy upon excitation, and the speed of light itself. In most measurement situations, the inner workings of streak cameras are dependent on processes that are on the same time scale as the signals they are created to measure, hence an in depth understanding of their temporal response functions is necessary.

Temporal response

Characterizing response functions usually implies the measuring of a so-called delta pulse (the 2 leftmost pulses of $S(t)$ of Fig. 2A). The measured signal, M, is then a convolution of said temporal response, F, with the signal itself, i.e., $M = S \circledast F$. Hence, the measured signal contains artifacts that have, simply put, smoothed the original signal, resulting in a temporal width of the delta pulse that is equal to the temporal width of the response function. The resulting effect on the measurement of e.g., exponential decays is a shortened linear regime, where a robust exponential fit can be performed, i.e., a shortened so-called fitting window (Fig. 2A). Deconvolution algorithms can and have been successfully applied to extract the true underlying signals [28]; however, this process can still be sensitive to hardware choices and assumptions (as thoroughly discussed in [3]). Moreover, any signal features that occur on time scales faster than the response of the detector are irretrievable.

If the response function is nonlinear such that it might have a nonlinear dependence on the signal itself, e.g., its intensity, the ability to retrace the original signal from the measured one quickly becomes a daunting task (illustrated in Fig. 2B). In this case, the response full width at half maximum (FWHM), τ (in picoseconds), is described by:

$$\tau = \begin{cases} 10 & \text{if } I \leq 0.5 \\ 20I & \text{if } I > 0.5 \end{cases} \quad (1)$$

As illustrated, the more intense delta pulse will obtain a much larger FWHM than the weaker one, running the risk of obscuring the latter. Furthermore, since the exponential decay spans across all intensities, the fitting window will have to be decreased in size and positioned carefully such that the correct decay constant can be extracted.

It turns out that streak cameras follow a similar response to the nonlinear case in Fig. 2B and Eq. 1, i.e., their temporal response is nonlinear as a function of the measured signal intensity. This is in stark contrast to the majority of temporally resolved detection systems such as e.g., high-speed cameras where the shutter speed and exposure time are indeed independent of recorded signal strength. Fortunately, there does exist a range where the response is constant, in this case between intensities 0 and 0.5. Within this region, the streak camera operates equivalently to Fig. 2A, hence it is dubbed the linear regime of the system, or the technical regime (due to its relation to τ_0). Unfortunately, the linear regime of streak cameras is usually quite limited, constraining its applicability in various measurements. Furthermore, in the presence of noise and other interfering sources, this limitation of linearity poses serious issues for the determination of quantities of interest; indeed, the application of e.g., a fitting window becomes a non-trivial problem.

Space-charge effect

The space charge effect originates from the Coulomb repelling force between electrons and has the effect of entwining the ideally independent, temporal resolution of streak cameras with the signal itself. A schematic representation is available in Fig. 3A, albeit in an exaggerated manner for illustrative purposes. A photon pulse of full width at half maximum zero ($FWHM_0$) in time is incident onto a photocathode, creating a cloud of electrons localized in space and time. As they travel, the Coulomb interaction causes this cloud to expand rapidly attaining a temporal footprint of $FWHM_2$ ($>FWHM_1 > FWHM_0$). When entering the streak tube, electrons are deflected as a

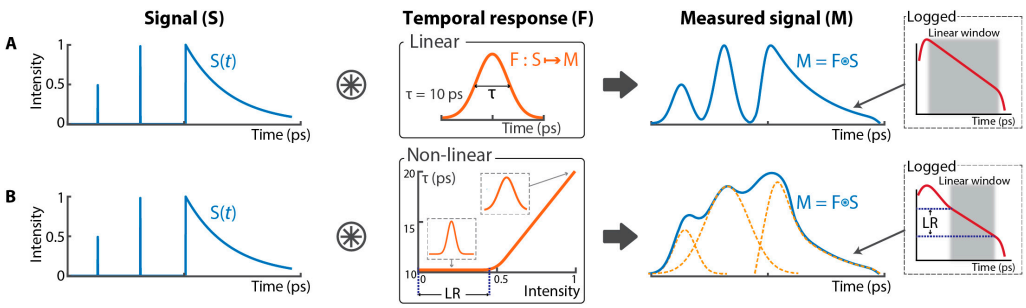


Fig. 2. Temporally resolved measurements. (A) A signal consisting of 2 delta pulses and an exponential decay are measured by a system with a constant temporal response of FWHM 10 ps. The convolution of the 2, $F \circledast S$, results in the measured signal, M, which has, simply put, been smoothed by the response function. In the case of the exponential decay, the result of this convolution is a shortened linear fitting window where the fit returns a constant value. (B) A nonlinear response function where, in this case, the FWHM of the response depends on the signal intensity, will result in different temporal widths of the 2 delta pulses. Furthermore, the linear regime of the exponential, where a fitting window is applicable, becomes considerably shorter.

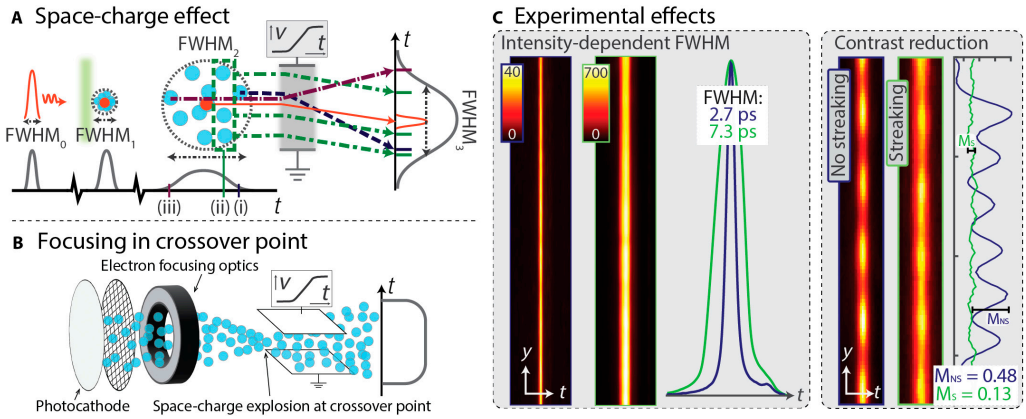


Fig. 3. Space-charge effect: in theory and practice. (A) The Coulomb repulsion between electrons that are created at the photocathode results in a spatial broadening of the electron bunch as they travel, dubbed the space-charge effect. This spatiotemporal expansion is measured as a temporal broadening of the initial pulse of light. (B) The focusing that is necessary to minimize spread at the streak tube (at the so-called crossover point) risks turning into a space-charge explosion, severely limiting the maximum signal intensity. (C) Experimental consequences of the space-charge effect. Left: An intensity-dependent temporal FWHM as a result of the effects described in (A). Right: A contrast reduction resulting from the focusing at the crossover point, a necessity when streaking is initiated.

function of their time of arrival; electrons in a temporal layer (a) are deflected downwards at a large angle, (b) are deflected downwards at a slightly smaller angle, and (c) are deflected upwards. However, their space-charge induced temporal distribution, which itself substantially broadens the signal, is not the only contributor to the response function. For electrons in temporal layer (b), their spatial distribution will also induce a measured broadening in time. Indeed, the 3 displayed electrons in layer (b) (within the green dashed box) will enter the microchannel plate at 3 different locations along the temporal axis, causing a further broadening of the measured FWHM such that $FWHM_3 > FWHM_2 > FWHM_1 > FWHM_0$. Hence, the measured FWHM of the photon pulse can be described as a convolution of the technical temporal transfer function, with the spatiotemporal space-charge induced transfer function. It turns out, as will be shown below, that the space-charge effect is what induces nonlinearities in the transfer functions of streak cameras similar to that of Fig. 2B and Eq. 1.

The space-charge effect, having been studied extensively due to its importance in a multitude of applications, is characterized by a power-law dependence ($\gamma = 0.5$) between electron bunch density and spatial size after a given time of flight [29]. This effect is illustrated experimentally in the left panel of Fig. 3C, where 2 laser pulses of equal temporal width (45 fs) but different intensities are imaged under identical conditions (streak rate and magnification). The temporal width (in pixels) of the recorded signal (left panel of Fig. 3C) has increased by a factor of ~3 for an increase in recorded intensity by a factor of ~15, clearly illustrating the strong dependence of temporal resolution on signal intensity.

An example of an industry standard solution to this issue is by focusing the electrons at the entry of the streak tube, at the so-called crossover point, thus minimizing their spatial spread [26]. However, the strong r^2 dependence of the Coulomb interaction results in a space-charge explosion at this point, seriously deteriorating the signal if one is not careful (Fig. 3B). Experimentally, this effect is shown in the right panel of Fig. 3C where the difference between the modulation depth of a sinusoidal intensity structure is severely affected when going from an unfocused electron beam (no streaking) to a focused beam (with streaking).

As has previously been mentioned, there does indeed exist a range of signal intensities where existing electro-optical solutions are adequate for resolving the issues caused by the space-charge effect, i.e., a range where the response is constant as a function of photon intensity. This is the region that is dubbed the linear regime of the streak camera; however, a severe trade-off between temporal resolution and the range of this regime is always present. A possible solution lies in attempting to isolate electrons that travel along the desired imaging path, or paraxially, throughout the streak camera system (illustrated as the red electron in Fig. 3A). The measured FWHM of the signal originating from these electrons would unavoidably be less affected by the spatial dimension of the space-charge effect, resulting in a more accurate rendition of the initial pulse. In this article, by imaging a periodic structure located at the slit throughout the system and realizing that only electrons that satisfy the paraxial imaging conditions will maintain said periodic structure, we are able to mitigate space-charge effects and improve the temporal characteristics of the system. The signal from electrons that exhibit these imaging properties is captured and isolated using the so-called PS technique.

Method

PS is a technique originally developed in order to increase the contrast of spectrometers by reducing stray light contributions originating from scattering within the instrument [30]. It has been further applied in order to, for example, create high-speed spectrograms of spatially separated spectra [31], achieve high-dynamic-range spectroscopy [32], and for improving the temporal contrast of streak cameras [33]. The technique involves

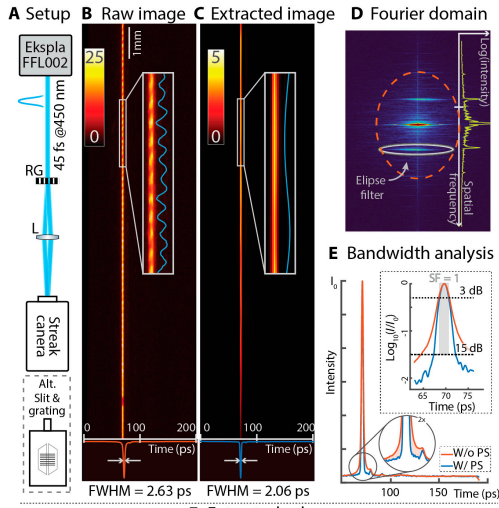
applying a sinusoidal spatial intensity modulation, of Michelson contrast one and frequency ω, to the signal of interest. The total, now modulated signal, is then imaged throughout the given measuring system and recorded. As the signal travels through the detection system, it can acquire an unwanted background signal. In the case of streak cameras, this could originate from the space-charge effect. Upon detection, the modulated signal will show up as discrete peaks in the spatial Fourier domain, while the low-frequency background will reside, to a large extent, in the DC component (this is usually dubbed $1/f$ noise). Consequently, this strategy effectively separates, in reciprocal space, the signal from the background acquired during detection. In order to extract said signal, lock-in detection is applied. In short, the signal peaks are shifted, in a phase-conserving manner, into the central DC component of the Fourier domain while simultaneously shifting the unwanted background away from the centre. By then low-pass filtering, the signal can hence be extracted. Note that, since the modulated signal was shifted before low-pass filtering, the applied modulation has now been eliminated from the extracted image and only the modulation envelope remains, which itself is a low-pass-filtered representation of the original signal. In general, this strategy has the effect of diminishing the contribution of signal that has not traveled along the detector's desired path and maintained the applied modulation. Thus, PS provides a practical solution to isolating signal from electrons that travel paraxially through the streak camera system such as the red ball in Fig. 3A. For a more detailed mathematical description of the algorithm, refer to Section S2.

Results

Pulse characterization for the determination of temporal resolution

Probing the temporal response of an ultrafast system in general requires an analysis involving the response of the system to a delta spike in time. For this streak camera system, with technical temporal resolution, τ_0, of 2 ps, a 45-fs, 450-nm laser pulse (Ekspla FFL002) is chosen as said delta probe. The PS concept, i.e., the technique used in order to isolate the signal from paraxially traveling electrons, is experimentally applied to this measurement by imaging a vertically rotated Ronchi grating (10 lp/mm) onto the entrance slit of the streak camera (Fig. 4A). A single thin lens is used for imaging rather than an objective, minimizing chirp and wavelength-dependent time lags [34]. Alternatively, a Ronchi grating can be directly mounted on the slit entrance (inset of Fig. 4A). The pattern and laser pulse are then imaged through the system at a streak rate of 10 ps/mm (which corresponds to 140 fs/px and $\tau_0 = 2$ ps).

The resulting image, Fig. 4B (jitter corrected average of 100 individual shots) with its sinusoidal structure in the y-direction, indicates a temporal response of 2.63 ps, extracted as the FWHM of the vertically collapsed signal. The temporal characteristics of this image are given by the convolution of the illuminating pulse (45 fs), the slit width (which can reach a minimum of ~8 px × 140 fs/px = 1.12 ps), the jitter correction algorithm (~2 px = 280 fs), and the electro-optic temporal response of the system (unknown but includes all effects related to space-charge). Hence, since the illuminating pulse contributes negligibly to this convolution, Fig. 4B can essentially be interpreted as the raw minimal temporal response of the streak camera system in question.



F Extracted values

	Without PS	With PS	Improvement
3 dB BW	392 GHz	486 GHz	24%
SF(3/15 dB)	5.6	2.3	143%

Fig. 4. Improved temporal resolution with PS. (A) Short femtosecond illumination (Ekspla FFL002) is directed toward the entrance slit of the streak camera in order to probe its temporal response. PS is incorporated by imaging a Ronchi grating (RG) directly onto the slit in a 2f configuration. (B) The raw temporal response of the system (2.63 ps) at $\tau_0 = 2$ ps, with the characteristic vertical modulation of the PS technique. (C) The PS extracted temporal response, (performed via spatial lock-in on [A]) results in a 27% improvement of FWHM (to 2.06 ps). (D) The spatial frequency domain of the raw signal depicting the modulated signal as individual peaks while simultaneously exposing the intrinsic low-pass filtering characteristics of the streak camera system (orange dashed circle). (E) The integrated temporal responses before and after extraction (w/o and w/ PS, respectively) depicting marked improvements in 3-dB bandwidth (1/FWHM) and shape factor (SF). Note how in the logarithmic inset, a high bandwidth is maintained even at high attenuations when using PS. (F) Table indicating the improvement of 3-dB bandwidth (BW) and SF upon application of PS.

A spatial lock-in at the modulation frequency of the imaged sinusoidal pattern, revealed as the high-frequency peaks in the Fourier domain of the raw image (Fig. 4D), is then performed. The indicated eliptical low-pass filter isolates said modulated signal such that only the information within the corresponding Fourier component is extracted. This is the step where the signal from electrons that satisfy the imaging conditions of the system are preserved while signal from unwanted electrons is diminished. The orientation and length of the filter's semimajor axis is chosen such that it coincides with the system's intrinsic low-pass response (marked as the orange dashed circle). Indeed, a semimajor axis larger than this would only contribute with noise and not a higher temporal resolution.

The resulting PS-extracted image, along with its integrated temporal response curve (Fig. 4C) displays a substantial improvement down to 2.06 ps, only slightly above τ_0. A more detailed bandwidth analysis is performed in Fig. 4E where the vertically integrated signals are plotted both linearly and logarithmically. A summary of the extracted values is also presented in the table of Fig. 4F. An improvement of the 3-dB response bandwidth

(1/FWHM) by 24% illustrates the efficiency of PS for filtering out temporally broadening artefacts hidden within the DC component of the signal. Upon further analysis, the improved steepness of the extracted response is evident in the logarithmic inset of Fig. 4E, where high bandwidth is maintained even at high attenuation (70 versus 210 GHz at 15 dB pre- and postextraction, respectively). This steepness can be quantified with the 3/15-dB shape factor (SF), i.e., the ratio between 3- and 15-dB bandwidth of the given signal. A shape factor of 1, as indicated by the light gray box, is considered the ideal case or infinitely steep. After extraction, an improvement of this quantity by 143% is realized, opening up for the detection of features with largely differing intensities at high bandwidth, showing improved dynamic range when PS is applied. This general amelioration of temporal characteristics persists over a large range of technical temporal resolutions (see Section S4). Indeed, as long as the space charge effect interferes with the signal, the amelioration offered by PS will endure.

Dynamic range analysis

As previously elaborated upon, the temporal response of a streak camera system is, due to the space-charge effect, directly affected by the signal intensity. Due to this dependence, there is a direct relation between the temporal resolution of the system and its dynamic range. Indeed, if a low temporal resolution is adequate for a given measurement situation, one can tolerate a larger space-charge-induced electron cloud to travel through the system, which in turn allows for higher signal intensities or dynamic ranges.

The dynamic range (DR) of a system for a given temporal resolution (τ) can be defined as:

$$DR_\tau = \frac{I_\tau^{\max}}{I_\tau^{\min}} = \frac{I_\tau^{\max}}{\epsilon_n}, \quad (2)$$

where $I_\tau^{\max/\min}$ is the maximum/minimum possible signal intensity for a given temporal resolution and ϵ_n is the background noise level of the system. Since the deterioration of temporal resolution is monotonic with respect to intensity, $I_\tau^{\min}$ is set as the noise floor, ϵ_n. Indeed, as the space-charge effect becomes less important for lower signal intensity, the highest detectable temporal resolution of a streak camera system as a whole is at a signal of signal-to-noise ratio, SNR, equal to 1. An equivalent argument can be used to understand how to probe $I^{\max}$. By probing the temporal response of the system with increasing signal intensity, the monotonic behavior of the space-charge effect will ensure that the measured temporal response for a given intensity will indeed be the maximized temporal resolution for the given signal intensity.

In order to investigate the effect of PS on the dynamic range, a lossy cavity consisting of one mirror and a 50/50 beamsplitter (of thickness 1 mm) was added to the beam path, creating a train of pulses whose intensities decrease by a factor of 1/2 but maintain their ~45 fs temporal width (Fig. 5A). The size of the cavity is set such that as many individual pulses as possible are visible within the temporal window of the streak while simultaneously avoiding any overlap. This is no new concept for characterizing the dynamic range as a function of τ within streak camera systems owing to the experiment's relatively simple yet informative footprint [35,36]. As previously mentioned, since the temporal resolution monotonically deteriorates with signal intensity, the detected FWHM of the pulses can be regarded as the temporal resolution limit, τ, for that intensity. Hence by recording pulse trains with ranging intensities and extracting the individual temporal profiles, the τ-dependent dynamic range can be extracted from a single streak.

The jitter-corrected, 100-shot average of such a pulse train, imaged at a streak rate of 25 ps/mm ($\tau_0 = 5$ ps), is shown in Fig. 5B, where the color axis has been adjusted in order to highlight

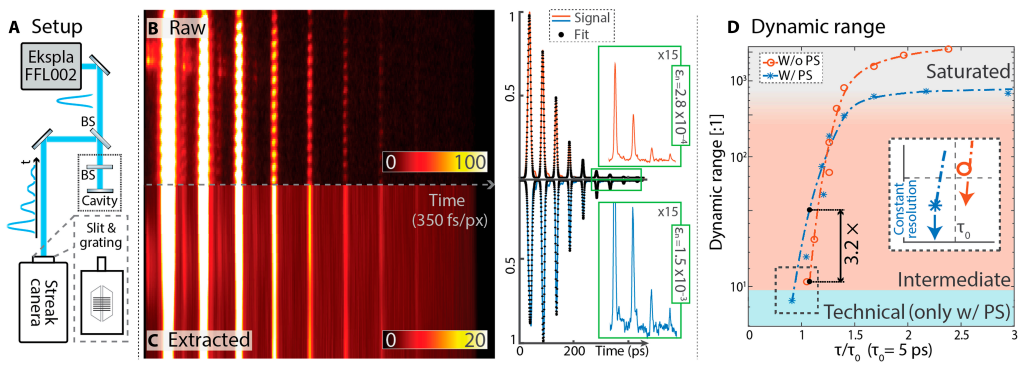


Fig. 5. Enhanced dynamic range with PS. (A) An optical cavity was added to the setup, creating a femtosecond pulse train of monotonically decreasing pulse intensities. (B) The raw, 100 times averaged and jitter corrected image along with its vertically integrated counterpart (to the right) imaged at $\tau_0 = 5$ ps (normalized to the maximum intensity). (C) The PS extracted pulse train and its vertically integrated counterpart. Note that the noise floor, ϵ_n, of the PS extracted signal is 7 times thicker than for the raw signal. (D) For a given maximum signal intensity the space-charge effect only permits a certain temporal resolution, illustrated as the red curve which corresponds to the fitted intensities and FWHM of (B). The dynamic range as a function of temporal resolution, where the PS extracted signal (blue curve) is shifted to the left, illustrates how PS effectively addresses the space-charge effect and allows for higher temporal resolutions at higher signal intensities. Furthermore, in contrast to w/o PS, the technique allows the system to reach τ_0 for nonzero dynamic range, opening up for a measurements within the region of linear temporal response.

the low-intensity pulses. One can immediately note how, even though each incoming pulse has a constant pulse duration, the detected FWHM decreases along the pulse train, clearly illustrating intensity-dependent temporal broadening. The lock-in and low-pass filtering algorithm used for PS (identical to Fig. 4D) is applied to this raw image resulting in Fig. 5C, along with its normalized integrated version to the right.

A gaussian pulse train is fitted to this raw integrated signal, resulting in intensity and FWHM values for each individual pulse. This, along with the noise floors (extracted as the standard deviation of a blank recording) are then inserted into Eq. 2 and plotted in Fig. 5D (see Section S3 for details). In general, a consistently left-shifted PS extracted curve at low τ/τ_0 indicates that PS simultaneously increases the dynamic range and temporal resolutions of the system over a broad range of measurement situations.

The calculation of the dynamic range, as in Eq. 2, is essentially a calculation of a temporal resolution-dependent SNR. For the last peak of the w/o PS curve (marked by the lower black dot at $\tau/\tau_0 = 1.06$ in Fig. 5D), the dynamic range is calculated as:

$$DR^{\mathrm{w/oPS}}(\tau) = \frac{I^{\mathrm{w/oPS}}_{5.3ps}}{\epsilon^{\mathrm{w/oPS}}} = \frac{3.0\times10^{-3}}{2.8\times10^{-4}} = 10.71 \qquad (3)$$

Since the maximum temporal resolution follows a monotonic behavior with intensity, a fitted value for the intensity of the w/ PS case at a temporal resolution of $\tau = 5.3$ ps, $I^{\mathrm{w/PS}}_{5.3ps}$, is used for the calculation of the dynamic range:

$$DR^{\mathrm{w/PS}}(\tau) = \frac{I^{\mathrm{w/PS}}_{5.3ps}}{\epsilon^{\mathrm{wPS}}} = \frac{5.0\times10^{-2}}{1.5\times10^{-3}} = 33.33 \qquad (4)$$

Hence, as marked in Fig. 5D, an amelioration of dynamic range by a factor of 3.1 (33.11/10.71) at a temporal resolution of 5.3 ps ($\tau/\tau_0 = 1.06$) is achieved when using PS. Even though the increase of the noise floor by a factor of 5 when using PS directly affects the dynamic range, this amelioration is nonetheless possible thanks to the shift of temporal resolution which outweighs the increase in noise floor (visualized by the left shifted blue curve in Fig. 5D).

We can divide the behavior of the dynamic range into 3 different regimes; saturated, intermediate, and technical. The low gradient of the saturated regime illustrates a dynamic range ceiling where the system cannot in general handle such high signal intensities. Here, the imaging conditions suffer from strong, unmanageable nonlinearities owing to a space-charge explosion. Practically the temporal resolution of a measurement in this regime cannot be inferred. This is analogous, however not identical, to how variations in signal intensity cannot be detected if a complementary metal-oxide semiconductor or CCD camera is saturated. In streak cameras, temporal variations of a signal cannot be detected if their intensities reach the saturated regime.

In contrast, the intermediate regime displays a steep dependence of dynamic range on the desired temporal resolution, where the PS extracted curve displays an improvement over the majority of the region. The significance of this improvement can be illustrated by an example. If one needs to perform a measurement where the physical process under investigation has a characteristic time scale of $\tau = 5.5$ ps ($1.1\tau_0$), the signal cannot vary in intensity by more than a factor of 14.8 when not using PS. However by applying PS and isolating the signal from paraxially traveling electrons, the signal can instead vary in intensity over a factor of 39.8, corresponding to a nearly 3-fold improvement in dynamic range.

There is indeed a point, around $\sim 1.3\tau_0$, where the enhancement of dynamic range for PS turns into minor deterioration. This stems from the space-charge effect deteriorating the signal to the point where the sinusoidal pattern's contrast decreases nonlinearly, resulting in lower amplitudes after PS extraction (see Section S5 for more information). To counteract this behavior, one may lower the streak rate in order to better match τ_0 with the characteristic temporal evolution of the event, shifting the measurement into a more desirable linear regime.

The third, technical regime, is only applicable to the extracted PS signal. Here, any signal that varies with a smaller dynamic range than this is measured at the technical temporal resolution i.e., at a constant temporal response. The importance of reaching this regime during measurements is illustrated in Fig. 2, where the temporal response within the technical regime corresponds to a constant response function. In stark contrast to the nonlinear response, in the technical regime deconvolution algorithms can be applied, and the measured temporal features in the signal are not dependent on the signal's intensity. This is an extremely important characteristic when determining sensitive quantities such as fluorescence lifetimes, where the fitting window can play an important role in its quantification.

Robust determination of fluorescence lifetimes

Fluorescence lifetime imaging is a key technique within the life sciences used, for example, in order to probe the chemistry of living cells [23]. The general principle involves the excitation of a molecule with a short laser pulse and following the fluorescence decay through the detection of the emitted photons. Performing an exponential fit to the detected fluorescence (mono- or multi-exponential depending on the system) reveals the lifetime, τ_L, of the molecule in question. This lifetime is heavily and robustly dependent on the molecule's local environment in terms of electric field, pressure, temperature, surrounding molecules, etc, and can thus be used as a strong indicator of, among other things, cancer [37] or structures within DNA sequences [38].

Extracting accurate τ_L from detected signals is a long-standing problem, where the troubles become more pronounced for lifetimes on the order of nanoseconds, in itself a regime of high importance within biomedical sciences. On the hardware side, low photon flux requires complicated systems such as Time Correlated Single Photon Counting, which, due to photon statistics, introduces a bias toward shorter lifetimes. Highly accurate timing requirements in the form of correct starting and stopping clocks strain electronics to their limits where jitter, nonuniform exposure windows in time (rather than tophats) and spatial nonuniformity of shutters introduce uncertainties in the extracted lifetimes [39,40]. All these hardware-based issues pose problems on the computational side where exponential fitting algorithms are prone to errors depending on the noise characteristics, temporal response functions, and the fitting window size [41].

Streak cameras present a good alternative as a detector for fluorescence lifetime signals as they can function in single photon counting mode and are specifically designed for fast signals, hence are engineered to have highly accurate clocking

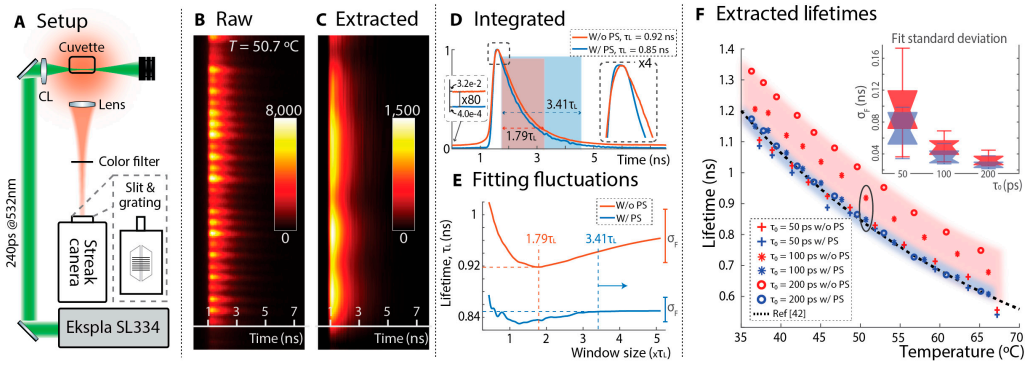


Fig. 6. Robustness of fluorescence lifetime imaging with streak camera/PS-based detection. (A) A 240-ps, 532-nm laser source excites a cuvette of Rhodamine B diluted in water (3×10^{-9} M). The detection is performed with a streak camera and PS-based system. (B and C) The raw and extracted image at a temperature of 50.7°C at $\tau_0 = 100$ ps. (D) The integrated lifetime curves, where the zoom-ins illustrates the effects of PS around the peak and a reduction of background level by a factor of ~80. (E) An exponential fit is then performed with close to identical start times and a varying window size. Without PS, the fitted lifetime never stabilizes while with PS, the fitted lifetime converges such that a robust determination of lifetime can be attained for large fitting windows. (F) The extracted lifetimes in comparison to literature [42] where PS performs markedly better for varying lifetimes and streak speeds. Furthermore, a consistently lower standard deviation of the fitted lifetimes, σ_F, is attained with PS for all streak rates in question.

electronics. However, the limited dynamic range and the nonlinear temporal resolution severely limits their use within the field, since, if one is not careful, multiple lifetimes could be extracted depending on the choice of fitting window (exemplified in Fig. 2B). In this article, we have presented a solution to these problems in the form of PS and Fig. 6 depicts its application to the estimation of Rhodamine B's temperature-dependent fluorescence lifetime.

In this experiment, a 240-ps, 532-nm pulse is directed and focused (as a sheet) into a cuvette of 3×10^{-9} molar Rhodamine B diluted in water. The fluorescence is then captured and imaged via a spherical lens and fluorescence filter onto the streak camera slit. Here, a glass Ronchi grating (5 lp/mm) is mounted, resulting in lifetime images with the characteristic PS stripes (Fig. 6A and B). Lock-in extraction is then performed resulting in Fig. 6C. The raw and extracted signals are then vertically integrated, normalized, and presented in Fig. 6D. Here, the shapes of the fluorescence decay with and without PS clearly differ, where the characteristic smoothing of the leading edge and the wider peak is visible for the w/o PS curve (red). Critically, albeit more subtle, an automatic background subtraction emblematic to the reduction of the DC component in the lock-in post-processing is visible by the downward shift of the blue curve by a factor of 80 (left zoom-in of Fig. 6D), allowing for robust lifetime extraction without complicated background estimation procedures.

A nonlinear least squares exponential fit is performed on these curves for varying fitting window sizes (with close to equal starting points) and the extracted lifetime τ_L is plotted as a function of said window size (Fig. 6E). The characteristic nonconverging curve for the lifetime as a function of fitting window size is visible for the case without PS, and the lifetime is chosen as the minimum value at a window size of $1.79\tau_L$, $\tau_L = 0.92$ ns. In contrast, the fitted lifetime on the PS signal does indeed converge for window sizes above $\sim 3\tau_L$ illustrating how, with PS, the fitting becomes substantially less sensitive to fitting parameters. The lower lifetime of 0.85 ns furthermore illustrates an overestimation of the lifetime for the case w/o PS. This is further corroborated in Fig. 6F, where irrespective of streak rate (depicted as the technical temporal resolution) and lifetime (varied with temperature) streak camera-based fluorescence lifetime determination with PS is markedly more robust and yields values in cohesion with literature [42]. Moreover, even though the standard deviations of extracted lifetimes with respect to window size, σ_F, increase with increased temporal resolution (attributed to less signal per pixel, see Section S6), they are nonetheless consistently lower for the PS signals, illustrating a higher level of precision.

Discussion

The lock-in procedure, characteristic to PS, only extracts signal that has maintained the applied sinusoidal intensity structure throughout the system, diminishing the signal from electrons that have not traveled along the desired optical path within the streak camera. In our case, this unwanted contribution constitutes of about 80% of the original signal (the extracted image intensities are about 5 times lower), illustrating the prominence of the space-charge effect within these streak camera systems. Fortunately, due to the intrinsic properties of lock-in detection, the DC background component is reduced by a factor of >50. This efficiently counteracts the lower extracted signal levels resulting in higher dynamic ranges nonetheless.

The low-pass filtering applied in the lock-in algorithm negatively affects the spatial resolution (in the vertical direction) of the final image. A higher grating frequency can be used to address this issue allowing for a wider band-pass filter and a more efficient filtering of unwanted signal [43]. However, the blurring caused by the space-charge effect reduces the contrast of the imaged modulation nonlinearly; hence, at high signal intensities, a trade-off between spatial resolution and filtering capabilities must be considered. Fortunately, since the application of PS is independent of the signal of interest, optical manipulation of said signal in the form of, e.g., magnification can be applied to address spatial resolution limitations.

There are multiple technical advantages of using PS in conjunction with electro-optical-based systems such as streak cameras. Firstly, since the technique increases the dynamic range of the streak camera, less signal intensity needs to be attenuated before the slit, opening up for single-shot measurements where previously unfeasible. This, in turn, relieves the need for complicated and unreliable jitter correction that is necessary upon signal averaging. Secondly, a larger dynamic range in conjunction with a higher temporal resolution allows for slower streak rates when measuring dynamic scenes. This opens up for larger temporal coverages within the camera window, or TRs, generating more possibilities for following an event over longer periods of time without losing temporal resolution within each instantaneous moment. Finally, due to its externally applied all-optical nature, one does not need to infringe on the fine electro-optical balances within the instrument itself. This allows for the concept of PS to be easily applied to other electro-optical-based imaging systems with similar results. Indeed, as long as imaging is involved, there is by default a desired optical path upon which one wants signal to travel, the only prerequisite of PS. Furthermore, this externally mounted optical approach opens up for the use of PS as a complement to other algorithms such as deconvolution algorithms. This could have the potential to further increase the robustness of, e.g., lifetime determination.

Here,we have presented an accessible, externally mounted all-optical approach based on the PS technique, that, when applied to streak cameras, is able to simultaneously achieve a 25% increase in temporal resolution, a >3× increase in dynamic range and a reduction of background noise levels by a factor of >50. This is accomplished by tagging and then algorithmically detecting photo-electrons that adhere to the imaging conditions of the streak camera system, hence substantially diminishing the contribution of signal that has been affected in undesirable ways, such as by the space-charge effect.

All in all, the every growing interest in ultrafast science, which itself necessitates ultrafast passive detection, has resulted in increased use of streak cameras in laboratories around the world. We believe that the all-optical approach presented herein will aid in extending streak camera applicability into the ultrafast as well as provide a new route for continued development of ultrafast electro-optical systems in general, and the streak camera in particular.

Acknowledgments

We would like to thank S. Nilsson for fruitful discussions about exponential lifetime fitting methods and applications and P. Summ from Optronis for clear explanations about the inner workings of streak cameras.
Funding: This work was funded by European Research Council (803634 and 852394) and Vetenskapsrådet (2019-05183,2021-04506).
Author contributions: V.K. and Y.B. performed the measurements and V.K. performed the data analysis. V.K. prepared the manuscript with input from all authors.
Competing interests: The authors declare that there is no conflict of interest regarding the publication of this article.

Data Availability

Data is available upon reasonable request.

Supplementary Materials

Sections S1 to S6
Figs. S1 to S8
Table S1

References

1. Höft H, Becker M, Kolb J, Huiskamp T. Double-propagation mode in short-gap spark. *Plasma Sources Sci Technol*. 2020;29(8):Article 085002.
2. Huiskamp T, Ton C, Azizi M, van Oorschot J, Höft H. Effective streamer discharge control by tailored nanosecond-pulsed high-voltage waveforms. *J Phys D Appl Phys*. 2021;55(2): Article 024001.
3. Invernizzi L, Duluard CY, Höft H, Hassouni K, Lombardi G, Gazeli K, Prasanna S. Peculiarities of measuring fluorescence decay times by a streak camera for ps-TALIF experiments in reactive plasmas. *Meas Sci Technol*. 2023;34(9):Article 095203.
4. Nandi S, Plésiat E, Zhong S, Palacios A, Busto D, Isinger M, Neoričić L, Arnold CL, Squibb RJ, Feifel R, et al. Attosecond timing of electron emission from a molecular shape resonance. *Sci Adv*. 2020;6(31):Article eaba7762.
5. Liang J, Wang LV. Single-shot ultrafast optical imaging. *Optica*. 2018;5(9):1113–1127.
6. Manin J, Skeen SA, Pickett LM. Performance comparison of state-of-the-art high-speed video cameras for scientific applications. *Opt Eng*. 2018;57:124105.
7. Suzuki M, Sugama Y, Kuroda R, Sugawa S. Over 100 million frames per second 368 frames global shutter burst CMOS image sensor with pixel-wise trench capacitor memory array. *Sensors*. 2020;20(4):1086.
8. Lazovsky L, Cismas D, Allan G, Given D. CCD sensor and camera for 100 Mfps burst frame rate image capture. Paper presented at: Airborne intelligence, surveillance, reconnaissance (ISR). Systems and Applications II; 2005 May 10; Orlando, Florida.
9. Nomerotski A, Chekhlov M, Dolzhenko D, Glazenborg R, Farella B, Keach M, Mahon R, Orlov D, Svihra P. Intensified Tpx3Cam, a fast data-driven optical camera with nanosecond timing resolution for single photon detection in quantum applications. *J Instrum*. 2023;18:C01023.
10. Linne MA, Paciaroni M, Berrocal E, Sedarsky D. Ballistic imaging of liquid breakup processes in dense sprays. *Proc Combust Inst*. 2009;32(2):2147–2161.
11. Peng P, Marceau C, Villeneuve D. Attosecond imaging of molecules using high harmonic spectroscopy. *Nat Rev Phys*. 2019;1(2):144–155.
12. Ehn A, Bood J, Li Z, Berrocal E, Aldén M, Kristensson E. FRAME: Femtosecond videography for atomic and molecular dynamics. *Light Sci Appl*. 2017;6(9):e17045.
13. Wang L, Wang P, Liang J. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nat Commun*. 2020;11(1):2091.
14. Nakagawa K, Iwasaki A, Oishi Y, Horisaki R, Tsukamoto A, Nakamura A, Hirosawa K, Liao H, Ushida T, Goda K, et al. Sequentially timed all-optical mapping photography (STAMP). *Nat Photonics*. 2014;8:695–700.
15. Courtney-Pratt JS. A new method for the photographic study of fast transient phenomena. *Research*. 1949;2(6):287–294.
16. Courtney-Pratt JS, A new photographic method for studying fast transient phenomena, Proceedings of the Royal Society of

London Series A-Mathematical and Physical Sciences 204, 27 (1950).
17. Bradley D, New GH. Ultrashort pulse measurements. *Proc IEEE*. 1974;62(3):313–345.
18. Velten A, Wu D, Jarabo A, Masia B, Barsi C, Joshi C, Lawson E, Bawendi M, Gutierrez D, Raskar R, et al. Femto-photography: Capturing and visualizing the propagation of light. *ACM Trans Graph*. 2013;32(4):44.
19. Gao L, Liang J, Lin L, Wang L. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature*. 2014;516(7529):74–77.
20. Adli E, Ahuja A, Apsimon O, Apsimon R, Bachmann AM, Barrientos D, Barros MM, Batkiewicz J, Batsch F, Bauche J, et al. Experimental observation of proton bunch modulation in a plasma at varying plasma densities. *Phys Rev Lett*. 2019;122(5–8):Article 054802.
21. Hosseinnia A, Raveesh M, Dominguez A, Ruchkina M, Linne M, Bood J. Single-shot coherent control of molecular rotation by fs/ns rotational coherent anti-Stokes Raman spectroscopy. *Opt Express*. 2022;30(18):32204–32214.
22. Krishnan RV, Masuda A, Centonze V, Herman B. Quantitative imaging of protein-protein interactions by multiphoton fluorescence lifetime imaging microscopy using a streak camera. *J Biomed Opt*. 2003;8(3):362–367.
23. Suhling K, Hirvonen LM, Levitt JA, Chung PH, Tregidgo C, le Marois A, Rusakov DA, Zheng K, Ameer-Beg S, Poland S, et al. Fluorescence lifetime imaging (FLIM): Basic concepts and some recent developments. *Med Photonics*. 2015;27:3–40.
24. Kalantar DH, Bell PM, Costa RL, Hammel BA, Landen OL, Orzechowski TJ, Hares JD, Dymoke-Bradshaw AK, Characterization of x-ray streak cameras for use on Nova In *22nd International Congress on High-Speed Photography and Photonics* , D. L. Paisley, A. M. Frank, eds., International Society for Optics and Photonics (SPIE, 1997), vol. 2869, p. 680–685.
25. Campillo A, Shapiro S. Picosecond streak camera fluorometry - A review. *IEEE J Quantum Electron*. 1983;19(4):585–603.
26. Shchelev M, Andreev SV, Greenfield DE, Degtyareva VP, Kopaev IA, Monastyrskiy MA. On some limitations on temporal resolution in imaging subpicosecond photoelectronics. *Quantum Elec*. 2015;45(5):455–461.
27. Wang Y, Gedik N. Electron pulse compression with a practical reflectron design for ultrafast electron diffraction. *IEEE J Sel Top Quantum Electron*. 2011;18(1):140–147.
28. Luo T, Fan R, Chen Z, Wang X, Chen D. Deblurring streak image of streak tube imaging lidar using Wiener deconvolution filter. *Opt Express*. 2019;27(26):37541–37551.
29. Tao Z, Zhang H, Duxbury PM, Berz M, Ruan C-Y. Space charge effects in ultrafast electron diffraction and imaging. *J Appl Phys*. 2012;111(4):Article 044316.
30. Kristensson E, Bood J, Alden M, Nordström E, Zhu J, Huldt S, Bengtsson PE, Nilsson H, Berrocal E, Ehn A. Stray light suppression in spectroscopy using periodic shadowing. *Opt Express*. 2014;22(7):7711–7721.
31. Richardson D. Emission spectroscopy with coded apertures for enhanced. *J Appl Phys*. 2021;130:Article 093105.
32. Kristensson E, Ehn A, Berrocal E. High dynamic spectroscopy using a digital micromirror device and periodic shadowing. *Opt Express*. 2017;25(1):212–222.
33. Bao Y, Kornienko V, Lange D, Kiefer W, Eschrich T, Jäger M, Bood J, Kristensson E, Ehn A. Improved temporal contrast of streak camera measurements with periodic shadowing. *Opt Lett*. 2021;46:5723–5726.
34. Ihlemann J, Helmbold A, Staerk H. Chromatic time lag in picosecond-streak-camera objectives. *Rev Sci Instrum*. 1988;59:2502–2503.
35. Friedman W, Jackal S, Seka W, Zimmermann J. Dynamic range and spatial resolution of picosecond streak cameras. Paper presented at: 12th Intl Congress on High Speed Photography; 1976 Aug 1–7; Toronto, Canada.
36. Hull DR, Freeman NJ. Dynamic range measurements on streak cameras with picosecond time resolution. *J Phys E Sci Instr*. 1980;13(6):685.
37. Karrobi K, Tank A, Fuzail MA, Kalidoss M, Tilbury K, Zaman M, Ferruzzi J, Roblyer D. Fluorescence Lifetime Imaging Microscopy (FLIM) reveals spatial-metabolic changes in 3D breast cancer spheroids. *Sci Rep*. 2023;13(1):3624.
38. Knemeyer J-P, Marmé N, Sauer M. Probes for detection of specific DNA sequences at the single-molecule level. *Anal Chem*. 2000;72(16):3717–3724.
39. Hirvonen L, Suhling K. Fast timing techniques in FLIM applications. *Front Phys*. 2020;8:Article 161.
40. Nilsson S, Kristensson E, Aldén M, Bood J, Ehn A. Fluorescence lifetime imaging through scattering media. *Sci Rep*. 2023;13(1):3066.
41. Fuhrmann N, Bruebach J, Dreizler A. On the mono-exponential fitting of phosphorescence decays. *Appl Phys B*. 2013;116(2):359–369.
42. Mercadé-Prieto R, Rivera L, Chen X. Fluorescence lifetime of Rhodamine B in aqueous solutions of polysaccharides and proteins as a function of viscosity and temperature. *Photochem Photobiol Sci*. 2017;16(11):1727–1734.
43. Kristensson E, Kristensson G. Physical explanation of the SLIPI technique by the large scatterer approximation of the RTE. *J Quant Spectrosc Radiat Transf*. 2017;189:112–125.
44. Hamamatsu, Guide to streak cameras (2008); https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/sys/SHSS0006E_STREAK.pdf

Paper VII

Research Article | Vol. 28, No. 7 / 30 March 2020 / Optics Express 9572
Optics EXPRESS

A versatile, low-cost, snapshot multidimensional imaging approach based on structured light

KAROLINA DOROZYNSKA,* VASSILY KORNIENKO, MARCUS ALDÉN, AND ELIAS KRISTENSSON

Department of Combustion Physics, Lund University, 223 63 Lund, Sweden
**karolina.dorozynska@forbrf.lth.se*

Abstract: The behaviour and function of dynamic samples can be investigated using optical imaging approaches with high temporal resolution and multidimensional acquisition. Snapshot techniques have been developed in order to meet these demands, however they are often designed to study a specific parameter, such as spectral properties, limiting their applicability. Here we present and demonstrate a frequency recognition algorithm for multiple exposures (FRAME) snapshot imaging approach, which can be reconfigured to capture polarization, temporal, depth-of-focus and spectral information by simply changing the filters used. FRAME is implemented by splitting the emitted light from a sample into four channels, filtering the light and then applying a unique spatial modulation encoding before recombining all the channels. The multiplexed information is collected in a single exposure using a single detector and extracted in post processing of the Fourier transform of the collected image, where each channel image is located in a distinct region of the Fourier domain. The approach allows for individual intensity control in each channel, has easily interchangeable filters and can be used in conjunction with, in principle, all 2D detectors, making it a low cost and versatile snapshot multidimensional imaging technique.

1. Introduction

Multidimensional optical imaging [1,2] is an approach which gathers multiple dimensions, up to nine [1], of information about a sample. Depth-of-focus, polarization, temporal and spectral information about samples are some of the dimensions of interest in optical imaging applications. Depth-of-focus information is required in volumetric imaging such as is obtained using scanning confocal microscopy [3]. Polarization imaging is useful for separating specular and diffuse reflections [4], 3-dimensional reconstruction and biological imaging [5] and studying transparent objects [6], i.e. stress, strain and imperfections. Temporal imaging is useful for studying highly dynamic samples such as combustion processes [7,8] and molecular dynamics. Spectral imaging has applications in combustion diagnostics [9], fluorescence imaging [10–12], in vivo biological imaging [13] and remote sensing [14]. Approaches able to capture more than 2 dimensions, conventionally the intensity distribution in x and y, have emerged but are accompanied by a variety of limitations such as high cost, limited use or poor point spread function and in non-snapshot cases, may not suitable for dynamic samples.

In order to image dynamic samples, high temporal resolution is required, which is met by snapshot imaging techniques [1,15]. These techniques acquire all the different dimensional information in a single exposure or through the use of parallelized detectors. Snapshot techniques typically discriminate the emitted light from a sample [5,6,16–18] although there are some techniques which encode the illumination source [19–22] and therefore discriminate based on the incident light on the sample. CASSI [16,18] and LATIS [17] are both snapshot multispectral techniques which employ dispersive elements for spectral discrimination. CASSI applies a coded pattern to the emitted light before spectral dispersion onto the detector. Each spectral element

#384535 https://doi.org/10.1364/OE.384535
Journal © 2020 Received 27 Nov 2019; revised 30 Jan 2020; accepted 8 Feb 2020; published 19 Mar 2020

Optics EXPRESS

has the same pattern but at slightly shifted horizontal positions on the detector, such that they can be differentiated from each other. LATIS uses a lenslet array to generate sub images, with void spaces in-between, which are then spectrally dispersed using a prism. At the detector the spectrally dispersed images fill the void spaces and the lenslet array can be rotated in order to most efficiently and compactly fill them. Alternatively to using dispersive elements there are commercial systems, from IMEC and Polarsens, where camera integrated filter arrays are employed. IMECs snapshot hyperspectral cameras use an on chip spectral filter array with up to 32 spectral channels, depending on the model used. Polarsens has two models for either polarization or polarization and multispectral sensitive image capture in a snapshot. The first has a polarization filter array of four polarization filters (0°, 45°, 90° and 135°) on the camera sensor and the second has the addition of a Bayer filter (RGB) overlaid over the polarization filter array. In methods where dispersive elements are used there is a non-linear spread of the point-spread-function of the shorter and longer wavelengths which results in a variation in the spectral resolution with respect to wavelength. Camera integrated systems overcome this, but limit users to single purpose (fixed dimensions) equipment often at high cost.

Ultrafast videography methods have temporal limitations [23] imposed by various elements within the setup. FDMI [24] and TMSD [25] use devices to encode individual frames of a dynamic scene with spatial modulations. Their temporal limitation is set by the update speed of a spatial light modulator (SLM), for FDMI, and a digital micromirror device (DMD), for TMSD. COSUP [26] applies a binary pseudo random pattern to the image of a dynamic scene using a DMD, however its temporal resolution is set by the mechanical rotation speed of a galvanometer scanner which temporally shears the encoded scene along the pixels of a camera. T-CUP [27] applies a similar encoding to COSUP, however it uses a streak camera to shear the scene across the pixels. This transition from mechanical limitations in the form of rotating mirrors to electronic limitations in the form of the streak camera, allows for T-CUP to enter the Tfps regime. MUSIC [28] applies spatial modulations on individual optically delayed paths that, in conjunction with a gated camera, has demonstrated a frame rate of about 1 Gfps. STAMP [29] and its follow-up SF-STAMP [30] are illumination techniques, thus limited by illumination pulse length. The techniques obtain high temporal resolution by dividing short laser pulses into a pulse train of discrete spectral components. Each individual component is then spatially separated with a dispersive optical element and as a result, they are, in their current state, restricted to the monitoring of coherence-maintaining events. The temporal resolution achievable in all these videography methods is rigid, requiring custom adjustments and calibration to vary the temporal properties, therefore lacking versatility for imaging a broad variety of dynamic scenes.

FDMI [24,31], FRAME [19–22], TMSD [25] and MUSIC [28,32] are all image multiplexing approaches based on frequency encoding. Using these techniques, high speed videography [20,24,25,28,31], field of view extension [32], volumetric and multispectral imaging [19,21,22] have been demonstrated from single exposure acquisition. Experimentally the encoding of the different images obtained in the single exposure is achieved by spatially modulating either the illumination sources or the light emitted from a scene. The sub images can then be computationally extracted from the multiplexed image using a computational analysis approach which is similar for FDMI, FRAME, TMSD and MUSIC alike. Recent work by our group [19–22] introduced and demonstrated FRAME (frequency recognition algorithm for multiple exposures) and its applicability in multidimensional imaging, capturing temporal, spatial and spectral information in different experimental setups. In these demonstrations, multiplexed images are obtained experimentally by using illumination sources with a spatial modulation which is maintained by the light from the sample. Each modulated signal corresponds to either, spatial, temporal or spectral information and is collected using a single detector and in a single exposure. Through computational analysis using the FRAME algorithm, the multiplexed images can be separated by demodulating the different sub images, found in distinct regions of the Fourier domain. These

Optics EXPRESS

demonstrations presented multispectral imaging where infinitely spectrally close source signals could be separated and temporal imaging of a pulse of light with 200 femtosecond resolution.

In this paper we present a versatile, snapshot multidimensional optical imaging approach setup based on spatial frequency encoding. This implementation of spatial frequency encoding - herein referred to as passive FRAME - works experimentally by splitting the light from the sample into different optical paths (channels), where the light is filtered and then uniquely encoded before being recombined and imaged onto the detector. Due to the encoding of the light, passive FRAME does not suffer from false light readings, an advantage it shares with other structured detection approaches and also integrated camera systems (e.g. PolarsensTM). This paper presents results from a single experimental setup where, by switching filters, we demonstrate depth-of-focus, polarization, temporal or spectral imaging. To the best of the authors' knowledge, unlike previous experimental realisations of image multiplexing approaches based on frequency encoding, this setup is the first with this versatility as well as the first to demonstrate depth-of-field, channel balancing, spectral imaging with linear unmixing analysis and polarization imaging. Additionally, since it is not an integrated camera approach it can be easily added to existing setups and used in conjunction with, in principle, any 2D detector, such as high-speed-, RGB- or intensified (time-gated) cameras.

2. Passive FRAME approach

FRAME is an imaging technique based on encoding light. Until now, the demonstrations of FRAME by our group [19–22] have been based on applying spatial modulations to the illumination sources in each setup, where spatial, temporal and spectral information have been encoded. In this paper we present, passive FRAME, an implementation of FRAME where instead, the light emitted from a sample is spatially encoded.

Figure 1 illustrates the camera view, i.e. the raw data, as well as the individual modulated images from each channel, 1-4, which it is comprised of. If the Fourier transform of the spatial domain images for each of the individual channels are taken then there are three distinct regions

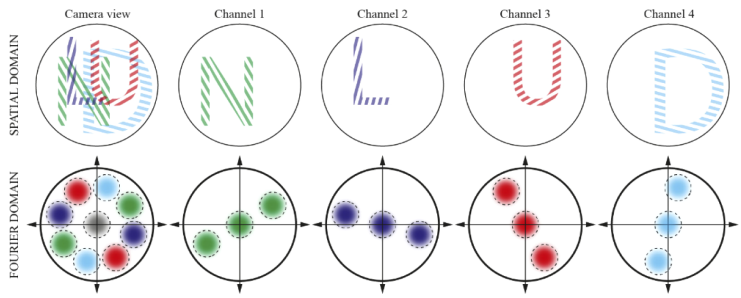


Fig. 1. Passive FRAME image encoding. The unique modulation and corresponding sample information in each channel (1-4) is illustrated in both the spatial and Fourier domains. Copies of the sample information are located in the centre of the Fourier domain as well as at the higher frequency regions corresponding to the applied modulation. The first image pair in the figure illustrates the camera view in a single exposure where all four channels are recorded in a snapshot, and its corresponding Fourier domain. Copies of all the channels can be found overlapping in the centre of the Fourier domain as well as non-overlapping copies in each higher frequency region and rotation angle (indicated by matching colors) corresponding to the unique spatial encoding of each channel (1-4).

Optics EXPRESS

seen in the Fourier domain. The central region corresponds to the frequency distribution of the image as obtained in standard non-modulated imaging. The two outer clusters correspond to the same image information but superposed multiplicatively with the spatial modulation frequency, ν, thus shifted to the corresponding higher frequency positions of that applied modulation. Since it is possible to control both the frequency and rotation of the spatial modulation applied to the light it is correspondingly possible to control where the pair of outer clusters, or modulated information, is 'placed' in the Fourier domain. Each channel, 1-4, has a unique encoding applied such that when all channels are captured simultaneously by the detector in a single exposure, the resulting information will be spatially separated in the Fourier domain. It is therefore possible to use a computational algorithm, based on the lock-in detection principle [9,22,33], in order to separate each of the image cluster pairs from the raw image, demodulate them and transform them back to the spatial domain. As a result multiple encoded images can be captured in a snapshot and separated in post processing. For more information on how to analyze modulated images see [1,21,34].

Experimentally in this paper, passive FRAME is achieved by splitting the light collected from the sample into separate channels, where each channel contains a filter and a line grating (Fig. 2). Beam splitters (BS) are used to split the light into four channels and the filter (polarization, spectral or glass plate) in each channel transmits only specific light, e.g. a certain polarization orientation. After the filter there is a line grating (20 lp/mm) where an image of the sample is incident such that it becomes encoded with the grating modulation and corresponding rotation. Each of the four channels are then recombined using BS's so that all the filtered images of the sample are spatially overlapped and imaged onto the detector.

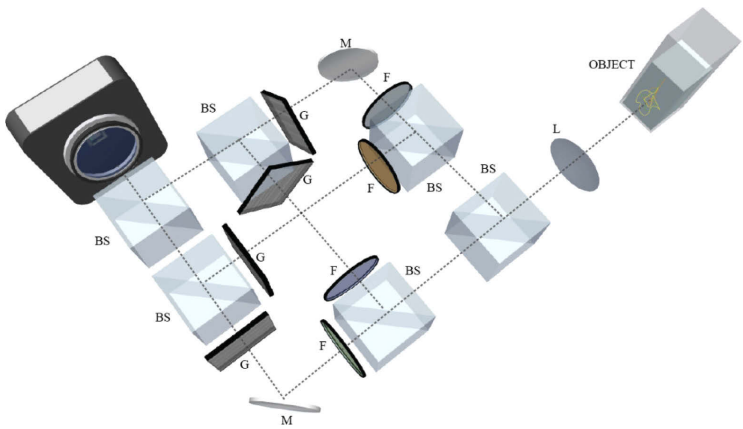


Fig. 2. Passive FRAME experimental setup. Schematic showing the four uniquely modulated and filtered (color, polarization or glass plate) channels through which light from the sample is recorded in the experimental setup. BS = Beam splitter, F = Filter (Color filter/Polarization filter/Glass Plate), M = Mirror, G = Grating, L = Imaging lens.

3. Results and discussion

Through the simple switching of filters in each of its channels we show experimentally that the passive FRAME setup can be used to image different depth-of-focus, polarization, temporal

Optics EXPRESS

or spectral properties instantaneously. Static samples are imaged using the depth-of-focus, polarization and spectral setups where the results are verified against their ground truths. Videography of a rotating computer fan is captured at 40 times the acquisition speed of the camera in the temporal setup. Finally, qualitative linear un-mixing is demonstrated using fluorescent dyes in both static and dynamic samples.

3.1. Depth-of-focus imaging

Volumetric imaging requires in focus image capture at a variety of depths and acquisition of such data in stationary samples can be time consuming. If the sample is dynamic, sequential capture of the different depths is not feasible and therefore snapshot approaches are required. By replacing the filters in the passive FRAME setup with glass plates each channel is reconfigured so that the imaging plane in each is shifted along the z-direction, thus capturing different depths-of-focus of a sample, in this case an arrangement of optical elements (Fig. 3).

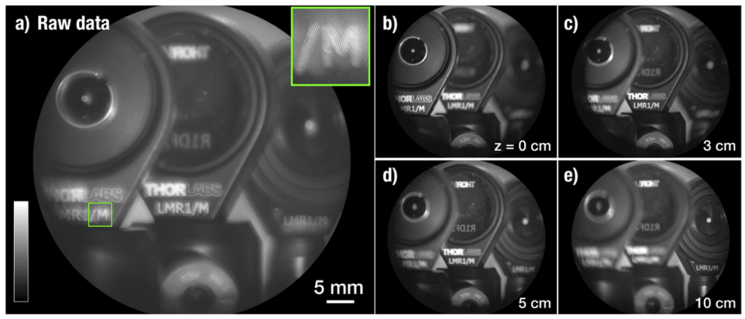


Fig. 3. Snapshot imaging of multiple depths-of-focus. An arrangement of optical components, the object, are illuminated with white light and imaged with each channel, 1-4, corresponding to a different z plane. (a) Raw data image comprised of all four modulated channels. (b) Extracted image corresponding to a z = 0 cm imaging plane. (c) Extracted image corresponding to a z = 3 cm imaging plane. (d) Extracted image corresponding to a z = 5 cm imaging plane. (e) Extracted image corresponding to a z = 10 cm imaging plane.

3.2. Polarization imaging

To demonstrate the polarization imaging of passive FRAME a white light source was used in a transmission imaging arrangement. The light was linearly polarized vertically (0°) before being incident on the sample, in this case some partially overlapping shapes cut from transparent tape. Figure 4 shows the raw data, ground truth and extracted images for the sample. In this case each of the different channels of the setup have linear polarization filters and therefore correspond to light polarized at 0°, 45°, 90° and 135°. As expected, the signals in channels 1 and 3 are inverted since channel 1 (0°) is fully transmissive and channel 3 (90°) is cross polarized. In addition, comparing the extracted with the ground truth images we obtain R^2 values of above 0.99, confirming that the passive FRAME approach is accurate and competent for polarization imaging.

3.3. Videography by time multiplexing

The passive FRAME setup was used for temporally resolved imaging of a dynamic sample, in this case a rotating computer fan, illuminated by a multi-colored pulse train, in transmission

Optics EXPRESS

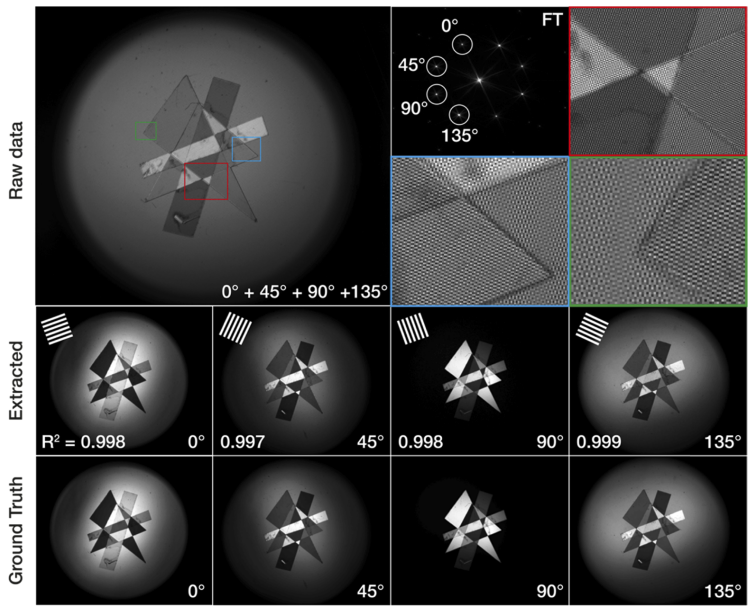


Fig. 4. Transmission polarization images of layered transparent tape shapes. Sample transmission illuminated by linearly polarized (0°) white light. *A raw data* image captured from all four modulated channels. *The Fourier Transform* of the raw data image. *Magnified regions* of the raw data image where the modulations from the different channels are visible. *The demodulated images* of channels 1, 2, 3 and 4, extracted from the raw data image, corresponding to linear polarizations of 0°, 45°, 90° and 135° respectively. *Ground truth images* for each polarization channel.

mode. Spectral filters in each channel were wavelength matched to each of the four pulses in the pulse train (see Fig. 13 in the appendix). The images from each channel therefore correspond to four different points in time. A single exposure image capturing the time sequence of the moving blade as it blocks the individual pulses was taken (Fig. 5(a)). Upon demultiplexing, a 2 kHz, four frame video, with a temporal resolution of 500 µs, was extracted (lower panel of Fig. 5). The results when using spectral filters paired with the multi-colored pulse train show how passive FRAME can be used to increase the maximum frame-rate of a camera, in this case by a factor of 40, demonstrating its temporal imaging capabilities.

3.4. Spectral imaging and signal intensity balancing across channels

To demonstrate the spectral sensitivity of the passive FRAME approach, an X-Rite Color checker nano target was imaged using white light in an absorption/reflection arrangement. Figure 6(a) shows the raw data image of the sample, where, in the magnified region, the applied modulations can be seen. Since each of the squares in the X-Rite target are different colors it is pertinent to examine more closely, their spectral characteristics, as obtained using the passive FRAME setup. Four different colored squares were selected and their spectral signatures are presented in

Optics EXPRESS

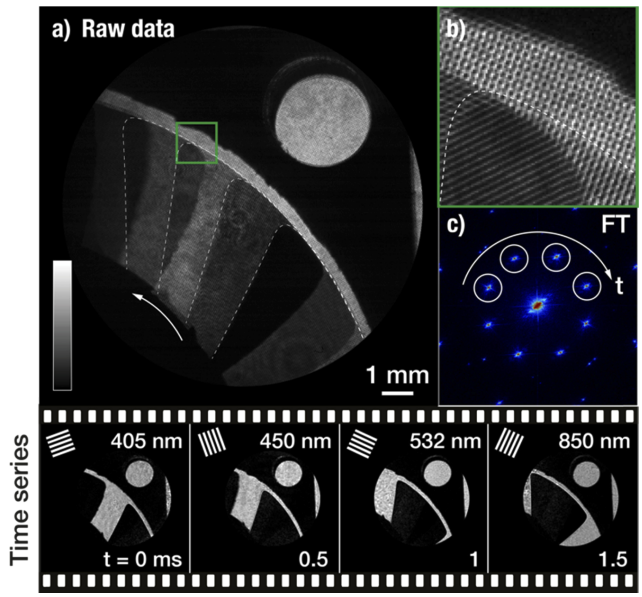


Fig. 5. **40x increase in camera speed using passive FRAME**. (a) A single-exposure transmission mode image of a rotating computer fan, illuminated by a multi-color pulse train. The dotted lines show the leading edge of the fan blade moving in the direction of the arrow as it is captured at four different points in time. (b) A magnification of the raw data. Four regions are visible corresponding to different combinations of the modulations. (c) The Fourier transform of the raw image where each cluster pair corresponds to a temporally separated snapshot image of the rotating fan blade. Upon demultiplexing, a 2 kHz time series of a rotating fan blade is shown in the lower panel. Each frame corresponds to the transmission of a pulse of given wavelength separated in time.

the bar plots (Fig. 6). It is possible to see that the spectral signatures for each of the regions is different, as we would expect due to their origin from different colored regions, demonstrating the setups spectral sensitivity and subsequent applicability for snapshot, multispectral imaging. Additionally it can be noted that the variation in the signal strengths across the channels, e.g. when looking at region of interest (ROI) 1, is quite large. When there is such a mixture of low and high intensity signals, the low intensity signals can suffer from cross talk from the neighbouring signals with higher intensity. Despite this, passive FRAME still achieves R^2 values of 0.97 or higher (Fig. 6) when comparing the extracted and ground truth images for the unbalanced results. Due to the experimental setup construction it is, however, possible improve the results by using OD filters in the different channels in order to balance the signals. By doing so the cross talk can be reduced yielding R^2 values of 0.99 for all channels, in the balanced results. Since the OD filter values are known, the channel intensities can then be rescaled to their original relative intensities, after extraction in the Fourier domain, in order to maintain the spectral sensitivity of the approach.

Optics EXPRESS

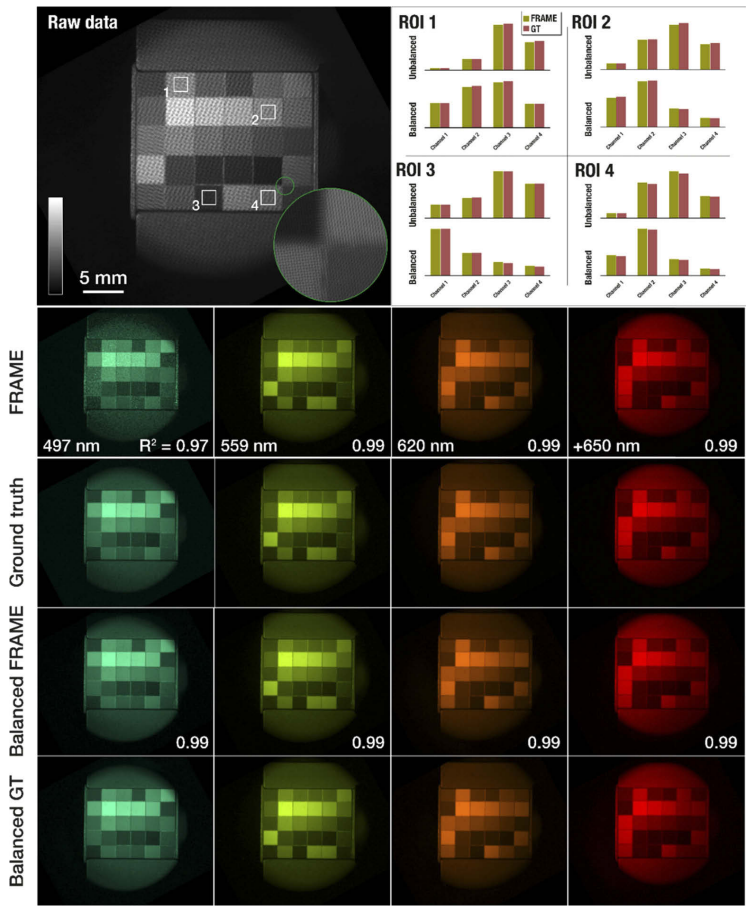


Fig. 6. Signal intensity balancing. An X-rite Color Checker imaged using white light illumination and four different spectral filters, one in each channel. *The raw image* of the target with all four modulated channels. The applied modulation patterns are visible in the magnified region. *Regions of interest* showing the spectral information of four regions of interest corresponding to four unique colors in the sample. Signal intensities for each of the channels are shown for both the unbalanced and balanced cases for the extracted and ground truth images. *Demodulated channel images* which are false colored images for channels 1 to 4 for the balanced and unbalanced extracted FRAME images compared with their corresponding ground truth images. R^2 values are given for the balanced and unbalanced cases.

Optics EXPRESS

3.5. Fluorescence imaging and linear unmixing

Linear unmixing is a mathematical method for separating different constituents in a sample and is a powerful technique for fluorophore identification in fluorescent imaging [13,35]. If a sample contains several different fluorophores which each have a characteristic spectral emission, linear unmixing can be used to determine which of these fluorophores the sample contains and in which relative quantitates. It is therefore possible to accurately image samples with multiple fluorophores with overlapping fluorescence signals and which are co-localized. Linear unmixing analysis is possible if one has access to a database of reference spectra pertaining to each of the different fluorophores that the sample may contain. By creating a series of simultaneous equations which represent the captured spectrum, and are defined using the reference spectra, they can then be solved to find the minimum values which give the minima of the reference spectra and thereby the relative concentrations of each fluorophore.

For the fluorescence imaging a blue LED was used in an absorption/reflection arrangement to image three different fluorescent dyes. The multiplexed image is shown in Fig. 7(a) along with the extracted images for each of the spectral channels. Analysis of the intensities of each of the dyes across the four spectral channels can be used to generate a catalogue of their spectral intensity profiles, or reference spectra. It is then possible to formulate simultaneous equations to describe the multiplexed image of the different dyes, in terms of these reference spectra, therefore linearly unmixing the dye letters. The result of this linear unmixing is that it becomes possible to identify which dye is present where, which is demonstrated in Fig. 7(b).

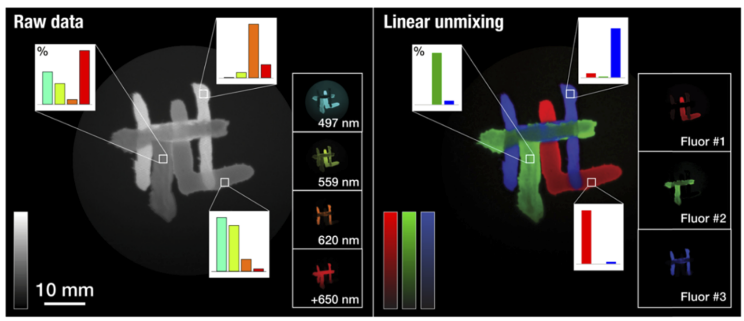


Fig. 7. Three fluorescent letters linearly unmixed. A sample of three fluorescent dyes illuminated with a blue LED (450 nm). *The raw image* of the dye letters L, T and H, captured using all four modulated channels with spectral filters. *The inserts* show the four channel intensities for each dye are shown in the bar plots, i.e. the spectral intensity profile of each dye. *Multispectral image* of the linearly unmixed dyes which are false colored red, green or blue.

Figures 8, 9 and 10 show a dynamic sample where the same three fluorescent dyes, characterised in Fig. 7, were dropped into water. A series of images were captured as a time series, whereby each image in the time series was an encoded snapshot image from all four spectral channels, i.e. allowing multispectral video capture. Figure 8(a) shows the raw data image from one of the moments in time captured in the time series. The Fourier transform of the image is seen in the inset where it is possible to see four pairs of modulated data corresponding to light captured from each of the four channels. The data is demodulated, as seen in Fig. 8(c), (d), (e) and (f), and then linearly unmixed and then combined into a single multispectral image as seen in Fig. 8(b). It can be seen from the linearly unmixed sub images, that at the chosen moment in time two of

Optics EXPRESS

the three dyes are present, fluorophores 2 and 3. Figure 9 shows another moment from the same time series where only two of the three dyes are present, now fluorophores 1 and 3. Figure 10 shows 10 demodulated and linearly unmixed images from the time series of the dynamic sample, and a video animating the data set can be seen in Visualization 1.

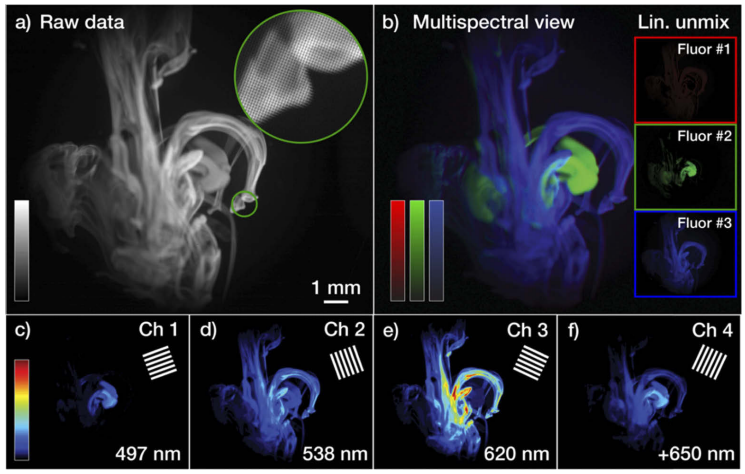


Fig. 8. Linearly unmixed dynamic sample. An image of a dynamic sample of three fluorescent dyes dropped into water (side view). Two of the three dyes are visible in the selected image from the time series. (a) A raw data image captured from all four modulated channels. (b) A multispectral image of the demodulated and linearly unmixed data. (c) The demodulated image from channel 1 only. (d) The demodulated image from channel 2 only. (e) The demodulated image from channel 3 only. (f) The demodulated image from channel 4 only.

Another dynamic sample was recorded, also using the same three fluorescent dyes, however dropped into shallow water and viewed from above. An image showing a moment in time where the first two dyes have already been dropped and allowed to disperse in the water is shown in Fig. 11, along with a third dye seen as a droplet on the surface of the water. Shortly after this moment in time the droplet rapidly explodes across the water and the full sequence can be seen in Visualization 2.

Optics EXPRESS

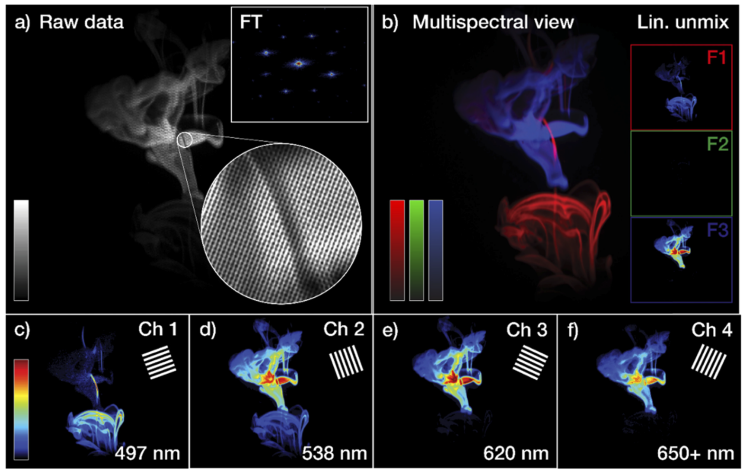


Fig. 9. Fluorescent dyes falling through water. An image of a dynamic sample of three fluorescent dyes dropped into water (side view). Only two dyes are visible in the selected image from the time series. (a) A raw data image captured from all four modulated channels. (b) A multispectral image of the demodulated and linearly unmixed data. (c) The demodulated image from channel 1 only. (d) The demodulated image from channel 2 only. (e) The demodulated image from channel 3 only. (f) The demodulated image from channel 4 only.

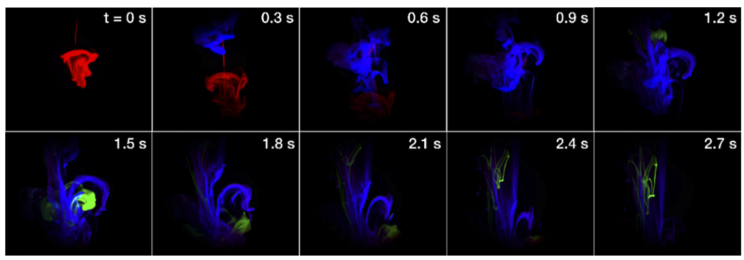


Fig. 10. Time series of linearly unmixed dynamic sample. A time series of multispectral snapshot images of a dynamic sample of three fluorescent dyes dropped into water (side view). At time 0s only one dye is present whereas at time 0.3s two are present and then at time 1.5s the first dye has passed and the latter two dyes are present. A video animating the data set is available online (Visualization 1).

Optics EXPRESS

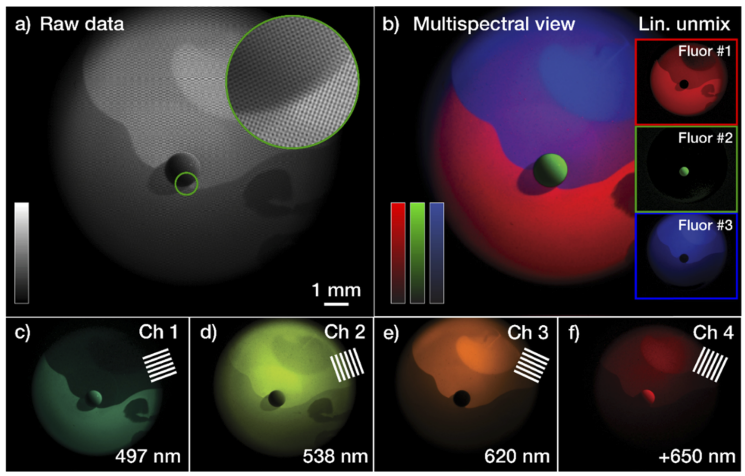


Fig. 11. Fluorescent dye droplet and two dispersed dyes. An image of a dynamic sample of three fluorescent dyes, two dropped and dispersed into water and one as a droplet on the water (top view). All three dyes are visible in the selected image from the time series. (a) A raw data image captured from all four modulated channels. (b) A multispectral image of the demodulated and linearly unmixed data. (c) The demodulated image from channel 1 only. (d) The demodulated image from channel 2 only. (e) The demodulated image from channel 3 only. (f) The demodulated image from channel 4 only. A video animating the data set is available online (Visualization 2).

4. Conclusion

In summary, this paper presents passive FRAME, a versatile multidimensional imaging approach. Imaging of depth-of-focus, polarization, temporal and spectral properties have been demonstrated experimentally. Imaging of these different dimensions was achieved by splitting light from a sample into four channels where it was then filtered and spatially encoded. The multidimensionality of the approach exhibits its applicability for imaging a broad variety of samples from biological fluorescence imaging to molecular dynamics.

Since passive FRAME is a snapshot technique it is much less sensitive to errors caused by movement in each of the extracted images and should the sample move during the acquisition, then all channels will be equally affected. It also means that multiple parameters can be acquired instantly and therefore dynamic samples can be imaged resulting in multidimensional videography being achieved. Conversely, if the samples are in fact stationary but in large quantities (many similar samples, e.g. biomedical samples) then the snapshot approach can reduce the total acquisition time, i.e. it has a high throughput.

A temporal resolution of 2 kHz has been demonstrated with the setup, corresponding to a 40x increase to the maximum frame rate of the camera. The setup is not restricted to transmission visualization since the only requirement is that temporally separated events are filtered in each of the channels. In addition, since the temporal limitation is set by the properties of the pulse train, broadband femtosecond laser sources could, in principle, push this technique into the GHz regime.

Optics EXPRESS

In the multispectral imaging format of passive FRAME, linear un-mixing, used for fluorescence imaging involving multiple different fluorophores, was demonstrated. To identify different fluorophores specific filters sets are used to accurately determine their spectral intensity profiles. Since a large catalogue of fluorophores exists, imaging systems are often developed using such specific filter sets, thus for specific sample/fluorophore applications. The presented approach therefore demonstrates how it is particularly compatible for fluorescence imaging since any filters can be used, which can be quickly and easily changed as required, for a broad variety of applications.

Due to the encoding being achieved in the four individual channels, it is not only possible to vary the filter types to achieve different dimensional imaging, but also to control the intensities in each channel by adding ND filters. This is useful when studying samples where there are large variations in the emission intensities. Since, for snapshot imaging, acquisition is performed in a single exposure, these weaker signals would normally be overwhelmed, however the combination of the encoding and the individual channels in passive FRAME means this can be overcome.

Finally, beyond the multidimensional imaging properties the setup is also versatile in its construction and combination with other lab equipment. Since it is not integrated into the detector system it is compatible with, in principle, all 2D detectors. As a result, the approach can be used in conjunction with user's current detectors, reducing the cost of implementing it into existing research and industry environments. The setup is also fully static and constructed from standard optical components. As technological advances come about, in both detectors and optical elements, the passive FRAME setup can be upgraded to take advantage of such improvements without the need to invest in an entirely new imaging system. Finally, passive FRAME is not fundamentally limited to four channels and various solutions are available to increase the optical throughput and number of images/dimensions instantaneously achievable, making it an even more powerful snapshot, multidimensional imaging approach.

Appendix

A1. Spatial resolution trade-off

Since FRAME applies a spatial modulation in order to encode the different dimensional information it sacrifices spatial resolution in the resulting images obtained. In some applications this compromise may prove too costly. The results in Fig. 12, showing an image of a resolution

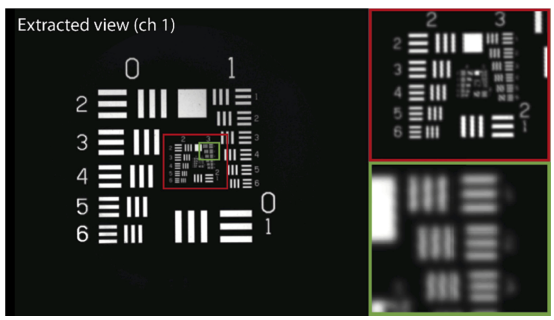


Fig. 12. Spatial resolution and field of view. Resolution target imaged using the passive FRAME setup. The demodulated image from channel 1 of the setup (left) along with magnified regions of the higher resolution area of the target (right).

Optics EXPRESS

target (Edmund Optics 38256) as acquired using the proof of concept passive FRAME setup, demonstrate that 11.30 lp/mm was resolvable and the corresponding field of view (FOV) for this resolution. It should be noted however that, depending on the application of the passive FRAME approach, detector used and other optical elements, the FOV and resolution achievable will vary, with the potential for even higher resolutions to be achieved.

A2. Pulse train characteristics

A multi-colored pulse train was used for temporally resolved imaging in transmission mode. It was generated by modulating individual diode lasers of wavelengths 405 nm, 450 nm, 532 nm and 850 nm with a separation of 500µs. The pulse train was optimized such that a maximum frame rate could be achieved while still respecting a reasonable definition for temporal resolution; the arrival of a pulse should coincide with the fall time, $\tau_{1/2}$, of the previous pulse. The measurement was performed using a Thorlabs DET10A\M fast photodetector yielding an average fall time of $\tau_{1/2} \approx 370\mu s$ for all four laser diodes (Fig. 13). This investigation, shows that the setup is indeed limited by the pulse train characteristics.

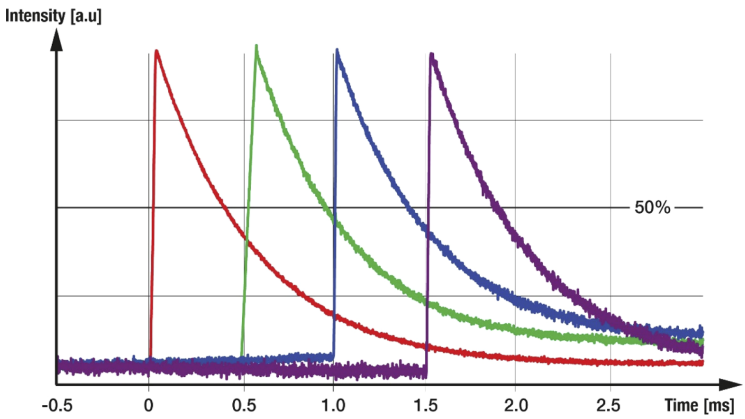


Fig. 13. Pulse Train Characteristics. Four individual, color coded pulses incident on the computer fan measured using a Thorlabs DET10A\M photodetector. The long fall time, calculated by finding the intersection between the pulse and the half the maximum intensity line, is on average 370µs for all the four laser diodes.

Funding

Vetenskapsrådet (121892); European Research Council (803634).

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. L. Gao and L. V. Wang, "A review of snapshot multidimensional optical imaging: measuring photon tags in parallel," Phys. Rep. **616**, 1–37 (2016).

2. A. Manakov, J. Restrepo, O. Klehm, R. Hegedus, E. Eisemann, H.-P. Seidel, and I. Ihrke, "A Reconfigurable Camera Add-On for High Dynamic Range, Multispectral, Polarization, and Light-Field Imaging," ACM Trans. Graph. **32**(4), 1 (2013).
3. J. T. Fredrich, "3D imaging of porous media using laser scanning confocal microscopy with application to microscale transport processes," Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy **24**(7), 551–561 (1999).
4. J. Kim and A. Ghosh, "Polarized Light Field Imaging for Single-Shot Reflectance Separation," Sensors **18**(11), 3803 (2018).
5. S. G. Demos and R. R. Alfano, "Optical polarization imaging," Appl. Opt. **36**(1), 150–155 (1997).
6. X. Xu, Y. Qiao, and B. Qiu, "Reconstructing the surface of transparent objects by polarized light measurements," Opt. Express **25**(21), 26296–26309 (2017).
7. J. Hunicz and D. Piernikarski, "Investigation of combustion in a gasoline engine using spectrophotometric methods," Proc. SPIE **4516**, 307–314 (2001).
8. P. Kauranen, S. Andersson-Engels, and S. Svanberg, "Spatial mapping of flame radical emission using a spectroscopic multi-colour imaging system," Appl. Phys. B: Photophys. Laser Chem. **53**(4), 260–264 (1991).
9. P. S. Hsu, D. Lauriola, N. Jiang, J. D. Miller, J. R. Gord, and S. Roy, "Fiber-coupled, UV–SWIR hyperspectral imaging sensor for combustion diagnostics," Appl. Opt. **56**(21), 6029–6034 (2017).
10. J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy," Nat. Methods **2**(12), 910–919 (2005).
11. D. M. Chudakov, S. Lukyanov, and K. A. Lukyanov, "Fluorescent proteins as a toolkit for in vivo imaging," Trends Biotechnol. **23**(12), 605–613 (2005).
12. C. E. Volin, B. K. Ford, M. R. Descour, J. P. Garcia, D. W. Wilson, P. D. Maker, and G. H. Bearman, "High-speed spectral imager for imaging transient fluorescence phenomena," Appl. Opt. **37**(34), 8112–8119 (1998).
13. T. Zimmermann, J. Rietdorf, and R. Pepperkok, "Spectral imaging and its applications in live cell microscopy," FEBS Lett. **546**(1), 87–92 (2003).
14. A. Jung, R. Michels, and G. Rainer, "Portable snapshot spectral imaging for agriculture," Acta agrar. Debr. **150**, 221–225 (2018).
15. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," Opt. Eng. **52**(9), 090901 (2013).
16. A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," Appl. Opt. **47**, B44–B51 (2008).
17. J. G. Dwight and T. S. Tkaczyk, "Lenslet array tunable snapshot imaging spectrometer (LATIS) for hyperspectral fluorescence microscopy," Biomed. Opt. Express **8**(3), 1950–1964 (2017).
18. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, "Video rate spectral imaging using a coded aperture snapshot spectral imager," Opt. Express **17**(8), 6368–6388 (2009).
19. E. Kristensson, Z. Li, E. Berrocal, M. Richter, and M. Aldén, "Instantaneous 3D imaging of flame species using coded laser illumination," Proc. Combust. Inst. **36**(3), 4585–4591 (2017).
20. A. Ehn, J. Bood, Z. Li, M. Aldén, and E. Kristensson, "FRAME: femtosecond videography for atomic and molecular dynamics," Light: Sci. Appl. **6**(9), e17045 (2017).
21. K. Dorozynska and E. Kristensson, "Implementation of a multiplexed structured illumination method to achieve snapshot multispectral imaging," Opt. Express **25**(15), 17211–17226 (2017).
22. Z. Li, J. Borggren, E. Berrocal, A. Ehn, M. Aldén, M. Richter, and E. Kristensson, "Simultaneous multispectral imaging of flame species using Frequency Recognition Algorithm for Multiple Exposures (FRAME)," Combust. Flame **192**, 160–169 (2018).
23. J. Liang and L. V. Wang, "Single-shot ultrafast optical imaging," Optica **5**(9), 1113–1127 (2018).
24. S. R. Khan, M. Feldman, and B. K. Gunturk, "Extracting sub-exposure images from a single capture through Fourier-based optical modulation," Signal Process. Image Commun. **60**, 107–115 (2018).
25. M. Gragston, C. D. Smith, and Z. Zhang, "High-speed flame chemiluminescence imaging using time-multiplexed structured detection," Appl. Opt. **57**(11), 2923–2929 (2018).
26. X. Liu, J. Liu, C. Jiang, F. Vetrone, and J. Liang, "Single-shot compressed optical-streaking ultra-high-speed photography," Opt. Lett. **44**(6), 1387–1390 (2019).
27. J. Liang, L. Zhu, and L. V. Want, "Single-shot real-time femtosecond imaging of temporal focusing," Light: Sci. Appl. **7**(1), 42 (2018).
28. M. Gragston, C. Smith, D. Kartashov, M. N. Shneider, and Z. Shang, "Single-shot nanosecond-resolution multiframe passive imaging by multiplexed structured image capture," Opt. Express **26**(22), 28441–28452 (2018).
29. K. Nakagawa and A. Iwasaki, "Sequentially timed all-optical all-optical mapping photography (STAMP)," Nat. Photonics **8**(9), 695–700 (2014).
30. T. Suzuki and R. Hida, "Single-shot 25-frame burst imaging of ultrafast phase transition of Ge2Sb2Te5 with a sub-picosecond resolution," Appl. Phys. Express **10**(9), 092502 (2017).
31. B. K. Gunturk and M. Feldman, "Frequency division multiplexed imaging," Proc. SPIE **8660**, 86600P (2013).
32. M. Gragston, C. D. Smith, J. Harrold, and Z. Zhang, "Multiplexed structured image capture to increase the field of view for a single exposure," OSA Continuum **2**(1), 225–235 (2019).
33. E. Kristensson, A. Ehn, and E. Berrocal, "High dynamic spectroscopy using a digital micromirror device and periodic shadowing," Opt. Express **25**(1), 212–222 (2017).
34. E. Berrocal, J. Johnsson, E. Kristensson, and M. Aldén, "Single scattering detection in turbid media using single-phase structured illumination filtering," J. Europ. Opt. Soc. Rap. Public. **7**, 12015 (2012).
35. T. Zimmermann, "Spectral imaging and linear unmixing in light microscopy," Adv. Biochem. Eng./Biotechnol. **95**, 245–265 (2005).

Paper VIII

scientific reports

Check for updates

OPEN

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

Karolina Dorozynska[1], Simon Ek[1], Vassily Kornienko[1], David Andersson[1], Alexandra Andersson[2], Andreas Ehn[1] & Elias Kristensson[1]✉

Fluorescence-based multispectral imaging of rapidly moving or dynamic samples requires both fast two-dimensional data acquisition as well as sufficient spectral sensitivity for species separation. As the number of fluorophores in the experiment increases, meeting both these requirements becomes technically challenging. Although several solutions for fast imaging of multiple fluorophores exist, they all have one main restriction; they rely solely on spectrally resolving either the excitation- or the emission characteristics of the fluorophores. This inability directly limits how many fluorophores existing methods can simultaneously distinguish. Here we present a snapshot multispectral imaging approach that not only senses the excitation and emission characteristics of the probed fluorophores but also all cross term combinations of excitation and emission. To the best of the authors' knowledge, this is the only snapshot multispectral imaging method that has this ability, allowing us to even sense and differentiate between light of equal wavelengths emitted from the same fluorescing species but where the signal components stem from different excitation sources. The current implementation of the technique allows us to simultaneously gather 24 different spectral images on a single detector, from which we demonstrate the ability to visualize and distinguish up to nine fluorophores within the visible wavelength range.

Fluorescence multispectral imaging is a powerful tool with applications in life science[1–4], photonics[5–7], engineering[8] and drug discovery[9]. The technique uses fluorescent markers to determine concentrations, structures, functions and interactions between different sample components.

As the demand for information increases[2,5,6], the diagnostic methods used require the means to increase the number of labels that can be simultaneously imaged. Traditionally, this is achieved using sequential imaging with one or several excitation wavelengths combined with optical filtering of the fluorescence signal[10]. In recent years, powerful scanning methods have been developed, such as the hyperspectral multicolor imaging reported by Jahr *et al.*, capable of distinguishing up to five fluorophores[7]. Another such example exploits multiplexed imaging based on stimulated Raman scattering[5,6]. In combination with engineered fluorescent probes whose spectral responses are well-separated, the visualization of a large number of distinctive species is enabled. However, both sequential and scanning approaches have a restricted temporal resolution and are therefore inefficient for the imaging of fast dynamic- or one-time events.

To meet the demand for imaging with higher temporal resolution, there has been an emergence of novel fluorescence-based fast, or even snapshot, multicolor imaging approaches[1,3,4,11–18]. In general, currently available solutions within this category are either based on sensing differences in the excitation spectra[1,3,4,12,15,17,18] or by spectrally resolving the fluorescence emission[7,13,14,19]. For example, Mahou *et al.* demonstrated in 2012 a multicolor two-photon excitation-based imaging system capable of visualizing three chromophores simultaneously[18]. Lavagnino *et al.* constructed a snapshot hyperspectral emission-based imaging technique with 60 spectral channels, demonstrating the parallel determination of three fluorescence markers[19]. In 2018, Garbacik *et al.* developed a time-multiplexed excitation-based multispectral imaging system that further advanced the field, allowing for the visualization of six fluorophores, demonstrated for scanning confocal imaging[15]. Both excitation and emission-based approaches are, however, ultimately limited by (1) the spectral overlap between the different fluorophores, known as the 'color barrier', and (2) the unavoidable decrease in signal to noise ratio (SNR) in

[1]Department of Combustion Physics, Lund University, 22363 Lund, Sweden. [2]Department of Physical Chemistry, Lund University, 22100 Lund, Sweden. ✉email: elias.kristensson@forbrf.lth.se

rapid/snapshot systems[20]. These limitations explain the restricted number of possible fluorescent markers within the visible/NIR spectrum that current technology can image simultaneously. However, even if many fluorescent markers may display large similarities in their excitation- or emission spectra, it is highly unlikely that *both* these characteristics are similar and hence being able to probe both properties simultaneously has the potential of greatly increasing the number of accessible fluorescent markers within the limited wavelength span offered by currently available cameras. To date, such multi-dimensional data can only be acquired using sequential/scanning imaging[5,6] or time-multiplexing with multiple detectors[16,17].

In this paper we report an alternative avenue for snapshot multispectral wide-field imaging that enables 'full optical path tracing', allowing us to sample both the excitation- and emission profiles of all fluorophores within a single acquisition and with a single detector. Unlike most existing configurations for multicolor imaging, our approach maintains the full temporal resolution of the detector and does not rely on any particular hardware, such as fast, time-resolved acquisition[4,15] or pulsed lasers[5,6]. Instead, our configuration is based on the Frequency Recognition Algorithm for Multiple Exposures (FRAME) concept[21], in which the excitation sources as well as the resulting fluorescence signals are spatially multiplexed, in two separate stages, yielding a doubly encoded intensity mapping of the sample. Using this novel approach alongside four different CW laser sources and four different spectral emission channels, we demonstrate that this combination yields a multiplexed image of 24 different images: one for each of the four lasers, one for each of the four spectral channels and 16 for all cross term combination. By analyzing these 24 images using linear unmixing, we demonstrate the ability to separate nine different fluorophores within the visible spectrum, despite significant overlap in their respective excitation and emission profiles. Furthermore, because we gain a more comprehensive and multi-dimensional picture of the fluorophores, we can achieve accurate distinction at a significantly lower SNR. To the best of the authors' knowledge, this is the first demonstration of full optical path tracing-from source to detector-using multiple tandem illumination sources and spectral emission channels for snapshot imaging, providing high-throughput, low phototoxicity and outstanding spectral sensitivity.

Results

Double modulated FRAME. FRAME is a multiplexing imaging technique that allows several images—in this case carrying spectral information—to be acquired in parallel using a single detector[12,13,21]. With FRAME, multiplexing of the image data is achieved by superimposing a unique spatial frequency (intensity modulation) onto each individual image prior to detection. The captured image thus contains the sum of all intensity-modulated images and although the image information overlaps in the spatial domain, the data is separated in reciprocal space. By applying a frequency-sensitive lock-in detection algorithm the multiplexed image information can be extracted and separated, albeit at a reduced spatial resolution[21]. The spatial resolution of any FRAME-system is thus not directly dictated by the pixel resolution given by the camera, but rather by the distance between the imposed spatial frequencies in reciprocal space, which in turn sets the limit for the dimensions of the filter function used in the lock-in detection algorithm[22]. For the current optical setup, the spatial resolution is ~ 3 lp/mm for the extracted frames and for a field-of-view of 21×17 mm^2.

FRAME has previously been demonstrated to be capable of multispectral snapshot imaging using either excitation[12] or emission-based[13] image multiplexing, where the intensity-modulation encoding was applied on the different laser intensity profiles or the different spectral regions of the emitted fluorescence signal, respectively. Figure 1a–d, demonstrates the data acquisition for the respective cases. In this paper we demonstrate a combined approach, dubbed double modulated FRAME (DM-FRAME) where encoding is simultaneously applied to both the excitation and the emission. This approach allows us to sample and sense the molecular response (fluorescence) for each excitation wavelength of all fluorophores present in the sample within a single snapshot. This kind of cross term information, which previously only could be acquired using sequential or scanning-based imaging, is contained within the recorded image as beatings between the intensity-modulations of the excitation sources and those on the detection side. The beating results in a satellite-like structure in reciprocal space (Fig. 1f) where each such Fourier peak carries unique image information arising from a specific combination of excitation wavelength and optical filter. This spatial frequency-encoded spectral data is accessible by means of a spatial lock-in detection computer algorithm (Supplementary Information), which demodulates and isolates each Fourier component, in this case yielding 24 images, all with different spectroscopic information. However, only 16 of these 24 images contain truly unique information, setting the upper limit on the number of distinguishable fluorophores to 16 with the current optical configuration. As the spectral information is obtained using spatial encoding rather than e.g. dispersive elements, DM-FRAME has the unique ability to differentiate between fluorescence light emitted at exactly the same wavelength, simultaneously emitted by the same fluorophore but stemming from different excitation sources.

Here the DM-FRAME approach is demonstrated for fluorescence multicolor imaging, in particular to distinguish fluorophores that have strongly overlapping excitation- or emission characteristics within the visible wavelength region. Our results show that the more comprehensive and detailed view of the molecular response obtained using DM-FRAME leads to an improved species-specificity that, in turn, both increases the number of fluorophores that can be imaged simultaneously and reduces acquisition times beyond the current limitations.

Unmixing fluorophores with highly overlapping spectral characteristics. If an experiment only requires a few fluorescent labels it is often possible to select ones that have well-separated spectral excitation and/or emission signatures and thereby use suitable illumination sources or spectral filters to distinguish them[7]. However, the possibility to differentiate between fluorophores using such approaches diminishes as the demand for simultaneous imaging of a greater number of markers grows[4,5]. To demonstrate the improved species-specificity with our DM-FRAME method, four fluorophores (coumarin, rhodamine, pyrene and triphenylmethane/

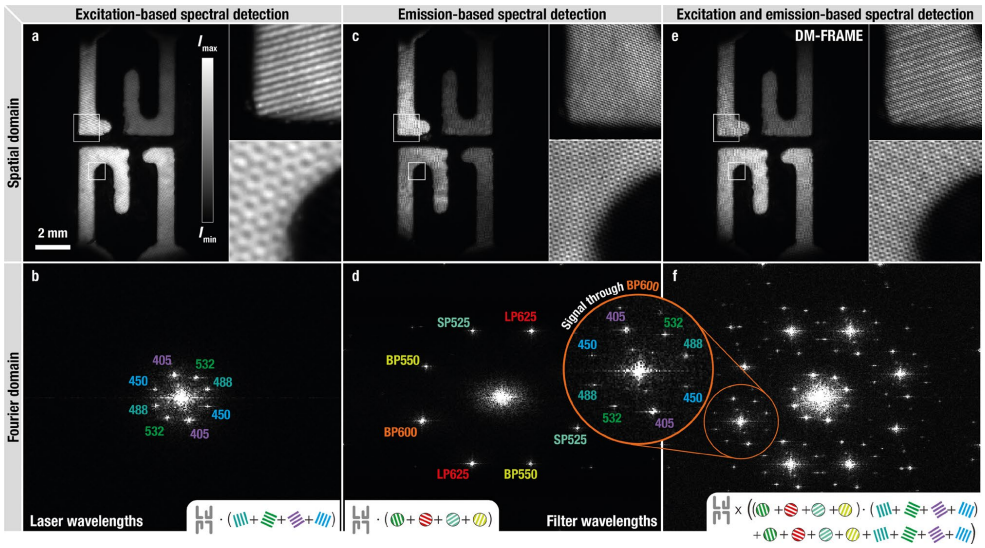


Figure 1. Example of DM-FRAME and its information storage capabilities. The fluorescence signals from four dyes are imaged using different spatial frequency multiplexing approaches. (**a**) The multispectral information is obtained by applying unique sine modulation patterns onto the intensity profile of four excitation light sources. (**b**) Fourier transform of (**a**). In (**c**) the multispectral information is obtained by guiding the light through four differently coded spectral emission channels (excitation at 450 nm). (**d**) Fourier transform of (**c**). (**e**) The two approaches combined; both the excitation sources and the detection channels are intensity-modulated. (**f**) Fourier transform of (**e**). The beating between the two applied intensity modulations yields satellite-like structures in the frequency domain. In total, this results in 24 pairs of peaks, each carrying unique 2D spectral information that can be accessed using a spatial frequency lock-in algorithm. SP = short-pass, BP = band-pass and LP = long-pass.

pyrene) with heavily overlapping spectral characteristics (Fig. 2a) were visualized using three different multispectral FRAME snapshot imaging configurations; (1) an excitation-based system with laser excitation at 405 nm, 450 nm, 488 nm and 532 nm, (2) an emission-based system equipped with optical filters, one shortpass (< 525 nm), two bandpass (525–575 nm and 575–625 nm) and one longpass (> 625 nm) and (3) a combination of these two systems, i.e. DM-FRAME. The results show that neither the excitation-based (Fig. 2b) nor the emission-based (Fig. 2c) systems are capable of accurately distinguishing the fluorophores through linear unmixing, even though the problem is mathematically solvable. This inability can be understood by examination of the similarities between the fluorescent responses-the base functions-of the individual fluorophores: pyrene and triphenylmethane/pyrene have an R^2-value of 0.98 for the excitation-based system whereas coumarin and pyrene have an R^2 value of 0.97 for the emission-based system. Distinguishing between fluorescent markers with such high R^2-values is feasible but requires high signal levels (Fig. 3a,b). Using the same optical components, i.e. the same excitation sources and optical filters, DM-FRAME overcomes the challenge of classifying markers with such similar spectral characteristics by extracting 24 different spectral images for each fluorophore using its unique full optical path tracing ability. Linear unmixing using these images (Fig. 2d) yields a 2D mapping of the four fluorophores with negligible misidentification and with a specificity that outperforms both excitation- and emission-based detection in all cases (Fig. 2e).

Spectral unmixing of nine fluorophores. Identifying a large number of fluorescent markers in a dynamic sample using multispectral imaging has been a long-standing challenge. Recent development towards fast hyperspectral imaging of multiple fluorophores are based on either scanning or sequential data acquisition, yet such approaches have a temporal resolution ultimately limited by the electronics or mechanical constraints[11]. The root challenge with simultaneous imaging of several fluorescent markers is that liquids and solids, unlike gas molecules, intrinsically have broad and featureless excitation- and emission profiles. In practice, it thus becomes difficult to find a range of suitable solid/liquid marker candidates that will not spectrally overlap significantly within the visible and near-IR region, where cameras are sensitive. Instead of attempting to circumvent the problem with broad excitation- and emission characteristics, DM-FRAME gathers a more comprehensive and complete view (24 spectral responses) of the probed markers, allowing us to fit a significantly higher number of fluorescent markers within the same camera-restricted wavelength span. To demonstrate this, we simultane-

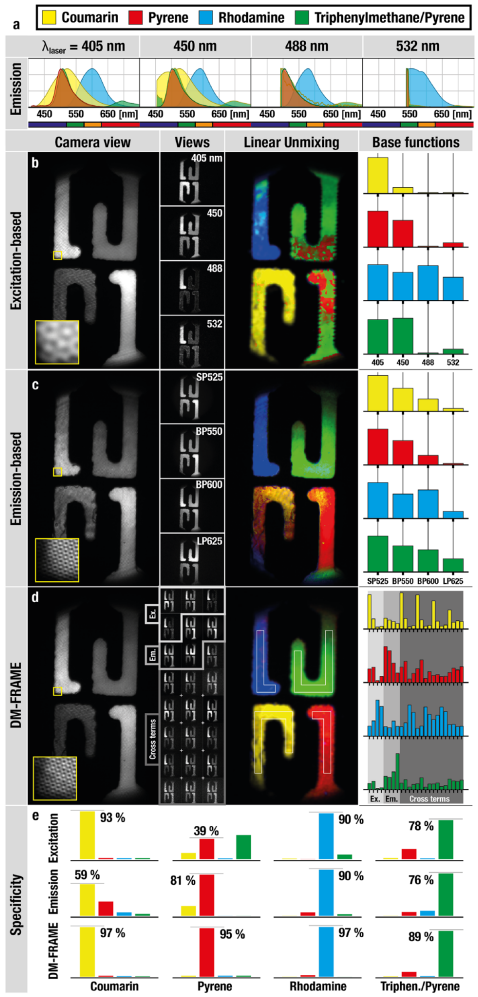


Figure 2. Improved spectral sensitivity gained by DM-FRAME. Measurements of four different dyes with very similar excitation and/or emission characteristics, analyzed using linear unmixing. (**a**) Spectral responses for each fluorophore at each excitation wavelength, together with an indication of the spectral bandwidths of the emission filters. (**b**) Excitation-based multiplexed imaging, (**c**) emission-based multiplexed imaging and (**d**) DM-FRAME. Each measurement contains sufficient information to theoretically identify the dyes but due to the strong spectral overlap, both the emission-based and excitation-based methods are prone to misidentification, see (**e**). However, when combining the methods using DM-FRAME, the faint differences between the dyes' spectral characteristics can be detected, yielding highly accurate identification (measured for the signal in the areas marked in **d**).

ously image a sample containing nine different fluorophores (Fig. 4a) with highly overlapping spectral features (Fig. 4e) and successfully perform linear unmixing to distinguish them all (Fig. 4c). Since the number of fluorophores is higher than the number of excitation- and emission channels combined, access to the 16 cross-term responses (Fig. 4b and d) turn the otherwise underdetermined system of equations into an overdetermined one. To the best of the authors' knowledge, this is the highest number of fluorescent markers simultaneously visualized using multispectral imaging.

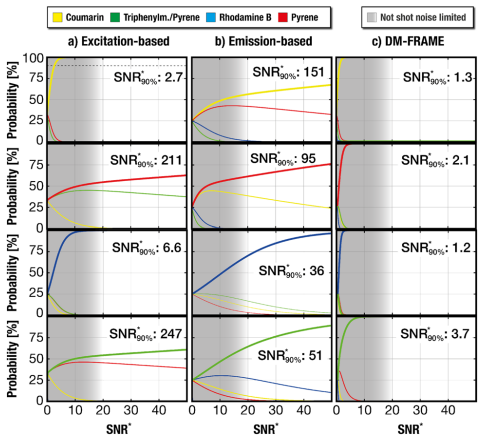


Figure 3. Estimations of the probability of correct fluorophore classification using the three different methods of Fig. 2, modeled for a purely shot noise limited system (indicated by the asterisk above the SNR). The indicated SNR* values correspond to an accurate classification of the fluorophores at a probability of 90%. (**a**, **b**). Neither the excitation- nor the emission-based detection scheme is able to accurately separate the full set of dyes at SNR* below 150. This susceptibility to noise induced uncertainties could hence be an explanation to the misidentifications observed in the linear unmixing of the corresponding cases in Fig. 2b,c, despite being mathematically solvable. (**c**) The ability of DM-FRAME to sense both emission- and excitation characteristics as well as the increased number of spectral channels it extracts yields a significantly improved species specificity, indicated by the fact the the curves reach 100% at low signal levels. Analysis of the background noise in our data indicates that the shot noise limited approximation holds for our imaging system at an SNR* above ~15 (illustrated by the gray shadowing) and that the shot noise can be approximated by a Gaussian distribution[23] above an SNR* ~ 5.

Multicolor imaging on a millisecond time scale. According to Cox and Sheppard, all imaging systems have a constant information capacity-bandwidth-that is, in turn, dictated by the system's spatial and temporal parameters[24]. Any of these parameters can be increased beyond their classical limit or replaced by another type of data by reducing the bandwidth of another parameter. For example, super-resolution microscopy trades its temporal bandwidth for an augmented spatial bandwidth[25]. Similarly, several optical methods aiming at fast multicolor imaging of several fluorescent markers rely on time-based multiplexing[3–6,15], thereby trading temporal resolution for access to spectroscopic data. For such approaches, the acquisition time unavoidably increases with the number of simultaneously measured markers, leading to a degraded temporal resolution. Emission-based methods have similar time constraints as these need to surpass a sufficient signal level in order to accurately differentiate the markers (see e.g. Fig. 4). Our solution-DM-FRAME-exploits the spatial domain to multiplex spectral data and we are therefore not bound by the same constraints as time-based methods, allowing us to record and separate several species simultaneously at record-breaking low integration times. We demonstrate this capability experimentally by (1) acquiring a multispectral 5.5 MPixel video at 4 Hz (Supplementary Video 1) and (2) recording and successfully de-multiplexing six species with an integration time of 1 millisecond (Fig. 5a). To the best of our knowledge, this is the fastest simultaneous multi-species acquisition of this many fluorophores to date.

A DM-FRAME configuration is highly adaptable and not restricted to a certain combination of excitation sources and optical filters. This versatility can be beneficial when probing a multiple number of fluorophores with low quantum yields as the excitation sources can be optimally selected so to provide maximum fluorescence emission. The results in Fig. 5, supported by the trends in Fig. 3, suggests that DM-FRAME is able to operate under such challenging, low signal conditions.

Finally, optimization of the detection configuration, i.e. replacing the beam splitters with more light-efficient dichroic mirrors, would yield a 16 fold increase in signal strength and thus a direct reduction in integration time, allowing for even faster data acquisition rates. Although the performance of DM-FRAME depends on e.g. pixel resolution and sample luminosity, its maximum acquisition rate is ultimately governed by the specifications of the camera system. It therefore holds potential for multicolor imaging of dynamic, one-time events in e.g. the kHz-MHz regime when combined with state-of-the-art high-speed video cameras[26] or FRAME-based methods for videography[27] that can offer high pixel resolutions.

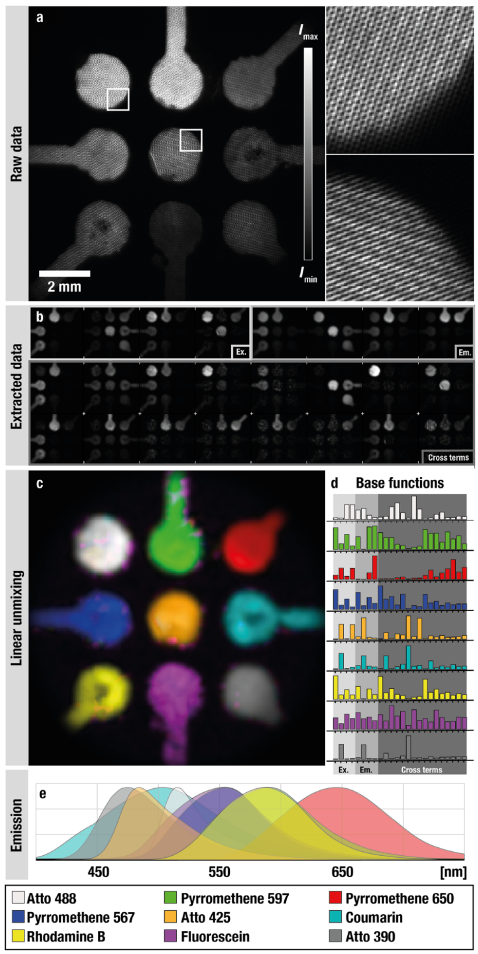


Figure 4. Linear unmixing of nine fluorophores. Nine fluorophores were probed using DM-FRAME where (**a**) is the raw unprocessed image as seen by the camera. The beating patterns, faintly visible in the magnified views, correspond to the doubly modulated information. (**b**) From the acquired data, 24 spectral images were extracted and subsequently used to perform the linear unmixing of the data. (**c**) The outcome of the spectral analysis, showing the nine different fluorophores in false colors. (**d**) The base functions (reference responses) for all nine fluorophores, where each fluorophore can be seen to acquire a unique footprint when imaged through the DM-FRAME system. (**e**) The spectrally resolved fluorescence signal (excitation 405 nm) for all nine fluorophores, indicating strong overlap of the featureless spectra.

Discussion

In summary, we have presented an optical arrangement-DM-FRAME-that allows the fluorescence of each probed fluorescent marker to be spectrally resolved in 2D for several different excitation wavelengths in parallel and with a single camera (illustration of the optical arrangement is given in Fig. 6). This feature allows the excitation of the fluorophores to be optimized, thereby providing high light-efficiency while simultaneously minimizing the risk for photobleaching. Technically, DM-FRAME performs two separate spatial multiplexing stages, one on the illumination side and one on the detection side. The acquired image is constructed by the product of these stages, i.e. a doubly multiplexed image. This technical advancement yields a powerful and effective means-full optical path tracing-to augment subtle differences in the spectral characteristics of different fluorescent markers otherwise only differentiable at high SNRs. Full optical path tracing thus opens up for multicolor imaging

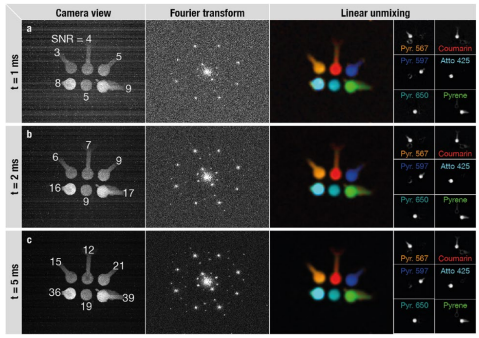


Figure 5. Multicolor imaging of six species on a millisecond time scale. Six species simultaneously visualized and separated using linear unmixing, with an acquisition time of (**a**) 1 ms, (**b**) 2 ms and (**c**) 5 ms. The results demonstrate the robustness of DM-FRAME and the ability to accurately separate fluorophores even at low SNR values (estimated SNR specified for each marker).

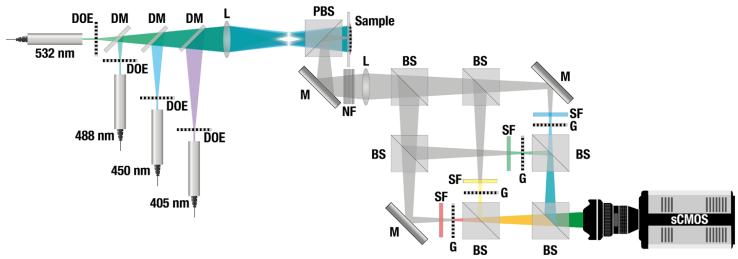


Figure 6. Schematic of the optical setup. The illumination consists of four CW laser sources, each guided through a diffractive optical element (DOE, company Holo/Or) which, when imaged onto the sample, creates a spatial intensity modulation pattern. The fluorescence emitted by the sample is then guided through a detection system where the light is split into four different paths. Each path contains a spectral filter (SF) and a transmission grating (G) which encodes the light. When the modulated fluorescence signal is transmitted through the gratings a unique beating pattern is created. This grants optical tracing abilities from the light source to the detector. DM = Dichroic Mirror, L = Lens, PBS = Polarizing Beam Splitter, M = Mirror, BS = Beam Splitter. NF = Notch Filter.

with high species-specificity using standard dyes and fluorescent markers. We demonstrate this capability here through the parallel detection of nine fluorophores-the highest number of simultaneously detected fluorophores to date-whose strongly overlapping fluorescence signals all reside within the visible wavelength region.

Full optical path tracing is accomplished by reducing the cameras spatial bandwidth. However, since the setup is all-optical, contains no moving parts and the data is acquired within a single snapshot, there are no direct requirements on the camera employed. The loss in spatial resolution can thus be partly remedied through either higher modulation frequencies on the illumination side[25] or the use of modern high-resolution cameras, especially since only one detector is required. This lack of requirements further allows DM-FRAME to be combined with advanced intensified cameras for UV-sensitive multispectral detection or high-speed video cameras for high-speed multicolor videography in the future. In addition, the ability to perform temporally-resolved imaging may also be used to determine the fluorescence lifetimes[28,29]. Including such information in the data analysis together with the spectroscopic data provided by DM-FRAME, could potentially provide an even more enhanced species-specificity.

Furthermore, even if DM-FRAME is demonstrated for macroscopic epi-fluorescence in this paper, the technique is not bound to neither macroscopic imaging nor such a configuration but could, in principle, be adapted for all methods compatible with structured illumination, such as selective plane illumination[30], two-photon excitation[31] and total internal reflection imaging[32].

Finally, fluorophore determination is not necessarily bound to imaging but could also be employed for e.g. screening and determination of protein concentration, where the parallelized detection provided by the DM-FRAME concept opens up for high-throughput spectroscopic analysis.

Methods

Optical setup. The experimental setup used consists of two encoding portions, illumination and emission encoding respectively. In the first encoding part, four continuous wave diode lasers operating at wavelengths of 405 nm (500 mW), 450 nm (2 W), 488 nm (120 mW) 532 nm (200 mW) are used. The lasers were chosen to be relatively evenly distributed over the more energetic part of the visible spectrum to enable excitation of the broad range of fluorophores used in the experiments. The lasers were further attenuated using ND-filters to match the power of the least powerful laser at 488 nm. The beams are initially expanded and collimated before propagating through individual diffractive optical elements (DOE), which splits each beam into two identical copies (angle between beams 0.4–0.6 degrees, corresponding to ~ 10 lp/mm). The beams are then made collinear by spatially overlapping in a recombination arm, consisting of a series of dichroic mirrors (Thorlabs DMLP505R, DMLP425R and BSW10R). Using a lens the beam copies are made to overlap at the sample plane, where they interfere and produce a spatial intensity modulation pattern, according to their wavelength and relative angle. This modulated light is what illuminates the sample and due to the linear response the fluorescence emissions maintain the same encoding. Note that a non-linear response, such as that produced with two-photon excitation, would still produce the necessary fundamental Fourier components, although the approach would require a different calibration procedure. The emitted intensity-modulated fluorescence then enters the second part of the setup, where the emission encoding occurs. Notch filters are used at the entrance to the emission encoding setup in order to block any stray excitation light. The fluorescence emission is split equally, using beam splitters (Thorlabs BS031), into four optical paths. Each path contains a spectral filter (chosen for these measurements to be Thorlabs filters: shortpass 525 nm, bandpass 550 ±25 nm, bandpass 600 ±25 nm and longpass 625 nm) followed by a transmission Ronchi grating (Edmund Optics 2″ × 2″ Ronchi Ruling, 20 lp/mm), such that each different spectral emission range will have unique spatial modulation encoding. The emission filters were chosen to be evenly distributed over the entire visible spectrum. All the emissions are then recombined using beam splitters, before being imaged on to an Andor Zyla 5.5 sCMOS camera equipped with a 1× Edmund Optics GOLD-TL telecentric lens.

The mathematics of a doubly modulated signal. FRAME, in any configuration, relies on applying a 2-dimensional intensity modulation, M, to an image signal:

$$M = \frac{1}{2}\left[\cos(\boldsymbol{\omega}\mathbf{r} + \phi) + 1\right], \tag{1}$$

where $\boldsymbol{\omega} = (\omega_x, \omega_y)$ corresponds to the spatial frequencies in the x and y direction, and $\mathbf{r} = (x, y)$ corresponds to the two spatial dimensions of an image. When the modulated laser profile is incident on the fluorophores, the resulting response I_f will consist of the underlying fluorophore response, A, superimposed with the laser modulation, M_l, such that:

$$I_f = AM_l = A\frac{1}{2}\left(\cos(\boldsymbol{\omega_l}\mathbf{r} + \phi_l) + 1\right). \tag{2}$$

$\boldsymbol{\omega_l}$ is the modulation frequency of the laser profile with an unknown phase ϕ_l. In DM-FRAME, this fluorescence response, I_f, travels through the imaging system, via spectral filters, to then be superimposed with another spatial modulation of frequency $\boldsymbol{\omega_e}$ and phase ϕ_e, resulting in a total signal of the form:

$$I = \tilde{A}M_lM_e = \tilde{A}\frac{1}{2}\left(\cos(\boldsymbol{\omega_l}\mathbf{r} + \phi_l) + 1\right) \times \frac{1}{2}\left(\cos(\boldsymbol{\omega_e}\mathbf{r} + \phi_e) + 1\right),$$

where $\tilde{A}$ is the spectrally filtered response. In the configuration presented in the manuscript, I corresponds to the final signal that is imaged on the sensor (e.g. Fig. 1a, c, e). Expanding the brackets and applying a product-to-sum trigonometric identity, I can be written as a sum of cosine terms:

$$I = \frac{1}{4}\tilde{A}\left(\cos\left[(\boldsymbol{\omega_l} + \boldsymbol{\omega_e})\mathbf{r} + \phi^+\right] + \cos\left[(\boldsymbol{\omega_l} - \boldsymbol{\omega_e})\mathbf{r} + \phi^-\right] + \cos(\boldsymbol{\omega_{lr}} + \phi_l) + \cos(\boldsymbol{\omega_{er}} + \phi_e)\right).$$

Hence the image will consist of the underlying fluorophore distribution not only modulated by the base frequencies applied by the excitation and emission, $\boldsymbol{\omega_l}$ and $\boldsymbol{\omega_e}$, but also their sum and difference: $(\boldsymbol{\omega_l} + \boldsymbol{\omega_e})$, $(\boldsymbol{\omega_l} - \boldsymbol{\omega_e})$. Performing spatial lock-in detection[21] on the emission or excitation frequency components (the $\boldsymbol{\omega_l}$ and $\boldsymbol{\omega_e}$ terms) results in images of the total excitation/emission profile while performing the lock-in detection on the frequency-sum components will reveal the fully optically traced information about the fluorophore distribution, i.e. which lasers cause the fluorophores to emit which signal. Examples of all the extracted images can for example be seen in the views of Fig. 2.

Linear unmixing. Linear unmixing is a mathematical approach[33,34], used in fluorescence imaging, for identifying the different fluorophores, as well as their relative quantities, present in a sample. Linear unmixing requires calibration, or reference spectra, for each of the different fluorophores that the sample may contain. These reference spectra contain the response of each fluorophore for a given laser excitation and spectral filter. The greater the difference in response, the greater the sensitivity becomes and it is therefore beneficial to select fluorophores that either have fluorescence in different spectral regions or that have large differences in their excitation spectra. These reference spectra are then used to solve a series of simultaneous equations, which represent the captured spectrum. This can be done using a least squares fitting method which aims to minimize the

square of the difference between the calculated and measured values. As a result the relative concentrations of each fluorophore in the mixed signal, can be determined.

Received: 22 June 2021; Accepted: 27 September 2021
Published online: 14 October 2021

References

1. Valm, A., Cohen, S., Legant, W., Melunis, J. & Hershberg, U. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167. https://doi.org/10.1038/nature22369 (2017).
2. Cutrale, F. *et al.* Hyperspectral phaser analysis enables multiplexed 5d in vivo imaging. *Nat. Methods* **14**, 149–152 (2017).
3. Cosco, E. D., Spearman, A. L., Ramakrishnan, S., Lingg, J. G. P. & Saccomano, M. Shortwave infrared polymethine fluorophores matched to excitation lasers enable non-invasive, multicolour in vivo imaging in real time. *Nat. Chem.* **12**, 1123–1130. https://doi.org/10.1038/s41557-020-00554-5 (2020).
4. Gomez, P. A., Garbacik, E. T., Otterstrom, J. J., Garcia-Parajo, M. F. & Lakadamyali, M. Excitation-multiplexed multicolor super-resolution imaging with fm-storm and fm-dna-paint. *Natl. Acad. Sci.* **115**, 12991–12996. https://doi.org/10.1073/pnas.1804725115 (2018).
5. Hu, F. *et al.* Supermultiplexed optical imaging and barcoding with engineered polyynes. *Nat. Methods* **15**, 194–200 (2018).
6. Wei, L. *et al.* Super-multiplex vibrational imaging. *Nature* **544**, 465–470 (2017).
7. Jahr, W., Schmid, B., Schmied, C., Fahrback, F. O. & Huisken, J. Hyperspectral light sheet microscopy. *Nat. Commun.* **6**, 7990. https://doi.org/10.1038/ncomms8990 (2015).
8. Schulz, C. & Sick, V. Tracer-lif diagnostics: Quantitative measurement of fuel concentration, temperature and fuel/air ratio in practical combustion systems. *Progr. Energy Combust. Sci.* **31**, 75–121 (2005).
9. Li, Y., Cu, Y. & Luo, D. Multiplexed detection of pathogen dna with dna-based fluorescence nanobarcodes. *Nat. Biotechnol.* **23**, 559–885 (2005).
10. Bearman, G. & Levenson, R. *Biological Photonics Handbook* (CRC Press, 2003).
11. Haugen, N. & Kudenov, M. Review of snapshot spectral imaging technologies. *Opt. Eng.* **52**, 090901. https://doi.org/10.1101/pdb.top071795 (2013).
12. Dorozynska, K. & Kristensson, E. Implementation of a multiplexed structured illumination method to achieve snapshot multi-spectral imaging. *Opt. Express* **25**, 17211–17226. https://doi.org/10.1364/OE.25.017211 (2017).
13. Dorozynska, K., Kornienko, V., Aldén, M. & Kristensson, E. A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Opt. Express* **28**, 9572–9586. https://doi.org/10.1364/OE.384535 (2020).
14. Deng, C. *et al.* Snapshot hyperspectral imaging via spectral basis multiplexing in Fourier domain. *Opt. Express* **26**, 32509–32521. https://doi.org/10.1364/OE.26.032509 (2018).
15. Garbacik, E. T., Sanz-Paz, M., Borgman, K. J., Campelo, F. & Garcia-Parajo, M. F. Frequency-encoded multicolor fluorescence imaging with single-photon-counting color-blind detection. *Biophys. J .* **115**, 725–736. https://doi.org/10.1016/j.bpj.2018.07.008 (2018).
16. Zhao, M., Li, Y. & Peng, L. Parallel excitation-emission multiplexed fluorescence lifetime confocal microscopy for live cell imaging. *Opt. Express* **22**, 10221–10232 (2014).
17. Sobhy, M. A. *et al.* Versatile single-molecule multi-color excitation and detection fluorescence setup for studying biomolecular dynamics. *Rev. Sci. Instrum.* **82**, 113702 (2011).
18. Mahou, P. *et al.* Multicolor two-photon tissue imaging by wavelength mixing. *Nature* **9**, 815–818. https://doi.org/10.1038/nmeth.2098 (2012).
19. Lavagnino, Z. *et al.* Snapshot hyperspectral light-sheet imaging of signal transduction in live pancreatic islets. *Biophys. J .* **111**, 409–417 (2016).
20. Valm, A. M., Oldenbourg, R. & Borisy, G. G. Multiplexed spectral imaging of 120 different fluorescent labels. *PLoS ONE* **11**, e0158495. https://doi.org/10.1371/journal.pone.0158495 (2016).
21. Kristensson, E. *et al.* Frame: Femtosecond videography for atomic and molecular dynamics. *Light Sci. Appl.* **6**, e17045 (2017).
22. Ek, S., Kornienko, V. & Kristensson, E. Long sequence single-exposure videography using spatially modulated illumination. *Sci. Rep.* **10**, 1–10 (2020).
23. Hensel, M., Parlow, T. & Grigat, R.-R. Modeling and real-time estimation of signal-dependent noise in quantum-limited imaging. In *Proceedings of 6th WSEAS International Conference on Signal Processing, Robotics Automation*, 183–191 (Corfu Island, Greece, 2007).
24. Cox, I. J. & Sheppard, C. J. R. Information capacity and resolution in an optical system. *J. Opt. Soc. Am. A* **3**, 1152–1158. https://doi.org/10.1364/JOSAA.3.001152 (1986).
25. Gustafsson, M. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microsc.* **198**, 82–87 (2000).
26. Crooks, J. *et al.* Kirana: A solid-state megapixel uCMOS image sensor for ultrahigh speed imaging. In Widenhorn, R. & Dupret, A. (eds.) *Sensors, Cameras, and Systems for Industrial and Scientific Applications XIV*, vol. 8659, 36–49. International Society for Optics and Photonics (SPIE, 2013).
27. Kornienko, V., Kristensson, E., Ehn, A., Fourriere, A. & Berrocal, E. Beyond mhz image recordings using leds and the frame concept. *Sci. Rep.* **10**, 1–10 (2020).
28. Van Munster, E. B. & Gadella, T. W. Fluorescence lifetime imaging microscopy (FLIM). *Microsc. Tech.* **8**, 8–9. https://doi.org/10.1007/b102213 (2005).
29. Ehn, A., Johansson, O., Arvidsson, A., Aldén, M. & Bood, J. Single-laser shot fluorescence lifetime imaging on the nanosecond timescale using a dual image and modeling evaluation algorithm. *Opt. Express* **20**, 3043–3056. https://doi.org/10.1364/OE.20.003043 (2012).
30. Keller, P. J. *et al.* Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. *Nat. Methods* **7**, 637–642 (2010).
31. Denk, W., Strickler, J. & Webb, W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76. https://doi.org/10.1126/science.2321027 (1990).
32. Kner, P., Chhun, B. B., Griffis, E. R., Winoto, L. & Gustafsson, M. G. L. Super-resolution video microscopy of live cells by structured illumination. *Nat. Methods* **6**, 339–342 (2009).
33. Zimmermann, T. Spectral imaging and linear unmixing in light microscopy. *Adv. Biochem. Eng. Biotechnol.* **95**, 245–265 (2005).
34. Zimmermann, T., Rietdorf, J. & Pepperkok, R. Spectral imaging and its applications in live cell microscopy. *FEBS Lett.* **546**, 87–92. https://doi.org/10.1016/S0014-5793(03)00521-0 (2003).

Acknowledgements
Swedish Research Council (2019-05183); European Research Council (803634).

Author contributions
E.K. invented the DM-FRAME concept. K.D., V.K., D.A and E.K constructed the optical setup and conducted the experiments. A.A provided assistance with the fluorescent markers. S.E. performed the linear unmixing data analysis. V.K. calculated the classification accuracy. K.D., S.E., V.K. and E.K. wrote the manuscript, with technical advice from A.E. and A.A. All authors have reviewed the manuscript.

Funding
Open access funding provided by Lund University.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-99670-6.

Correspondence and requests for materials should be addressed to E.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Le métier est entré.